ALMA MATER STUDIORUM – UNIVERSITÀ DI BOLOGNA

Department of Computer Science and Engineering (DISI)

Second Cycle Degree in Artificial Intelligence

Master's Thesis

Decoding Spatial Pain Anticipation from EEG using Deep Learning

Research carried out in collaboration with the Department of Psychology, Center for Studies and Research in Cognitive Neuroscience

Supervisor

Prof. Francesca Starita

Co-supervisors

Prof. Andrea Acquaviva
Dr. Francesco Barchi

Candidate

Bahar Hamzehei

Academic Year 2025/2026

Abstract

This thesis investigates whether electroencephalography (EEG) recorded while a participant waits for a possible painful stimulus contains information about where on the body that pain is expected. The study uses a Pavlovian threat-learning task in which visual cues signalled one of three meanings: expected pain on the left arm, expected pain on the right arm, or safety. The classification problem was therefore defined as a three-class distinction between left-threat, right-threat, and safety during the cue period.

The analysis focuses on Block 1, the acquisition phase of the experiment, because this is the phase in which cue meanings were stable. The Block 1 raw scan covered 30 recording files, while the final machine-learning export contained 29 retained recording identifiers. The final empirical claims therefore refer to the exported dataset of 29 retained recording identifiers and 1711 clean cue epochs, with almost balanced class counts: 571 left-threat epochs, 572 right-threat epochs, and 568 safety epochs.

Three model families were compared using a subject-wise train/validation/test split and train-only normalization: a baseline temporal convolutional network (TCN), an attention-augmented TCN, and a compact convolutional neural network trained on short-time Fourier transform (STFT) spectrograms. In the full-band analysis, the attention TCN obtained the strongest held-out performance, with 39.93% test accuracy and macro-F1 0.3867. This is above the balanced three-class chance level of 33.33%, but the margin is modest. In the exploratory analyses, the alpha-restricted attention TCN achieved the highest test accuracy overall, reaching 40.27% and macro-F1 0.3975.

Overall, the results suggest that cleaned cue-period EEG contains some information about the expected spatial content of pain, but that this information is weak and difficult to generalize across participants. The main contribution of the thesis is therefore twofold: it provides a cautious empirical decoding result and a transparent preprocessing and evaluation pipeline for spatially specific EEG decoding in a technically challenging dataset.

Declaration on the Use of Generative Artificial Intelligence

During the preparation of this thesis, generative artificial intelligence tools were used as writing-support tools. Their use was limited to improving grammar, clarity, structure, and readability; helping to reorganize explanatory passages; and supporting the review of methodological descriptions and code-related explanations.

All scientific decisions, data preprocessing choices, analyses, model implementation, interpretation of results, and final conclusions were made and verified by the author. No generative artificial intelligence tool was used to generate experimental data, alter results, fabricate references, or replace the author's scientific judgement. The author takes full responsibility for the accuracy, originality, and integrity of the thesis.

List of Abbreviations

AI	Artificial Intelligence
CS	Conditioned stimulus
CS+L	Cue associated with expected pain on the left arm
CS+R	Cue associated with expected pain on the right arm
CS−	Safe cue, never paired with pain
DC	Direct-current correction event annotated in the recording stream
EEG	Electroencephalography
EMG	Electromyography
FDI	First dorsal interosseous muscle
MEP	Motor-evoked potential
MNE	Open-source Python environment for EEG/MEG analysis
NPZ	NumPy compressed array archive
SCR	Skin conductance response
STFT	Short-time Fourier transform
TCN	Temporal convolutional network
TMS	Transcranial magnetic stimulation
US	Unconditioned stimulus (painful electrical stimulation)

Contents

List of Tables

List of Figures

Chapter 1

Introduction

1.1 Motivation

Pain anticipation is a useful case for studying how the brain prepares for events before they occur. Pain is not only a response to tissue stimulation after the fact; it is shaped by expectation, prior learning, attention, and action readiness [11, 19, 12, 3, 13]. When a cue predicts that pain may occur, the nervous system can prepare before sensory confirmation is available. This anticipatory state can influence subjective pain, defensive behaviour, attention, and physiological readiness [11, 19, 12, 7].

The present thesis focuses on a more specific question: whether anticipation contains spatial information about the expected bodily location of pain. In everyday situations, it matters not only that a harmful event may happen, but also where it is expected. Anticipating pain on the left arm is not the same as anticipating pain on the right arm, because attention and motor preparation can be organized around the threatened body site [5, 6, 21, 22, 3, 27]. For this reason, a decoding problem that distinguishes left-threat, right-threat, and safety is more demanding than a simple threat-versus-safety contrast.

This question is also methodologically relevant. EEG signals are high-dimensional, noisy, and strongly participant-specific. Apparent decoding performance can therefore be inflated if preprocessing is insufficiently transparent, if subjects appear in more than one data partition, or if normalization uses information from held-out data [8, 9, 10, 18]. For that reason, dataset construction, artifact handling, and leakage control are reported as integral parts of the analysis rather than as implementation details separate from the scientific claim.

1.2 Experimental task and analytical focus

The dataset comes from a Pavlovian threat-learning experiment. In the acquisition phase, participants saw visual cues that signalled one of three meanings: pain expected on the left arm, pain expected on the right arm, or safety. The two threat

cues were partially reinforced, meaning that the painful stimulation did not occur on every threat trial. The safe cue was never followed by pain. The EEG interval analysed in this thesis is the cue period, when the participant has seen the cue and can anticipate the possible outcome, but before the outcome itself occurs.

The broader experiment also included a reversal phase, in which previously learned cue meanings changed. Reversal is theoretically important because it involves updating and relearning. It is, however, a more complex target for a first decoding analysis, because cue-related EEG may then reflect current expectancy, residual learning from the previous mapping, uncertainty, or updating processes. For this reason, the main thesis analysis focuses on Block 1, the acquisition phase, where cue meanings are most stable and the labels are easiest to interpret.

A further practical issue is that the recordings contained DC-correction events. These are annotated events related to abrupt signal correction or instability in the recording stream. The main concern is not simply the event marker itself, but the possibility that the signal remains unstable for a short time afterwards. A cue beginning too soon after such an event may therefore be present in the annotation stream but unsuitable as a learning sample. The post-DC exclusion tail used in this thesis was introduced to remove such cue epochs conservatively.

1.3 Problem statement

The central question of the thesis is:

> Can EEG activity recorded during the cue period in Block 1 of a Pavlovian threat-learning experiment be used to decode the expected spatial content of pain, distinguishing left-threat, right-threat, and safety at the single-epoch level under a leakage-free subject-wise evaluation protocol?

This question is deliberately narrow. First, the analysis targets the cue period rather than the response to the painful stimulation itself. Second, it focuses on Block 1 acquisition rather than treating acquisition and reversal as interchangeable. Third, the evaluation is conservative: data are split by participant, normalization is estimated from training data only, and held-out test performance is used for the main interpretation. Fourth, the study compares time-domain and time-frequency representations without changing the cleaned source dataset.

In practical terms, the analysis required five connected steps: reconstructing cue-period samples from raw annotations; removing epochs invalidated by duration

anomalies or DC-related contamination; exporting fixed-size time-domain and STFT representations; enforcing a subject-wise split with train-only normalization; and interpreting the resulting performance in light of the modest sample size and the subtle nature of the target.

1.4 Main contributions

The thesis makes four main contributions.

First, it builds a traceable preprocessing and export pipeline for Block 1 cue-period EEG. Each retained epoch is linked back to subject, recording, block, timing, label, and annotation context.

Second, it treats DC-related contamination explicitly. Rather than assuming that the annotation boundary fully captures signal usability, the analysis estimates a post-DC exclusion tail from signal recovery and applies it in an order-aware cleaning procedure.

Third, it evaluates models under a grouped subject-wise protocol with train-only normalization, validation-based model selection, and a held-out test set with no subject or recording overlap.

Fourth, it compares several modelling choices: a baseline TCN, an attention-augmented TCN, a spectrogram CNN, and exploratory alpha-only, theta-only, and C3/C4-only variants. The comparison is interpreted cautiously, with emphasis on whether the result survives subject-wise evaluation rather than on maximizing a single headline score.

1.5 Structure of the thesis

Chapter 2 reviews the theoretical and methodological background: pain anticipation, Pavlovian threat learning, spatial bodily anticipation, alpha and theta activity, central-channel restrictions, and EEG decoding. Chapter 3 describes the experimental paradigm, the dataset, and the analytical scope. Chapter 4 presents the dataset construction and cleaning procedure. Chapter 5 describes the input representations, models, training procedure, and evaluation protocol. Chapter 6 reports the results. Chapter 7 discusses the implications and limitations. Chapter 8 concludes. The appendices provide reproducibility notes, hyperparameter grids, selected configurations, supplementary diagnostic figures, and compact result tables.

Chapter 2

Theoretical Background and Related Work

2.1 Pain anticipation as predictive and embodied processing

Pain is increasingly understood as a process shaped by expectation, context, attention, and action relevance rather than as a passive readout of tissue status [11, 19, 12, 3, 13]. Expectations can influence pain before sensory confirmation is available, and anticipatory states can change both subjective experience and physiological readiness [11, 19, 12, 7]. This makes the cue period a meaningful target for analysis: it is the interval in which the participant has information about possible harm but has not yet received the outcome.

Predictive accounts of pain are important for the present thesis because they treat anticipation as part of the pain-related process rather than as irrelevant pre-stimulus activity. If a cue has acquired threat value, the brain may already organize attention, preparation, and expectation around the predicted event. In this setting, decoding cue-period EEG is not an attempt to classify a large sensory response. It is an attempt to test whether a learned anticipatory state contains recoverable information at the single-epoch level.

The present work also adopts an embodied perspective. The target is not only whether pain is expected, but where it is expected. Bodily location has behavioural consequences: preparing for pain to the left arm and preparing for pain to the right arm may involve different attentional and motor states. If these states are reflected in EEG, then a classifier may be able to distinguish left-threat from right-threat, even though the two classes share the general meaning of threat.

2.2 Pavlovian threat learning, acquisition, and reversal

Pavlovian threat learning provides a controlled way to study anticipation. Initially neutral cues acquire predictive meaning through association with an aversive outcome, and once learning has occurred, conditioned stimuli can evoke neural, behavioural, and physiological responses before the outcome occurs [14, 15, 20, 28]. In the present task, two threat cues predicted pain on different arms and one cue predicted safety. Because the threat cues were partially reinforced, the cue signalled expected pain location even though pain was not delivered on every threat trial.

It is useful to distinguish acquisition from reversal. Acquisition is the phase in which the cue-outcome mapping is learned and remains stable. Reversal introduces a different problem: the participant must update or inhibit a previously learned mapping and adapt to changed contingencies [16, 17, 20]. Reversal can therefore involve uncertainty, prediction error, and cognitive control in addition to current expectancy [15, 16, 17]. These processes are scientifically important, but they make the label interpretation less direct.

For that reason, the present thesis uses Block 1 acquisition as the main empirical scope. This does not mean that reversal is unimportant. Rather, it means that acquisition provides the cleaner setting for a first subject-wise EEG decoding analysis of spatial pain anticipation. If the spatial content of anticipation is difficult to recover even when cue meanings are stable, then reversal should be treated as a separate and more complex problem rather than merged uncritically into the main analysis.

2.3 Spatially specific threat representation and bodily anticipation

Evidence from experimental pain research supports the idea that threat anticipation can be spatially anchored to the body. Anticipating pain at a particular body location can bias tactile processing toward the threatened site, suggesting that attention and readiness are organized around the expected locus of harm [21, 22]. Importantly, this spatial organization does not need to be limited to a single point on the skin. It may operate at a functional body-part or body-side scale [22].

This literature motivates the three-class target used here. A binary threat-versus-safety classifier could perform above chance while ignoring bodily location. The left-threat/right-threat/safety formulation is more demanding because it asks whether

the anticipatory state contains information that separates the two threatened sides while also distinguishing both from safety. The task is therefore aligned with the scientific question rather than with a simpler but less informative classification problem.

2.4 Alpha-band dynamics and somatomotor preparation

Somatomotor alpha activity is especially relevant to spatial anticipation. Alpha-band activity over sensorimotor cortex has been linked to anticipatory regulation, selective bodily preparation, and pain expectation [4, 5, 6]. More generally, alpha modulation has been interpreted as a gating mechanism that changes regional engagement or inhibition depending on task demands [23, 27]. In pain-related contexts, this makes alpha a plausible candidate for carrying information about the expected side of stimulation.

The alpha-restricted analysis in this thesis was therefore not an arbitrary exploratory filter. It tested whether a frequency range with a plausible role in somatomotor preparation would improve or clarify decoding of left-threat, right-threat, and safety. A stronger alpha result would be consistent with the idea that the spatial component of anticipation is partly expressed through sensorimotor preparatory dynamics. At the same time, any such result must remain cautious, because the analysis is still based on classification performance rather than direct physiological localization.

2.5 Theta-band dynamics, threat processing, and updating

Frontocentral theta is also relevant, but likely for a different reason. Midfrontal theta has often been associated with conflict monitoring, uncertainty, control allocation, and adaptive updating [24]. In threat-learning paradigms, these functions are important when cue meanings are being learned, challenged, or revised. Theta may therefore support classification by distinguishing threatening from safe anticipatory states or by indexing control-related engagement, even if it is less directly tied to bodily laterality than alpha.

The theta-restricted analysis was included for this reason. It provides a comparison

between a frequency range expected to be more related to general threat and control processes, and a frequency range more closely linked to somatomotor preparation. The comparison is not intended to prove a neural mechanism by itself. Instead, it helps interpret whether the strongest decoding results are more consistent with a body-specific preparatory account or with a broader threat-processing account.

2.6 Rationale for the C3/C4 exploratory analysis

The C3/C4-restricted analysis was motivated by the spatial-motor nature of the hypothesis. C3 and C4 are central electrodes commonly associated with activity over left and right sensorimotor regions. If the relevant information were concentrated mainly in lateralized sensorimotor activity, then a reduced analysis using these two channels might preserve useful information while removing unrelated signals.

This restriction is intentionally conservative and exploratory. It is not a claim that C3 and C4 fully capture the neural sources of pain anticipation. Scalp EEG is spatially blurred, and anticipatory states may involve distributed networks rather than isolated electrodes. The C3/C4 analysis therefore serves as a stress test for the spatial-motor account: it asks whether the problem can be solved from a very narrow central-channel subset, or whether broader scalp information remains useful.

2.7 EEG decoding and leakage-controlled evaluation

Deep learning has become common in EEG analysis, but the literature repeatedly stresses that EEG classification is vulnerable to leakage, unstable validation, and overoptimistic interpretation [8, 9, 10, 18]. This is especially important when samples from the same participant are numerous. If train and test partitions share participants, a model may partly learn participant-specific structure rather than the phenomenon of interest.

For this reason, subject-wise evaluation is central to the present thesis. The aim is not to determine whether a model can recognize patterns from participants it has already seen, but whether cue-related information generalizes to held-out participants. Train-only normalization is equally important: distributional information from validation or test data must not influence the scaling parameters used during development. These safeguards make the decoding problem harder, but they make the interpretation more trustworthy.

Pain-related decoding studies show that pre-stimulus or anticipatory brain activity can contain behaviourally relevant information [25, 26]. The present thesis extends that logic to a different target: the spatial content of pain anticipation in a three-class cue-period EEG problem. The expected effect is subtle, so the interpretation of above-chance performance must remain modest and must be supported by transparent preprocessing and evaluation.

Chapter 3

Experimental Paradigm, Dataset, and Analytical Scope

3.1 Overview of the experiment

The experiment was designed to study how learned threat cues shape anticipation of limb-specific pain. Participants completed a Pavlovian threat-learning task while EEG was recorded. The broader study also included TMS-induced motor-evoked potentials and skin conductance responses, but the present thesis analyses only EEG during the cue period.

During each trial, a visual cue indicated one of three conditions. One cue predicted possible pain on the left arm, one cue predicted possible pain on the right arm, and one cue indicated safety. The two threat cues were reinforced on 70% of trials, whereas the safety cue was never reinforced. The cue period lasted 4.500 s; this interval is treated as the anticipation window in the thesis. Each block contained 20 trials per cue type, for 60 trials per block.

The broader task included both acquisition and reversal. In Block 1 acquisition, the cue meanings were stable. In Block 2 reversal, cue meanings changed and participants had to update learned associations. Because the goal of this thesis is to establish a first controlled decoding analysis of spatial pain anticipation, the main empirical analysis is restricted to Block 1 acquisition.

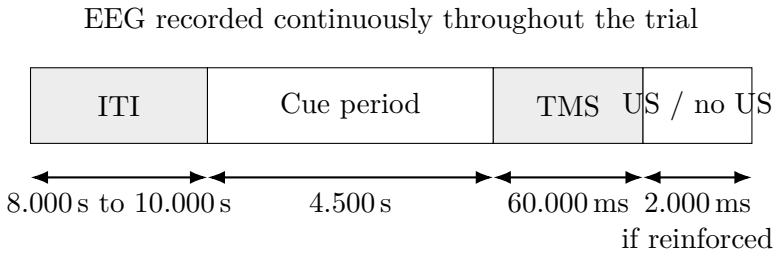

Figure 3.1: Simplified trial structure of one experimental trial. The full cue period was treated as the anticipation window of interest in the present thesis. TMS and outcome events occur after the cue interval and are not part of the classifier input.

Table 3.1: Summary of the experimental design relevant to the present thesis.

Component	Description
Initial recording set	30 raw Block 1 recording files were scanned; the final machine-learning export contained 29 retained recording identifiers
Blocks	Block 1 acquisition; Block 2 reversal
Cue classes	CS+L, CS+R, CS−
Trials per block	20 per cue type; 60 total per block
Threat reinforcement	70% for CS+L and CS+R; 0% for CS−
Cue duration	4.500 s
Inter-trial interval	8.000 s to 10.000 s
Present thesis	Block 1 cue-period EEG only; three-class decoding of left-threat, right-threat, and safety cues

3.2 Operational scope of the study

The analysis began from the Block 1 recordings and reconstructed ordered annotation-to-annotation segments. The Block 1 raw scan covered 30 recording files, while the final machine-learning export contained 29 retained recording identifiers. The final empirical claims therefore refer to this exported dataset and its 1711 clean cue epochs. From the scanned recordings, 3654 ordered segments were reconstructed. Among them, 1740 were task-candidate cue epochs before contamination-aware cleaning. After the cleaning rules were applied, 1711 cue epochs remained and 29 were removed.

The final machine-learning dataset was almost perfectly balanced across the three target classes: 571 left-threat epochs, 572 right-threat epochs, and 568 safety epochs. This balance is important because it means that the later model comparison is not dominated by strong class imbalance.

Table 3.2: Dataset progression for the final Block 1 analysis.

Stage	Count
Raw Block 1 recording files scanned	30
Retained recording identifiers in the final learning export	29
Ordered segments reconstructed	3654
Task-candidate cue epochs	1740
Removed after cleaning	29
Retained clean task epochs	1711
Class 0 (left threat)	571
Class 1 (right threat)	572
Class 2 (safety)	568

3.3 Label space and interpretation of the classes

The original stimulus event codes were mapped to the three-class target space shown in Table 3.3. Codes corresponding to real and non-reinforced left-arm cues were collapsed into the same left-threat class; the same was done for right-arm cues. This is appropriate because the classifier is intended to decode the anticipatory meaning of the cue, not whether pain was ultimately delivered on that specific trial.

Table 3.3: Mapping from stimulus event codes to fine labels and final classes.

Event code	Fine label in implementation	Final class
S11	`stim_left_real`	Left-threat (class 0)
S110	`stim_left_fake`	Left-threat (class 0)
S12	`stim_right_real`	Right-threat (class 1)
S120	`stim_right_fake`	Right-threat (class 1)
S13	`stim_no`	Safety (class 2)

Throughout the thesis, the classes are referred to as left-threat, right-threat, and safety. Code-level labels are retained only where exact implementation details matter.

3.4 Analytical branches and scope

The main empirical branch is the full-band Block 1 analysis. Three exploratory Block 1 branches are also reported: alpha-only, theta-only, and C3/C4-only. These variants change the input restriction while keeping the general cleaning and evaluation

logic fixed. They are used to interpret whether decoding is stronger when the signal is restricted to theoretically relevant frequency bands or central electrodes.

A jointly trained Block 1+Block 2 analysis is reported only as supplementary context and is not treated as a main empirical result. Reversal changes the interpretation of the labels and introduces updating-related processes. For that reason, the thesis keeps the main conclusions anchored to the Block 1 acquisition export.

Table 3.4: Analytical branches retained in the thesis and their role.

Analysis	Status	Role
Full-band Block 1	Main analysis	Principal test of three-class cue-period decoding from cleaned EEG
Alpha-only Block 1	Exploratory	Tests whether alpha-band input is especially informative for spatial pain anticipation
Theta-only Block 1	Exploratory	Tests whether theta-band input carries broader threat/control-related information
C3/C4-only Block 1	Exploratory	Tests whether a narrow central-channel subset is sufficient for the spatial-motor hypothesis
Block 1+Block 2	Supplementary only	Not used for main conclusions because reversal changes the interpretation of the labels

3.5 Configuration choices and analytical scope

Several thresholds and settings were determined before model interpretation. The cue-epoch length was design-fixed at 4.500 s. The stimulus-duration tolerance was estimated from the empirical distribution of cue durations. The post-DC exclusion tail was estimated from signal recovery around DC-correction events. In the executed confirmatory Block 1 branch, the applied post-DC exclusion tail was 1.827 s. A closely related combined-block estimate of 1.836 s existed for the supplementary two-block branch, but it does not govern the main Block 1 results.

This distinction is important because the thesis reports the value that actually governed each executed branch. The main empirical claims use the Block 1 export, the Block 1 cleaning configuration, and the Block 1 subject-wise split.

Chapter 4

Methodology: Segmentation, Preprocessing, and Dataset Construction

Dataset construction was organized around a specific validity question: whether each annotated cue could be used as a complete and uncontaminated cue-period epoch. A DC-correction event is an annotated event in the recording stream associated with abrupt signal correction or instability. Because signal recovery may continue after the marker itself, a cue beginning immediately after such an event may be present in the annotation stream while still being unsuitable for learning. The cleaning procedure therefore uses whole-epoch rejection for two DC-related cases: cues interrupted by a DC-correction event and cues beginning within the estimated post-DC exclusion tail.

4.1 Parameter choices and justification types

Table 4.1 summarizes how the main preprocessing and modelling parameters were justified. Some values were fixed by the task design, some were estimated from the data analysed here, some were inherited from the executed configuration, and some were retained as conservative implementation defaults.

Table 4.1: Classification of key methodological parameters by type of justification.

Parameter or decision	Type of justification	Rationale
Cue-epoch target duration = 4.500 s	Design-fixed	Directly inherited from cue duration in the task
Restriction to Block 1	Scope decision	Final exported dataset and split summaries all correspond to Block 1
Stimulus tolerance = 0.300 s	Primary data-driven	Estimated from the observed distribution of cue-duration deviations in the dataset-audit analysis
Post-DC exclusion tail	Signal-based estimate used in the executed branches	The confirmatory Block 1 branch uses 1.827 s; the supplementary jointly trained Block 1+Block 2 branch uses 1.836 s
Auxiliary DC minimum gap = 0.050 s	Conservative default	Used as a small guard threshold in the state-machine logic
Band-pass and resampling	Retained preprocessing choices	Stable signal extraction at 250.000 Hz with fixed band limits per run
STFT configuration	Retained but justified engineering choice	Read from the active configuration and held fixed across the relevant comparison
Search spaces for model hyperparameters	Conservative bounded search	Chosen for interpretable comparison under limited subject count and finite compute

4.2 Raw BrainVision loading and header-sidecar repair

Each `.vhdr` file was first read with `mne.io.read_raw_brainvision`. When this initial read failed, the header was repaired by parsing the `DataFile=` and `MarkerFile=` references, searching for plausible sidecar locations relative to the recording directory, the configured raw directory, and the raw-data parent directory, and then writing a repaired copy into `_vhdr_fixed` before repeating the read. This step is methodologically important because it shows how data loading remained robust to broken header paths without silently discarding files.

The loading stage was therefore handled explicitly rather than assumed to be trivial. Even though no files were ultimately skipped in the main Block 1 analysis, header repair remained part of the implemented method and is reported for reproducibility.

4.3 Signal-based estimation of DC recovery

The DC-tail recovery analysis estimates signal recovery directly from the data surrounding each detected DC marker. For each event, it extracts a fixed pre/post window around the DC onset, filters the data in the same general passband used by the main preprocessing pipeline, computes a smoothed RMS-based recovery trace, and defines recovery according to the point at which the post-DC signal falls back below a threshold derived from a prestimulus baseline period. The Block 1 recovery analysis used the following settings:

- band = 0.5–100 Hz,
- target resampling frequency = 250.000 Hz,
- pre-window = 5.000 s,
- post-window = 12.000 s,
- baseline window = $[-4.5, -0.5]$ s relative to the DC event,
- baseline percentile = 95,
- stability requirement = 0.500 s,
- RMS smoothing window = 0.050 s,
- safety margin added after the percentile summary = 0.500 s.

This stage yielded a recovery-event table, a failed-files table, a scan-summary JSON file, a valid-events CSV, a percentile-summary CSV, and four diagnostic plots: a recovery-time histogram, a boxplot, a marker-gap-versus-recovery scatter plot, and a median-RMS visualization. Together, these materials document how the recovery estimates were obtained, even though the raster images themselves are not embedded in the notebook JSON.

Table 4.2: Block 1 signal-based DC recovery summary.

Quantity	Value
VHDR files found	30
Files processed	30
Files with DC markers	29
Files without DC markers	1
Failed files	0
Total DC markers	144
Valid recovery estimates	129
Skipped: insufficient pre-window	14
Skipped: insufficient post-window	1
Median recovery	0.988 s
95th percentile recovery	1.327 s
Recommended Block 1 tail	1.827 s

Table 4.3: Comparison of the post-DC tail recommendations and the value ultimately used in the final modelling analyses.

Source	DC markers	Valid recoveries	Recommended tail
Block 1 DC-recovery analysis	144	129	1.827 s
Combined-block DC-recovery analysis	310	273	1.836 s
Confirmatory Block 1 full-band branch	—	—	1.827 s
Supplementary Block 1+Block 2 branch	—	—	1.836 s

4.4 Boundary-based segmentation and segment ontology

The segmentation logic is boundary-based and order-aware. Rather than taking a fixed-length crop after every stimulus onset regardless of what follows, the segmentation reconstructs segments from one relevant annotation boundary to the next. This is crucial because the contamination rules depend on temporal order. A stimulus can become invalid because it is interrupted by a DC event within its own interval, or because it begins too soon after a previous DC event even when its nominal duration looks normal.

Let the ordered relevant annotations in a recording be $\{(\tau_i, d_i)\}_{i=1}^{M}$, where τ_i is onset time and d_i is annotation description. The segmentation pass constructs ordered intervals

$$g_i = [\tau_i, \tau_{i+1}), \tag{4.1}$$

after filtering the annotation stream to the labels that matter for the final rules: task stimuli, TMS markers, DC-correction markers, and explicit new-segment boundaries. In this way, temporal context is maintained while only a subset of intervals becomes eligible for the learning dataset.

Each reconstructed interval is stored with metadata including subject, block identifier, recording identifier, source file path, within-recording order, raw annotation description, previous and next annotation descriptions, start time, end time, duration, fine label, task-candidate status, and later good/bad status. This metadata makes every retained or rejected sample auditable after model training.

4.5 Cue duration, duration tolerance, and whole-epoch rejection

The design-fixed cue period lasted 4.500 s. This value defines the intended temporal extent of a valid anticipation sample and was inherited from the task design rather than tuned on decoding performance. The analysis aims to classify the anticipatory state induced by the full cue, so incomplete cue intervals were not treated as shorter valid samples.

A task segment was accepted only if its duration remained close enough to the design value to count as a plausible realization of the full cue period. The selected tolerance was 0.300 s, giving an accepted range of 4.200 s to 4.800 s. This threshold

was estimated from the dataset-audit analysis. Across 1740 candidate cue segments, 1725 plausible durations entered the deviation summary; the 95th percentile of absolute deviation from 4.500 s was 0.242 s, and the 99th percentile was 0.254 s. The final 0.300 s threshold therefore allowed small annotation imprecision while excluding clearly abnormal cue windows.

Table 4.4: Empirical basis of the stimulus-duration tolerance.

Quantity	Value
Task candidates examined	1740
Plausible durations used for estimation	1725
95th percentile of absolute deviation	0.242 s
99th percentile of absolute deviation	0.254 s
Selected tolerance	0.300 s
Accepted cue-duration range	4.200 s to 4.800 s

The same logic determined the unit of rejection. If a cue was interrupted before the full anticipation interval had elapsed, the entire cue epoch was rejected; no pre-interruption fragment was retained as a separate learning sample. The same whole-epoch rule was applied when a cue began within the post-DC recovery window. After reconstruction, task segments were therefore checked for the following intrinsic anomalies before later state-machine contamination rules were applied:

1. a task cue shorter than the accepted range was marked invalid;
2. if the short segment ended because the next boundary was a DC-correction event, the reason was recorded as `stim_interrupted_by_dc`;
3. if the next boundary was a TMS marker, the reason was recorded as `stim_interrupted_before_tms`;
4. a task cue longer than the accepted range was marked `stim_too_long`; and
5. TMS segments themselves were never treated as machine-learning samples.

Formally, the admissible duration interval was

$$\begin{aligned} \mathrm{EPOCH}_{\mathrm{MIN}} &= \mathrm{EPOCH}_{\mathrm{LEN}} - \mathrm{STIM}_{\mathrm{TOL}}, \\ \mathrm{EPOCH}_{\mathrm{MAX}} &= \mathrm{EPOCH}_{\mathrm{LEN}} + \mathrm{STIM}_{\mathrm{TOL}}. \end{aligned} \tag{4.2}$$

For the present study, $\mathrm{EPOCH}_{\mathrm{LEN}} = 4.500\,\mathrm{s}$ and $\mathrm{STIM}_{\mathrm{TOL}} = 0.300\,\mathrm{s}$, so any task segment outside [4.200 s, 4.800 s] was intrinsically invalid.

4.6 Signal-based post-DC exclusion tail

The post-DC exclusion tail determines how much time after a DC-correction marker should still be treated as potentially contaminated, even if the next cue has already begun. Two closely related signal-based estimates were obtained, but only the Block 1 value governed the main empirical analysis.

The Block 1-specific DC-recovery analysis found 144 DC markers, 129 of which yielded valid recovery estimates. Its recovery-time distribution had a 95th percentile of 1.327 s, and with the retained 0.500 s safety margin it recommended 1.827 s. The combined-block recovery analysis found 310 DC markers, 273 valid recovery estimates, a 95th percentile of 1.336 s, and a recommended value of 1.836 s. In the confirmatory Block 1 full-band branch, the applied value was 1.827 s. In the supplementary jointly trained Block 1+Block 2 branch, the applied value was 1.836 s. The threshold discussion therefore distinguishes explicitly between the confirmatory Block 1 branch and the supplementary two-block branch.

Table 4.5: Comparison of the two signal-based DC-tail estimates carried into the modelling analyses.

Quantity	Block 1 analysis	Block 1+2 analysis
Total DC markers found	144	310
Valid DC recovery events	129	273
Median recovery	0.988 s	0.992 s
95th percentile recovery	1.327 s	1.336 s
Safety margin added	0.500 s	0.500 s
Recommended `DC_TAIL_S`	1.827 s	1.836 s
Used in confirmatory Block 1 branch	Yes	No
Used in supplementary Block 1+Block 2 branch	No	Yes

The two estimates are close in magnitude, supporting the stability of the recovery-based logic. The value 1.827 s remains the threshold that governed the Block 1 classification analyses.

4.7 Illustrative clean-versus-contaminated cue examples

The cleaning strategy was also checked visually. Each principal modelling analysis included a dedicated visualization step that rebuilt one retained clean cue epoch and

one rejected cue epoch directly from the segmentation state. This made it possible to inspect how the cleaning rules operated on concrete signal examples rather than only on summary counts.

The selection logic is explicit. The helper function `_choose_examples()` first scans the ordered segment lists and searches for a contaminated task-stimulus epoch rejected for `stim_interrupted_by_dc`. If none is available, it falls back to the first available `dc_within_tail` example. It then tries to find a clean epoch with the same fine label as the contaminated example so that the visual comparison is not confounded by class identity. Only if no label-matched clean epoch exists does it fall back to any retained clean task epoch. In other words, the figure selection is neither random nor retrospectively hand-picked. It follows a fixed priority rule that prefers the contamination mechanism most directly tied to whole-epoch interruption and then attempts to keep the task label constant across the clean-versus-contaminated comparison.

Once an example is chosen, the original BrainVision file is reopened, the full 4.500 s stimulus window is extracted from the raw signal, the data are resampled to the thesis sampling rate, and a small set of representative channels is plotted. The channel-selection helper uses a fixed preference list — `Fz`, `Cz`, `Pz`, `Oz`, `Fp1`, `Fp2`, `C3`, `C4`, `P3`, and `P4` — and keeps up to six channels in the full-channel analyses. The C3/C4-only analysis reuses the same visualization machinery but limits the retained signal itself to those two channels. Dashed vertical lines are added for DC markers that occur inside the plotted stimulus window, and the summary CSV written next to the figure stores the example type, subject, block, recording identifier, file path, within-recording order, fine label, rejection reason, epoch duration, number of DC markers inside the displayed window, and the delay from the previous DC marker when available.

Each modelling analysis therefore produced a representative clean-versus-contaminated example figure together with a compact summary table describing the selected epochs. These figures serve as quality-control evidence rather than incidental debugging output. They show that the rejected cue epochs are not simply unusual durations, but concrete examples of cue windows that are interrupted by DC or that begin before post-DC recovery has elapsed.

4.8 Order-aware state machine for contamination handling

Contamination handling is implemented through a state machine over the ordered segment sequence. A task epoch is invalid not only when a DC event occurs inside it, but also when it begins before the signal has recovered from a preceding DC event.

Let t_i^{start} be the onset time of candidate stimulus epoch i, and let $t_{\text{last DC}}$ be the onset of the most recent DC event preceding that epoch in the ordered stream. The state machine rejects the entire stimulus epoch if

$$0 \leq t_i^{\text{start}} - t_{\text{last DC}} < \text{DC\_TAIL}. \tag{4.3}$$

In the confirmatory Block 1 branch documented here, DC_TAIL = 1.827 s; in the supplementary jointly trained Block 1+Block 2 branch, DC_TAIL = 1.836 s. In both cases, the rejection reason is stored as `dc_within_tail`. This rule makes the cleaning decision depend on temporal proximity to DC-correction events rather than on annotation boundaries alone.

A second whole-epoch rule concerns interruption. If a task stimulus begins and a DC marker occurs before the cue interval is complete, the resulting boundary-to-boundary task segment is too short. In that case, the reason `stim_interrupted_by_dc` invalidates the entire cue epoch. Again, no part of the interrupted interval is salvaged as a separate learning sample.

4.9 Auxiliary DC minimum gap (`DC_MIN_S`)

A smaller parameter appearing in the code is `DC_MIN_S = 0.050`. Unlike `STIM_TOL_S` and `DC_TAIL_S`, this threshold is not one of the primary data-driven scientific results of the pipeline. It is better described as a conservative auxiliary default. A threshold of 0.050 s is small enough to catch clearly degenerate marker gaps without materially determining which normal cue epochs survive dataset construction.

4.10 Filtering and resampling

After a task epoch passed the segmentation and contamination checks, the signal extraction stage applied a Butterworth band-pass filter and resampled the retained segment to a common sampling rate of 250.000 Hz. In the main full-band and C3/C4

runs, the passband was 0.5–100 Hz with filter order 4. In the alpha-only and theta-only variants, the same extraction logic was reused but the passband was restricted to 8–12 Hz and 4–8 Hz respectively.

These filter settings are best described as retained preprocessing choices rather than newly optimized thresholds. Their role was to provide stable and comparable extraction while preserving the frequency content most relevant to the present question and attenuating very slow drift and higher-frequency noise that were not central to the analysis [1, 2]. Resampling to 250.000 Hz harmonized all retained epochs to a common temporal resolution and yielded a fixed cue length of 1125 samples for a 4.500 s epoch.

The implementation used reproducible filtering and resampling functions rather than an unspecified generic operation. Filtering was performed on the segment data after extraction from the raw object. The helper function `_safe_sos_filter` constructed a fourth-order Butterworth band-pass in second-order-sections form. When the segment was long enough, the code used `sosfiltfilt` to obtain a zero-phase result; otherwise it fell back to `sosfilt`. Resampling was handled by `_resample_exact`: if the ratio between the input and target sampling rates was an integer decimation factor, the code used `signal.decimate(..., zero_phase=True)`; otherwise it used `signal.resample_poly`.

4.11 Extraction of fixed-size time-domain tensors

Let $x_c(t)$ denote the signal from channel c. After filtering and resampling, the preprocessing stage padded or truncated each retained epoch to a fixed length corresponding to 4.500 s at 250.000 Hz, namely 1125 time points. The code therefore exported a time-domain tensor

$$\mathbf{X}^{\text{time}} \in \mathbb{R}^{N \times C \times T}, \tag{4.4}$$

where $N = 1711$, $T = 1125$, and $C = 64$ in the full-band, alpha-only, and theta-only runs, while $C = 2$ in the C3/C4-only run.

4.12 STFT representation and retained settings

The time–frequency dataset is derived from exactly the same cleaned epochs. The STFT configuration loaded by the main Block 1 analysis was:

- Hann window,
- `nperseg` corresponding to 3.000 s, that is 750 samples at 250.000 Hz,
- overlap corresponding to 2.800 s, that is 700 samples,
- `nfft` equal to four times `nperseg`, namely 3000,
- frequency truncation at 40.000 Hz.

These values were held fixed within the relevant comparison. In the main analysis, the STFT tensor had shape $(1711, 64, 481, 8)$, with 481 frequency bins spanning 0–40 Hz and eight time bins whose saved coordinates ranged from approximately 1.500 s to 2.900 s. In the C3/C4-only restriction, the corresponding shape was $(1711, 2, 481, 8)$.

4.13 Exported outputs and provenance records

Beyond the model-ready arrays, the study produced a set of outputs that can be grouped into six categories:

1. cleaned time-domain and time-frequency datasets for downstream modelling;
2. dataset-construction summaries describing how many segments were scanned, rejected, and retained;
3. per-recording metadata tables carrying subject, recording, order, label, timing, and quality-control information;
4. split manifests and normalization statistics derived from the training partition only;
5. search summaries, selected-model reports, and checkpoint metadata for each model family; and
6. diagnostic figures such as class-balance plots, learning curves, compact validation-versus-test summaries, and confusion matrices.

Taken together, these outputs make it possible to relate the final classification results back to the dataset-construction process that produced them. A retained sample can still be traced to its recording context, and the summary tables reported in the thesis remain linked to records generated directly during the study.

4.14 Diagnostic reports and figure inventory

In addition to the final arrays, the study also generated a substantial set of diagnostic reports. For the full-band analysis, and again for the alpha, theta, and C3/C4 variants, this included split-distribution summaries, class-balance diagnostics, hyperparameter-search tables, compact held-out summary plots, train/validation learning curves, learning-rate traces, confusion matrices, and plain-text classification reports for the selected models.

These materials provide a visual record of model behaviour beyond the headline result tables. The main comparisons are discussed in the body of the thesis, while additional supplementary figures are gathered in Appendix D. Table 4.6 summarizes the main kinds of material generated during the study.

Table 4.6: Main categories of analytical output used throughout the thesis.

Output category	Typical contents	Use in the thesis
DC-recovery diagnostics	Recovery-time distributions, marker-gap summaries, and threshold notes	Motivate and justify the post-DC exclusion tail
Dataset-audit summaries	Duration histograms, removal summaries, and per-recording counts	Support the cleaned-dataset audit and the stimulus-tolerance discussion
Illustrative quality-control figures	Representative clean and contaminated cue examples	Show how the contamination rules map onto concrete signal traces
Split and balance reports	Partition counts, class distributions, and sampler summaries	Document the grouped split and the near-balanced class structure
Training diagnostics	Loss, accuracy, macro-F1, learning-rate, and confusion-matrix figures	Describe the optimization behaviour of the selected models
Compact result summaries	Validation-versus-test comparison plots and summary tables	Provide a concise overview of setting-specific performance

4.15 Result of the cleaning stage

The Block 1 dataset-audit analysis showed that, after all rules were applied, 29 task epochs were removed. Importantly, these removals were concentrated in exactly the two DC-related categories shown in Table 4.7.

Table 4.7: Task-epoch removals by rule in the executed Block 1 dataset-audit analysis.

Removal reason	Count
`stim_interrupted_by_dc`	16
`dc_within_tail`	13
Total removed task epochs	29

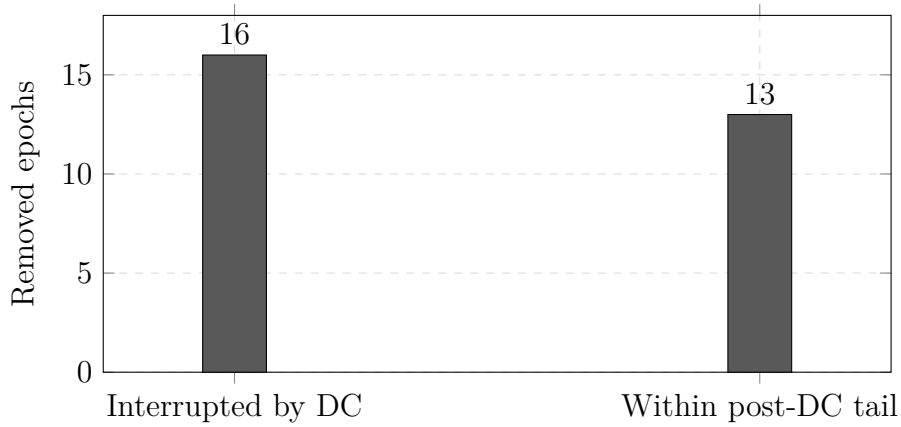

Figure 4.1: Removed task epochs by rule. The removals were concentrated in the exact DC-related mechanisms that motivated the contamination-control strategy.

Because the removals were targeted rather than diffuse, the cleaning stage can be interpreted as conceptually specific rather than broadly punitive.

Chapter 5

Representations, Models, Training Procedure, and Evaluation Protocol

5.1 Time-domain and time–frequency inputs

The main Block 1 run exported two matched representations from the same cleaned source data.

Time domain. The raw model input is the multichannel signal tensor of shape (N, C, T). In the main run, $N = 1711$, $C = 64$, and $T = 1125$. This representation preserves waveform structure directly and leaves feature extraction to the neural network.

Time–frequency domain. The STFT representation encodes each epoch as a channel-by-frequency-by-time tensor. In the full-band main run, the shape is $(1711, 64, 481, 8)$. This makes oscillatory structure explicit while compressing the temporal axis into a small number of windows.

The methodological value of this pairing is that it separates representation from cleaning. Any performance difference between the time-domain and STFT-based models cannot be attributed to different retained sample sets in the reported run.

5.2 Grouped split, safety fallback, and train-only normalization

The data were partitioned subject-wise with repeated `GroupShuffleSplit` attempts until a valid split satisfying a minimum-per-class condition was found. In the present study, the resulting split mode was `grouped_ok(seed=42,min_per_class=2)`. The intended grouped split was therefore achieved without resorting to the sample-wise fallback path kept in reserve as a safety mechanism.

The final partition sizes were:

- training: 1117 samples from 19 subjects,
- validation: 296 samples from 5 subjects,
- test: 298 samples from 5 subjects.

There was zero overlap of subjects and zero overlap of recording identifiers across train, validation, and test.

Normalization was computed on the training partition only and then applied to validation and test. For the time-domain tensor, the normalized signal takes the generic form

$$\hat{x}_{nct} = \frac{x_{nct} - \mu_c^{\text{train}}}{\sigma_c^{\text{train}} + \varepsilon}, \tag{5.1}$$

where the mean and standard deviation are estimated exclusively from the training data. The same principle was applied to the spectrogram representation. This prevents leakage of distributional information from held-out subjects into development.

Table 5.1: Grouped split summary for the final main Block 1 analysis.

Partition	Samples	Unique subjects	Class counts
Train	1117	19	373 left-threat, 372 right-threat, 372 safety
Validation	296	5	99 left-threat, 100 right-threat, 97 safety
Test	298	5	99 left-threat, 100 right-threat, 99 safety

5.3 Weighted sampler and class balance

Training used a weighted random sampler built from inverse class frequency on the training partition only. In practice, however, the training set was already almost perfectly balanced: the class-balance report gave 373, 372, and 372 samples across the three classes, with a max/min non-zero imbalance ratio of approximately 1.0027. The weighted sampler should therefore be interpreted as a retained training-stabilization choice rather than as a substantive rebalancing intervention.

5.4 Training procedure

All neural models were trained under the same broad training logic, although the time-domain and time–frequency branches differ slightly in early-stopping patience

and in the batch-size search range. The random seed was fixed at 42 for Python, NumPy, and PyTorch, and deterministic CuDNN behaviour was enabled where available. Training was carried out on CPU. Optimization used AdamW with the learning rate and weight decay defined by the current hyperparameter configuration. The loss was cross-entropy with label smoothing 0.05. Validation macro-F1 was the sole model-selection signal.

The common training loop was:

1. construct a weighted-sampler training loader and deterministic validation/test loaders;
2. train for up to 80 epochs;
3. after each epoch, evaluate on the validation set and append train loss, validation loss, train accuracy, validation accuracy, validation macro-F1, and the current learning rate to the history object;
4. pass validation macro-F1 to a *ReduceLROnPlateau* scheduler configured in maximization mode, with reduction factor 0.5 and patience 3;
5. save the best model state whenever validation macro-F1 improves by more than 10^{-4};
6. stop if the number of non-improving epochs reaches the patience threshold; and
7. reload the best validation state before the final validation and test evaluation.

The patience threshold was 12 epochs for the time-domain TCN families and 10 epochs for the spectrogram CNN family. During backpropagation, gradients were clipped to a maximum ℓ_2 norm of 1.0. Each trial retained the best validation state as a checkpoint, and the search stage also generated summary tables, learning curves, confusion matrices, plain-text classification reports, and a compact JSON summary of the selected models.

Table 5.2: Training-loop settings used in the modelling analyses.

Setting	Value used in the runs
Random seed	42 for Python, NumPy, and PyTorch
Device used in this study	CPU
Optimizer	AdamW
Learning-rate schedule	Reduce-on-plateau schedule on validation macro-F1 (factor 0.5, patience 3)
Maximum epochs per trial	80
Early-stopping patience	12 epochs for time-domain models; 10 epochs for spectrogram models
Gradient clipping	Maximum norm of 1.0
Model-selection signal	Validation macro-F1
Checkpointing	Best validation state retained for each trial
Outputs produced after search	Search tables, summary JSON, classification reports, learning curves, and confusion matrices

5.5 Loss function and evaluation metrics

The models were trained with a cross-entropy objective using label smoothing of 0.05. In generic form, the smoothed target distribution for class k among K classes can be written as

$$q_i = \begin{cases} 1 - \alpha + \frac{\alpha}{K}, & i = k, \\ \frac{\alpha}{K}, & i \neq k, \end{cases} \tag{5.2}$$

with $\alpha = 0.05$ and $K = 3$ in the present study.

Model selection was based on validation macro-F1, while accuracy was also monitored and reported. Macro-F1 was chosen as the main selection metric because it weights classes equally and is less likely than plain accuracy to hide class-specific degradation in a three-class problem:

$$\text{Macro-F1} = \frac{1}{K} \sum_{k=1}^{K} \text{F1}_k. \tag{5.3}$$

The test set was evaluated only for the final selected model within each family.

5.6 Baseline TCN architecture

The time-domain baseline model is a temporal convolutional network with three dilated convolutional stages. It applies input batch normalization, then three `Conv1d` blocks with kernel sizes drawn from the current configuration, batch normalization, GELU activations, dilation factors 1, 2, and 4, and `MaxPool1d(2)` after each stage. The backbone is followed by adaptive average pooling, dropout, a small fully connected hidden layer, and the final classifier.

The best full-band baseline trial used channel widths $(64, 96, 128)$, kernel sizes $(9, 7, 5)$, dropout 0.2, learning rate 10^{-4}, weight decay 2×10^{-4}, and batch size 8.

5.7 Attention-augmented TCN

The second time-domain family extends the baseline with a learned temporal attention-pooling layer. After the convolutional backbone produces a feature sequence $\mathbf{h}_t$, the attention module computes scalar scores over time, applies a softmax, and forms a weighted pooled vector:

$$\alpha_t = \frac{\exp(s_t)}{\sum_j \exp(s_j)}, \qquad \mathbf{z} = \sum_t \alpha_t \mathbf{h}_t. \tag{5.4}$$

This allows the model to emphasize the most informative parts of the cue interval while remaining close to the baseline architecture.

The best full-band attention TCN used channel widths $(64, 96, 128)$, kernel sizes $(11, 9, 7)$, dropout 0.2, attention hidden size 32, learning rate 10^{-4}, weight decay 10^{-4}, and batch size 8.

5.8 Spectrogram CNN

The time–frequency family uses a compact convolutional neural network over spectrograms. It contains three two-dimensional convolutional blocks with batch normalization and GELU activations. Pooling is applied only along the frequency axis, with a pooling kernel of 2 and stride 2 in frequency while leaving the time axis unchanged. In this way the model preserves the small number of STFT time windows while progressively aggregating spectral structure. The backbone is followed by adaptive average pooling, dropout, and a final linear classifier.

The best full-band spectrogram CNN used base channels 32, dropout 0.3, learning

rate 3×10^{-4}, weight decay 10^{-4}, and batch size 32.

5.9 Architecture detail and parameter counts

Because the analysis records preserve parameter counts for every evaluated trial, the thesis can report model capacity directly rather than only qualitatively. In the main full-band run, the selected baseline TCN contained 150,755 trainable parameters, the selected attention TCN 199,972, and the selected spectrogram CNN 111,651. In the C3/C4-only variant the parameter counts changed because the number of input channels was reduced to two: the selected baseline TCN contained 114,919 parameters, the selected attention TCN 154,120, and the selected spectrogram CNN 23,859.

These counts matter because they show that the attention model's advantage in the main run was not obtained by an order-of-magnitude increase in capacity. They also clarify that the C3/C4-only restriction changes not only the input representation but also the parameterization of the early layers.

Table 5.3: Parameter counts of the selected models in the present study.

Setting	Model family	Trainable parameters
Full-band	Baseline TCN	150,755
Full-band	Attention TCN	199,972
Full-band	Spectrogram CNN	111,651
Alpha-only	Baseline TCN	89,475
Alpha-only	Attention TCN	70,564
Alpha-only	Spectrogram CNN	111,651
Theta-only	Baseline TCN	68,995
Theta-only	Attention TCN	199,972
Theta-only	Spectrogram CNN	111,651
C3/C4-only	Baseline TCN	114,919
C3/C4-only	Attention TCN	154,120
C3/C4-only	Spectrogram CNN	23,859

5.10 Search-space design

The modelling procedure relied on bounded grid-based search spaces and then sampled manageable subsets from the full Cartesian grids. For the time-domain

families, the full baseline grid contained 32 configurations and the full attention grid 64 configurations; the study evaluated 12 baseline trials and 16 attention trials. For the spectrogram family, the full grid also contained 32 configurations and the study evaluated 16 sampled trials.

These ranges were not meant to exhaust the full model-design space. Their purpose was to support a restrained and interpretable comparison under fixed preprocessing and a limited number of subjects. Tables 5.4 and 5.5 summarize the selected ranges.

Table 5.4: Time-domain hyperparameter search space used in this study.

Parameter	Values explored
Channels	$(32, 64, 96)$, $(64, 96, 128)$
Kernels	$(9, 7, 5)$, $(11, 9, 7)$
Dropout	0.20, 0.30
Learning rate	10^{-4}, 3×10^{-4}
Weight decay	10^{-4}, 2×10^{-4}
Batch size	8
Label smoothing	0.05
Attention hidden size	16 or 32 (attention model only)

Table 5.5: Time–frequency hyperparameter search space used in this study.

Parameter	Values explored
Base channels	16, 32
Dropout	0.20, 0.30
Learning rate	10^{-4}, 3×10^{-4}
Weight decay	10^{-4}, 10^{-3}
Batch size	16, 32
Label smoothing	0.05

5.11 Exploratory variants

Beyond the main full-band analysis, the thesis reports three targeted exploratory variants.

Alpha-only variant. The alpha-only analysis used an 8–12 Hz band-pass filter while keeping the same segmentation logic, class mapping, grouped split sizes, and bounded search structure.

Theta-only variant. The theta-only analysis used a 4–8 Hz band-pass filter and includes the time-domain baseline and attention models as well as the spectrogram CNN. All three theta model families are therefore reported.

C3/C4-only variant. The C3/C4-only analysis retained only channels C3 and C4, producing tensors of shape $(1711, 2, 1125)$ in the time domain and $(1711, 2, 481, 8)$ in the STFT domain.

The exploratory variants are informative but remain secondary to the main full-band analysis. They change the input restriction while holding the broader cleaning and evaluation logic fixed.

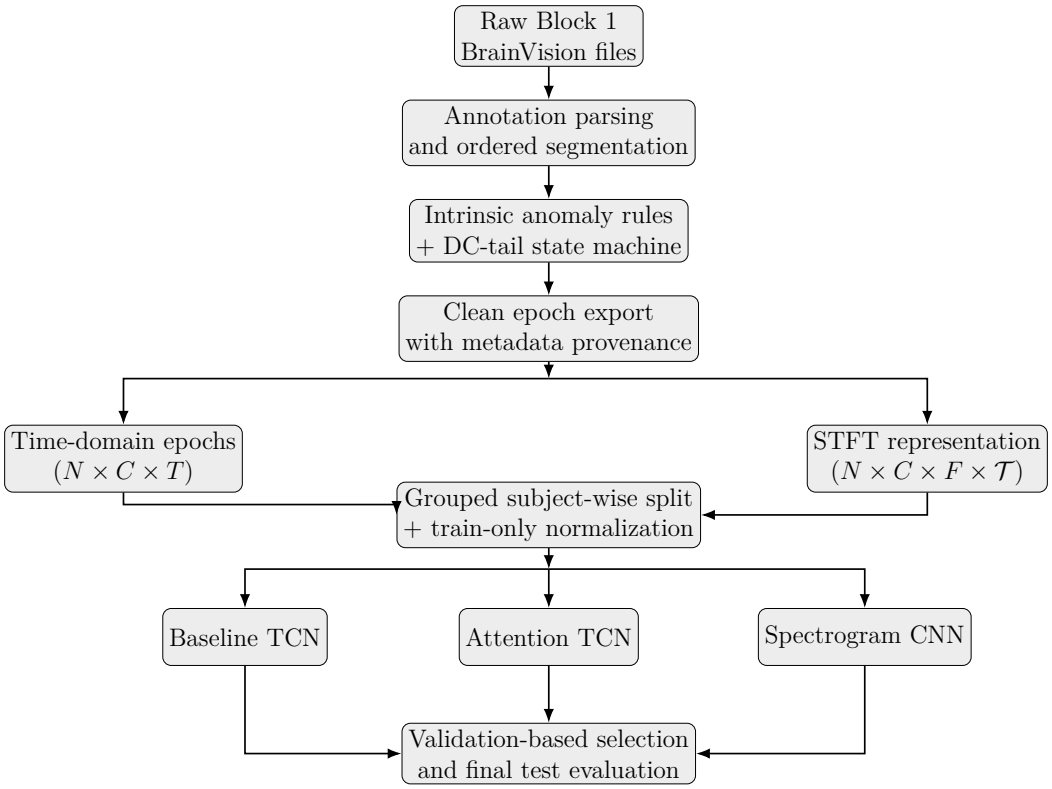

Figure 5.1: Overview of the present study pipeline. The classifier stage is the final step in a longer chain of justified transformations.

Chapter 6

Results

6.1 Quality and balance of the final exported dataset

The cleaned Block 1 export contained 1711 cue-period epochs from 29 retained recording identifiers. The three classes were almost perfectly balanced: 571 left-threat epochs, 572 right-threat epochs, and 568 safety epochs. The classification results are therefore not driven by a dominant class, and the DC-related cleaning rules did not distort the class distribution.

The subject-wise split preserved this balance. The training partition contained 1117 samples from 19 subjects, the validation partition 296 samples from 5 subjects, and the test partition 298 samples from 5 subjects. There was no subject or recording overlap between partitions.

6.2 Main full-band decoding results

Table 6.1 reports the main full-band Block 1 results. The three model families showed different validation-to-test behaviour.

The baseline TCN reached 39.53% validation accuracy and validation macro-F1 0.3950. On the held-out test set, it dropped to 36.58% accuracy and macro-F1 0.3637. This is above the balanced chance level of 33.33%, but the margin is small.

The attention TCN achieved slightly lower validation performance than the baseline TCN, with 38.85% validation accuracy and validation macro-F1 0.3839. However, it generalized better to the held-out test participants, reaching 39.93% test accuracy and macro-F1 0.3867. This was the strongest full-band held-out result.

The spectrogram CNN had the strongest validation score, with 42.57% validation accuracy and validation macro-F1 0.4253. Its test performance, however, dropped to 35.23% accuracy and macro-F1 0.3535. This suggests that the STFT representation captured validation-set regularities that did not transfer as well to unseen participants.

Table 6.1: Main full-band Block 1 results from the present study.

Model family	Val. macro-F1	Val. acc. (%)	Test macro-F1	Test acc. (%)
Baseline TCN	0.3950	39.53	0.3637	36.58
Attention TCN	0.3839	38.85	0.3867	39.93
Spectrogram CNN	0.4253	42.57	0.3535	35.23

Overall, the full-band results support a cautious interpretation. The best model exceeded chance by about 6.6 percentage points, but the classification problem remained difficult. The models did not produce sharp class separation, and the validation-to-test pattern shows why held-out evaluation is necessary.

6.3 Validation-to-test generalization behaviour

Table 6.2 summarizes the change from validation to test. The attention TCN was the most stable model: its test accuracy was slightly higher than its validation accuracy, and its macro-F1 was almost unchanged. The baseline TCN showed a moderate decrease, and the spectrogram CNN showed the largest decrease.

Table 6.2: Validation-to-test change for the main full-band models. Positive values indicate improvement from validation to test.

Model family	Δ Accuracy (test – val, pp)	Δ Macro-F1 (test – val)
Baseline TCN	-2.95	-0.0313
Attention TCN	+1.08	+0.0027
Spectrogram CNN	-7.34	-0.0718

A reasonable interpretation is that the attention mechanism helped the time-domain model retain useful temporal information while allowing it to weight parts of the cue interval adaptively. The spectrogram CNN made oscillatory structure explicit, but the retained STFT representation compressed the cue into a small number of time bins. In this dataset, that representation did not provide the most robust subject-wise generalization.

6.4 Diagnostic figures and confusion matrices

The main full-band analysis produced diagnostic figures for the selected time-domain models and for the selected spectrogram model. These figures are included because

they show more than the scalar scores in Table 6.1: they show the training curves, validation behaviour, learning-rate schedule, and final test confusion matrices.

Figure 6.1: Time-domain diagnostic figures for the main full-band analysis. The panels show train/validation loss, train/validation accuracy, validation macro-F1, learning-rate schedules, and held-out test confusion matrices for the selected baseline and attention TCN models. Confusion-matrix cells are raw held-out test-trial counts, not percentages; colour intensity increases with the number of trials in the cell.

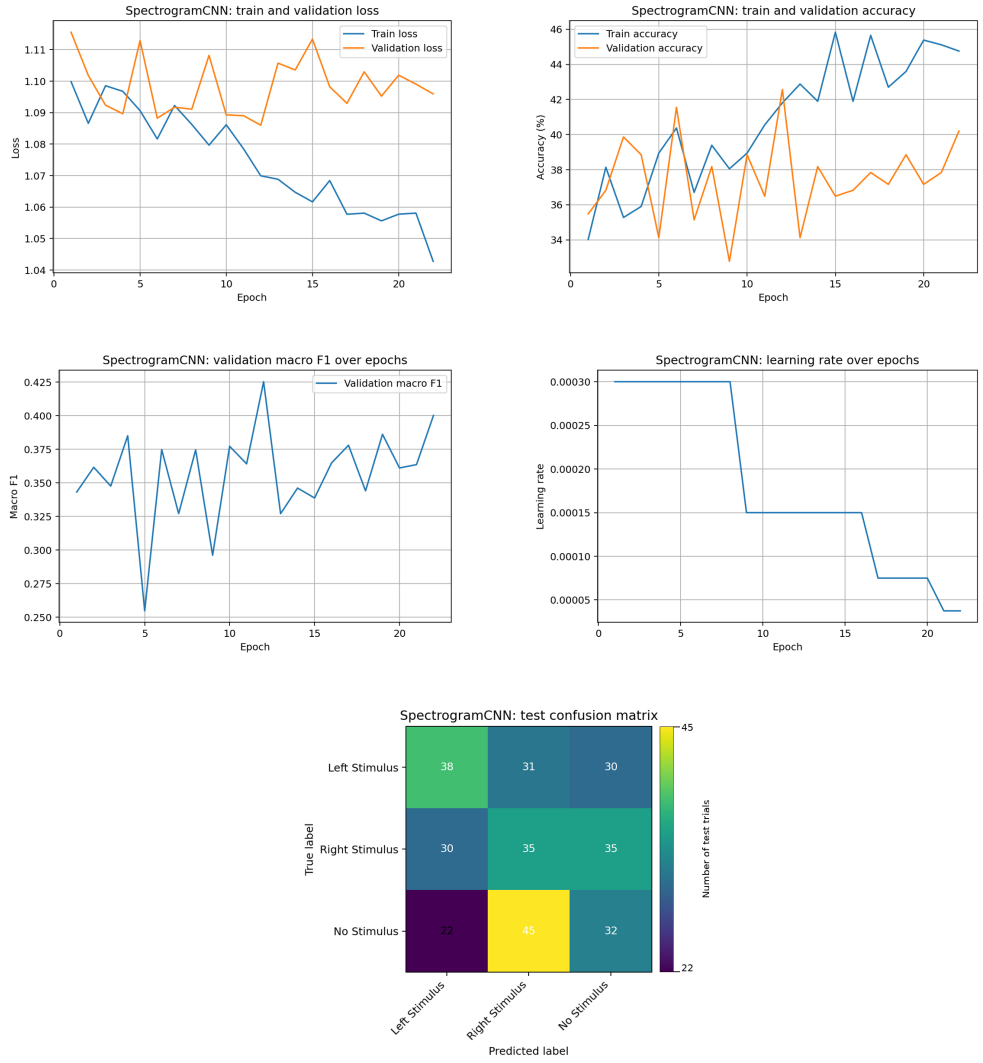

Figure 6.2: Time-frequency diagnostic figures for the main full-band analysis. The panels show train/validation loss, train/validation accuracy, validation macro-F1, learning-rate schedule, and the held-out test confusion matrix for the selected spectrogram CNN. Confusion-matrix cells are raw held-out test-trial counts, not percentages; colour intensity increases with the number of trials in the cell.

The confusion matrices show limited class separation. The attention TCN, the best full-band model, classified safety more clearly than the two threat-direction classes. Left-threat and right-threat trials were still frequently confused with safety or with each other. This pattern is consistent with the aggregate metrics: the models extracted some structure from the EEG, but the class boundaries remained weak.

The interpretation should therefore remain proportionate. Cue-period EEG did not provide an easy or highly reliable single-trial classifier of expected pain location. Rather, under conservative subject-wise evaluation, the results indicate recoverable above-chance information. The modest confusion-matrix structure limits stronger claims about separability.

6.5 Exploratory variant results

The exploratory analyses asked whether performance changed when the signal was restricted to a theoretically motivated frequency band or to a narrow central-channel subset. The results are reported in Tables 6.3, 6.4, and 6.5.

Table 6.3: Results for the alpha-restricted analysis.

Model family	Val. macro-F1	Val. acc. (%)	Test macro-F1	Test acc. (%)
Baseline TCN	0.4297	42.91	0.3448	35.23
Attention TCN	0.4121	41.22	0.3975	40.27
Spectrogram CNN	0.3705	37.50	0.3047	30.87

Table 6.4: Results for the theta-restricted analysis.

Model family	Val. macro-F1	Val. acc. (%)	Test macro-F1	Test acc. (%)
Baseline TCN	0.3625	36.49	0.2980	31.54
Attention TCN	0.3859	38.85	0.3678	36.91
Spectrogram CNN	0.3978	40.54	0.3574	37.25

Table 6.5: Results for the C3/C4-restricted analysis.

Model family	Val. macro-F1	Val. acc. (%)	Test macro-F1	Test acc. (%)
Baseline TCN	0.3615	36.15	0.3616	36.58
Attention TCN	0.3632	37.16	0.2808	33.22
Spectrogram CNN	0.3897	39.19	0.3281	33.22

The alpha-only attention TCN achieved the strongest held-out result overall, with 40.27% test accuracy and macro-F1 0.3975. This is only slightly stronger than the full-band attention TCN, so it should not be overstated. Still, it is consistent with the idea that alpha-band activity contains information relevant to spatially specific anticipation.

The theta-only results remained informative but did not surpass the best alpha-only model. The theta spectrogram CNN reached 37.25% test accuracy, and the theta attention TCN reached 36.91%. This pattern is compatible with theta contributing to broader threat-related or control-related processes rather than being the strongest source of side-specific information.

The C3/C4-only analysis did not improve performance. The best C3/C4 result was the baseline TCN at 36.58% test accuracy and macro-F1 0.3616. This suggests that the spatially relevant information is not reducible to the two central electrodes alone. Sensorimotor channels may be relevant, but the useful signal appears to benefit from a broader scalp representation.

6.6 Answer to the research question

The answer to the research question is positive, but limited. Under a Block 1-only, subject-wise, train-only-normalized protocol, cleaned cue-period EEG contains enough information to support above-chance decoding of left-threat, right-threat, and safety. The strongest confirmatory full-band result was the attention TCN with 39.93% test accuracy and macro-F1 0.3867. The strongest exploratory result was the alpha-only attention TCN with 40.27% test accuracy and macro-F1 0.3975.

At the same time, the result is modest. The confusion matrices show weak class separation, and the thesis does not claim statistical significance against a permutation null distribution because such tests were not computed. The findings should therefore be read as evidence of recoverable but limited spatial information in cue-period EEG, not as evidence of a highly reliable applied classifier.

Chapter 7

Discussion

7.1 Data cleaning as part of the scientific result

Data cleaning was part of the evidential basis of the study. In this dataset, cue-period decoding could only be interpreted after the validity of the cue epochs had been established. The final removals were concentrated in two DC-related categories: cue epochs interrupted by DC-correction events and cue epochs beginning within the post-DC exclusion tail. This concentration supports the view that cleaning was targeted rather than broadly punitive.

The post-DC exclusion tail was especially important. A DC-correction event may produce instability that extends beyond the marker itself. Removing cues that begin too soon after such an event reduces the risk that the classifier learns residual signal disturbance rather than anticipatory EEG. The threshold used in the confirmatory Block 1 branch was 1.827 s, derived from the signal-based recovery estimate plus a safety margin. This value is close to, but distinct from, the supplementary combined-block estimate of 1.836 s.

The cleaning procedure also preserved traceability. Retained and rejected samples were linked to subject, recording, timing, annotation context, and rejection reason. This makes the final dataset auditable and strengthens the interpretation of the model results.

7.2 Interpreting the full-band model comparison

The main full-band comparison shows that the attention TCN generalized best to held-out participants. The baseline TCN remained above chance but showed a moderate validation-to-test drop. The spectrogram CNN had the best validation score, but its test performance decreased more strongly. This pattern argues against selecting models by validation performance alone and supports the use of a held-out subject-wise test set.

The attention model's advantage should not be interpreted as a large architectural

breakthrough. The improvement was modest. A plausible interpretation is that temporal attention helped the model focus on informative portions of the cue period while preserving more of the original waveform structure than the compact STFT representation. In contrast, the spectrogram CNN may have captured features that were useful in the validation split but less stable across unseen participants.

The confusion matrices are essential for keeping the interpretation proportionate. Even the best model did not classify all three classes well. Safety was recognized more clearly than left-threat and right-threat, and the two threat-direction classes remained difficult. This supports the main conclusion of the thesis: the signal is informative, but weak.

7.3 Alpha, theta, and bodily specificity

The alpha-restricted analysis produced the strongest held-out result, with the alpha-only attention TCN reaching 40.27% test accuracy and macro-F1 0.3975. This fits the theoretical expectation that alpha-band activity may reflect sensorimotor preparation and bodily attention [4, 5, 6, 23, 27]. The result does not prove that alpha is the sole carrier of spatial anticipation, but it suggests that the alpha range may improve the signal-to-noise ratio for this specific decoding problem.

Theta remained informative but did not dominate. This is consistent with a view of theta as more related to general threat processing, control, uncertainty, or updating [24, 16, 17]. In a three-class task, such processes may help distinguish threat cues from safety, but they may not be the most selective source of left-versus-right information.

The C3/C4-only analysis provides an important caution. It would be too simple to assume that a spatial-motor hypothesis can be tested fully with only two central electrodes. The reduced channel set did not improve performance, suggesting that the decodable pattern is not confined to C3 and C4. The useful information may include broader scalp-distributed components as well as central sensorimotor activity.

7.4 Strengths of the present study

The first strength is the conservative evaluation design. The subject-wise split, zero subject/recording overlap, train-only normalization, and validation-based model selection reduce common sources of optimistic bias in EEG machine learning.

The second strength is methodological transparency. The thesis reports how the

raw annotations were converted into ordered segments, how cue-duration anomalies were handled, how the DC-tail threshold was estimated, and how retained epochs were exported.

The third strength is comparative framing. The thesis does not rely on a single model family. It compares time-domain and time-frequency approaches and includes targeted exploratory restrictions while keeping the main interpretation anchored to the same Block 1 cleaned dataset.

7.5 Limitations

The results are based on one grouped train/validation/test split rather than repeated grouped cross-validation. The held-out test set is therefore important, but uncertainty around the reported metrics is not fully quantified.

Permutation-based significance testing and confidence intervals were not computed. The thesis therefore does not claim formal statistical significance against a null distribution. The results are interpreted descriptively and cautiously.

The subject count is modest. Although the dataset contains 1711 clean epochs, these epochs come from 29 retained recording identifiers. Subject-wise EEG generalization is difficult at this scale.

The no-tail versus with-tail downstream ablation was not completed because the paired comparison datasets were not available. The post-DC tail is strongly justified by signal recovery and by the targeted rejection profile, but the thesis cannot claim a completed decoding comparison with and without the tail.

Finally, the main empirical scope is Block 1 acquisition. This is a deliberate choice that improves label interpretability, but it means that the results do not establish how the same approach would behave during reversal.

7.6 Future work

Future work should first replicate the analysis under repeated grouped resampling or grouped cross-validation with confidence intervals. This would help quantify uncertainty around the reported test performance. A second priority is a completed no-tail versus with-tail decoding comparison, because it would make the downstream effect of the DC exclusion rule explicit.

A third direction is more interpretable alpha-focused modelling. Spatial attention, channel-group analyses, or saliency methods could help determine whether the alpha

advantage reflects central sensorimotor activity, broader scalp-distributed patterns, or both. Finally, reversal should be analysed as a separate problem rather than merged into the main acquisition analysis. A matched Block 2 study could ask whether spatial anticipation remains decodable when cue meanings are being updated.

Chapter 8

Conclusion

This thesis asked whether cue-period EEG in Block 1 of a Pavlovian threat-learning experiment contains information about the expected spatial content of pain. The results support a positive but restrained answer. Under a subject-wise evaluation protocol with train-only normalization, cleaned EEG allowed above-chance decoding of left-threat, right-threat, and safety, although the performance was modest.

The best full-band model was the attention-augmented TCN, which reached 39.93% test accuracy and macro-F1 0.3867. The alpha-restricted attention TCN achieved the strongest exploratory result, reaching 40.27% test accuracy and macro-F1 0.3975. These values are above the balanced chance level of 33.33%, but the confusion matrices and validation-to-test behaviour show that the classes remain difficult to separate.

The main contribution of the thesis is the combination of a transparent cleaning pipeline, explicit handling of DC-correction events, careful dataset provenance, and conservative subject-wise evaluation. Within that framework, the results suggest that the spatial content of pain anticipation is present in cue-period EEG, but only weakly and with substantial limits on generalization.

Appendix A

Reproducibility Notes and Data Dictionary

A.1 Shapes of the exported arrays

The main exported arrays had the shapes reported in Table A.1. The same split sizes were used across the full-band, alpha-only, theta-only, and C3/C4-only variants.

Table A.1: Main tensor shapes reported by the final analyses.

Setting	Time-domain shape	STFT shape	Notes
Full-band Block 1	$(1711, 64, 1125)$	$(1711, 64, 481, 8)$	Main confirmatory run
Alpha-only	$(1711, 64, 1125)$	$(1711, 64, 481, 8)$	8–12 Hz bandpass before export
Theta-only	$(1711, 64, 1125)$	$(1711, 64, 481, 8)$	4–8 Hz bandpass before export
C3/C4-only	$(1711, 2, 1125)$	$(1711, 2, 481, 8)$	Channel restriction only

A.2 Event and segment categories encountered in the raw scan

The main Block 1 analysis reported the global segment counts shown in Table A.2 across the scanned recordings before cleaning. Presenting these counts as a table rather than a long bullet list makes the symmetry of the annotation stream easier to see.

These counts confirm that the candidate task space is symmetric at the annotation

Table A.2: Global segment counts reported by the final main analysis before cleaning.

Segment category	Count
New recording segment	30
DC-correction segment	144
No-stimulus cue segment	580
TMS-after-no-stimulus segment	580
Left-stimulus cue (reinforced)	406
TMS after left-stimulus cue (reinforced)	406
Right-stimulus cue (reinforced)	406
TMS after right-stimulus cue (reinforced)	406
Right-stimulus cue (non-reinforced)	174
TMS after right-stimulus cue (non-reinforced)	174
Left-stimulus cue (non-reinforced)	174
TMS after left-stimulus cue (non-reinforced)	174

level: before cleaning, the stream contains 580 left-cue instances, 580 right-cue instances, and 580 safety-cue instances.

A.3 Per-recording removal profile from the dataset-audit analysis

The dataset-audit analysis provides a compact per-recording removal summary. In the visible printed rows, most recordings retained all 60 candidate task epochs, whereas a small number accounted for the removals observed in the final audit. The most affected printed example was `03as_block1` with 55 kept out of 60 candidate epochs, corresponding to 5 removed epochs (8.33%). The printed rows for `02gd_block1`, `09dm_block1`, and `10lc_block1` each showed 1 removed epoch out of 60 (1.67%). This supports the broader interpretation that the cleaning stage was targeted rather than globally punitive.

A.4 Audit note on raw-scan versus exported dataset counts

Two verified counts should be kept conceptually distinct. The Block 1 raw scan covered 30 recording files, while the final machine-learning export contained 29 retained recording identifiers. The final empirical claims therefore refer to the

exported dataset of 29 retained recording identifiers and 1711 clean cue epochs.

Table A.3: Compact audit note distinguishing the Block 1 raw scan from the exported dataset used for learning.

Pipeline stage	Verified count
Raw Block 1 recording files scanned	30
Retained recording identifiers in the final learning export	29
Retained clean cue epochs in final machine-learning dataset	1711
Raw-scan/export distinction	The raw scan covered 30 recording files, while the final learning export contained 29 retained recording identifiers
Scope of the final empirical claims	The exported Block 1 acquisition dataset used for learning

Appendix B

Hyperparameter Search Grids

This appendix reports the bounded search grids used in the final analyses.

B.1 Time-domain baseline grid

The full baseline TCN grid comprised the Cartesian product of the following values:

- channels in $\{(32, 64, 96), (64, 96, 128)\}$;
- kernels in $\{(9, 7, 5), (11, 9, 7)\}$;
- dropout in $\{0.20, 0.30\}$;
- learning rate in $\{10^{-4}, 3 \times 10^{-4}\}$;
- weight decay in $\{10^{-4}, 2 \times 10^{-4}\}$;
- label smoothing $= 0.05$;
- batch size $= 8$.

This yields 32 configurations, of which 12 were sampled in the study.

B.2 Time-domain attention grid

The full attention-TCN grid added one more factor:

- attention hidden size in $\{16, 32\}$.

All other factors matched the baseline grid, yielding 64 configurations in total, of which 16 were evaluated in each analysis.

B.3 Spectrogram grid

The full spectrogram CNN grid comprised:

- base channels in $\{16, 32\}$;
- dropout in $\{0.20, 0.30\}$;
- learning rate in $\{10^{-4}, 3 \times 10^{-4}\}$;
- weight decay in $\{10^{-4}, 10^{-3}\}$;
- batch size in $\{16, 32\}$;
- label smoothing $= 0.05$.

This yields 32 configurations, of which 16 were evaluated in each analysis.

Appendix C

Selected Configurations and Supplementary Result Tables

C.1 Best configurations reported by the final analyses

Table C.1 summarizes the selected configurations across the main and exploratory Block 1 settings.

Table C.1: Best configurations reported by the final analyses.

Setting	Model family	Configuration
Full-band	Baseline TCN	channels=(64,96,128), kernels=(9,7,5), dropout=0.2, lr=1e-4, wd=2e-4, batch=8
Full-band	Attention TCN	channels=(64,96,128), kernels=(11,9,7), dropout=0.2, att_hidden=32, lr=1e-4, wd=1e-4, batch=8
Full-band	Spectrogram CNN	base_channels=32, dropout=0.3, lr=3e-4, wd=1e-4, batch=32
Alpha-only	Baseline TCN	channels=(32,64,96), kernels=(11,9,7), dropout=0.3, lr=3e-4, wd=2e-4, batch=8
Alpha-only	Attention TCN	channels=(32,64,96), kernels=(9,7,5), dropout=0.2, att_hidden=16, lr=3e-4, wd=2e-4, batch=8
Alpha-only	Spectrogram CNN	base_channels=32, dropout=0.3, lr=3e-4, wd=1e-4, batch=32
Theta-only	Baseline TCN	channels=(32,64,96), kernels=(9,7,5), dropout=0.3, lr=3e-4, wd=1e-4, batch=8

Setting	Model family	Configuration
Theta-only	Attention TCN	channels=(64,96,128), kernels=(11,9,7), dropout=0.3, att_hidden=32, lr=1e-4, wd=2e-4, batch=8
Theta-only	Spectrogram CNN	base_channels=32, dropout=0.2, lr=3e-4, wd=1e-4, batch=16
C3/C4-only	Baseline TCN	channels=(64,96,128), kernels=(9,7,5), dropout=0.2, lr=1e-4, wd=1e-4, batch=8
C3/C4-only	Attention TCN	channels=(64,96,128), kernels=(11,9,7), dropout=0.3, att_hidden=16, lr=1e-4, wd=2e-4, batch=8
C3/C4-only	Spectrogram CNN	base_channels=16, dropout=0.3, lr=1e-4, wd=1e-3, batch=16

C.2 Supplementary best-result tables

Tables C.2 and C.3 report the best result for each model family in each main or exploratory Block 1 setting.

Table C.2: Best results for the main full-band, alpha-only, and theta-only settings.

Analysis setting	Model family	Params	Val. macro-F1	Val. acc. (%)	Test macro-F1	Test acc. (%)
Main full-band	Baseline TCN	150755	0.3950	39.53	0.3637	36.58
Main full-band	Attention TCN	199972	0.3839	38.85	0.3867	39.93
Main full-band	Spectrogram CNN	111651	0.4253	42.57	0.3535	35.23
Alpha-only	Baseline TCN	89475	0.4297	42.91	0.3448	35.23
Alpha-only	Attention TCN	70564	0.4121	41.22	0.3975	40.27
Alpha-only	Spectrogram CNN	111651	0.3705	37.50	0.3047	30.87
Theta-only	Baseline TCN	68995	0.3625	36.49	0.2980	31.54
Theta-only	Attention TCN	199972	0.3859	38.85	0.3678	36.91
Theta-only	Spectrogram CNN	111651	0.3978	40.54	0.3574	37.25

Table C.3: Best results for the C3/C4-only exploratory setting.

Analysis setting	Model family	Params	Val. macro-F1	Val. acc. (%)	Test macro-F1	Test acc. (%)
C3/C4-only	Baseline TCN	114919	0.3615	36.15	0.3616	36.58
C3/C4-only	Attention TCN	154120	0.3632	37.16	0.2808	33.22
C3/C4-only	Spectrogram CNN	23859	0.3897	39.19	0.3281	33.22

Appendix D

Supplementary Diagnostic Figures

The main full-band diagnostic figures are included in Chapter 6. This appendix gathers the corresponding alpha-only, theta-only, and C3/C4-only diagnostics. Confusion-matrix cells in these figures are raw held-out test-trial counts, not percentages; colour intensity increases with the number of trials in the cell.

D.1 Alpha-restricted analysis

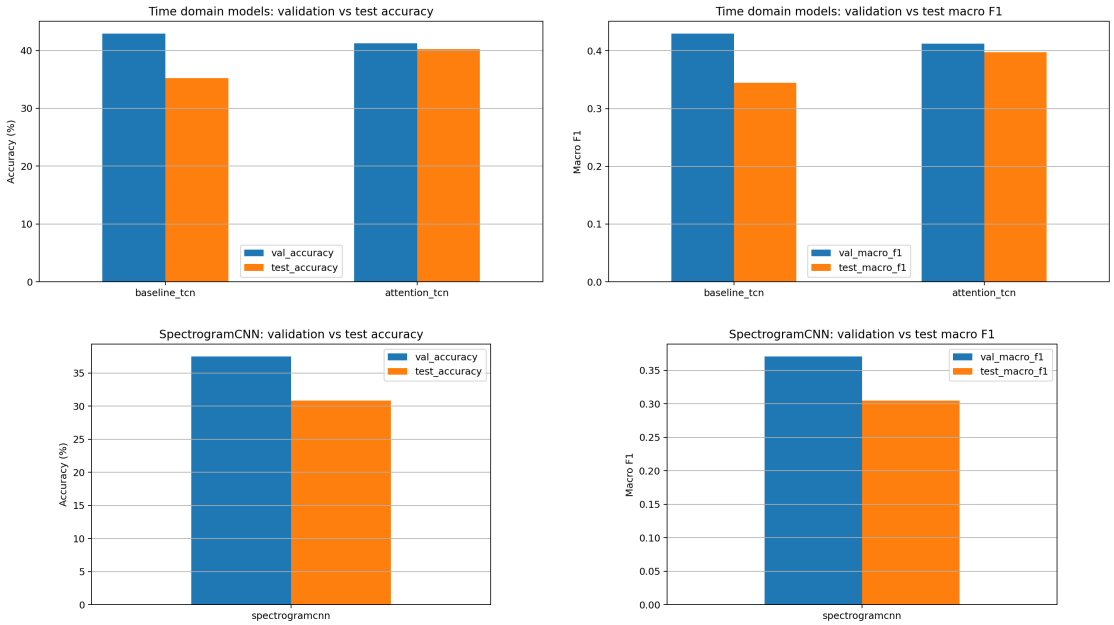

Figure D.1: Summary figures for the alpha-restricted analysis. The grid presents held-out validation-versus-test comparisons for the time-domain and spectrogram analyses.

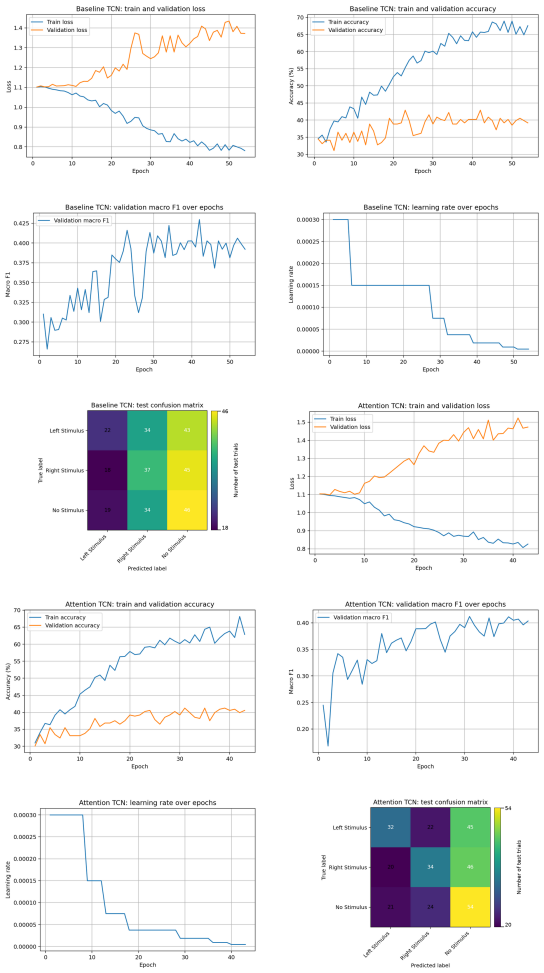

Figure D.2: Time-domain diagnostic figures for the alpha-restricted analysis, including learning curves, learning-rate schedules, and held-out test confusion matrices for the selected baseline and attention TCN runs. Confusion-matrix cells are raw held-out test-trial counts, not percentages; colour intensity increases with the number of trials in the cell.

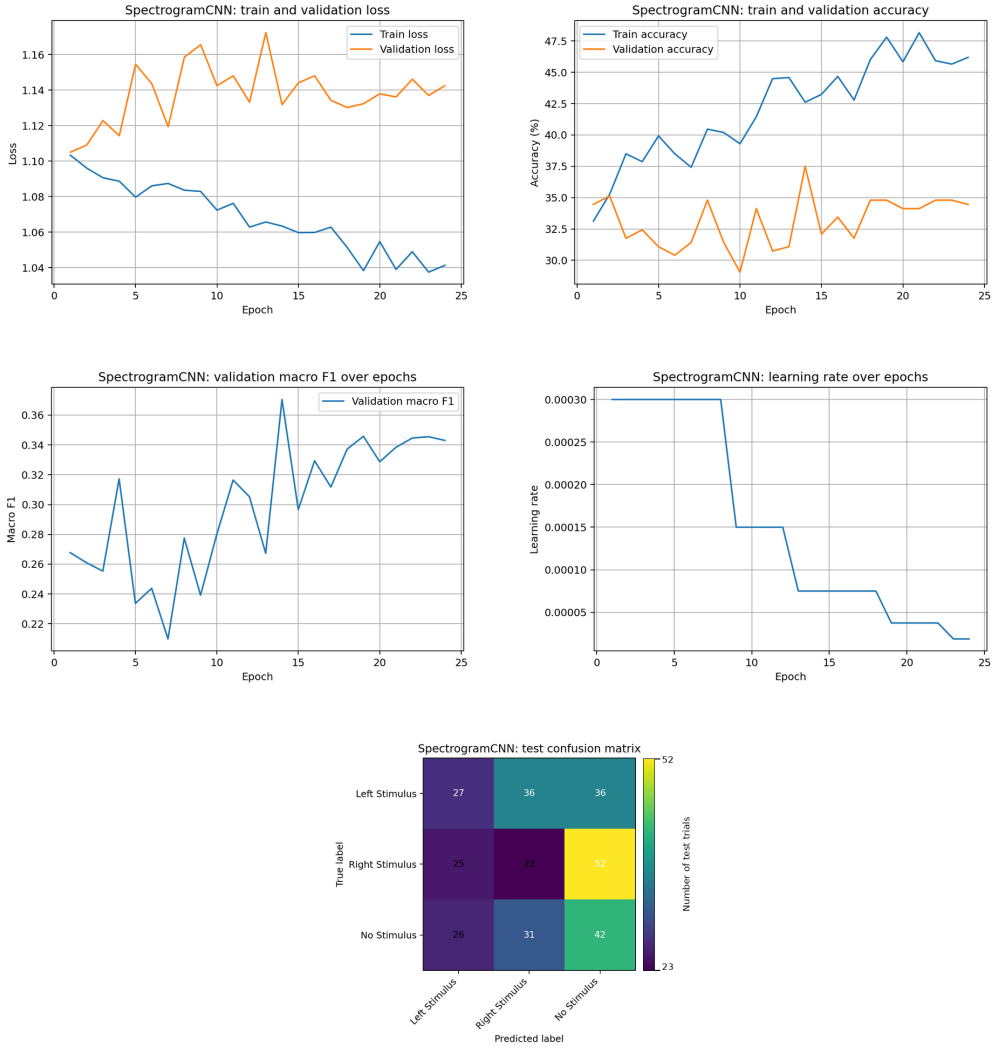

Figure D.3: Time-frequency diagnostic figures for the alpha-restricted analysis, including learning curves, learning-rate schedule, and held-out test confusion matrix for the selected spectrogram CNN. Confusion-matrix cells are raw held-out test-trial counts, not percentages; colour intensity increases with the number of trials in the cell.

D.2 Theta-restricted analysis

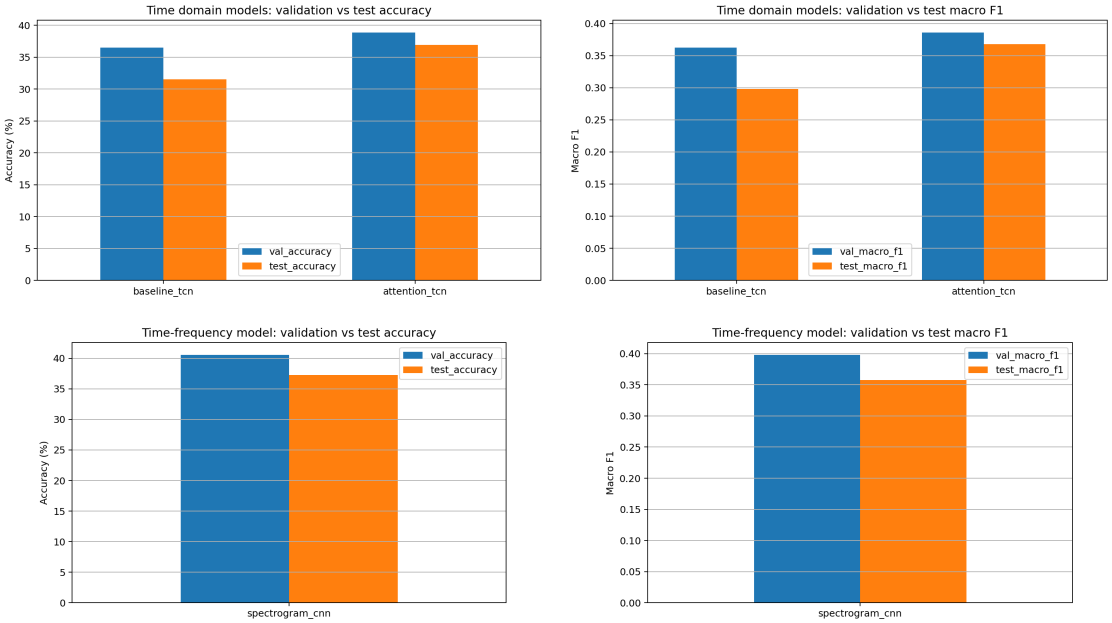

Figure D.4: Summary figures for the theta-restricted analysis. The grid shows held-out summaries for the theta-restricted analyses in both representations.

Figure D.5: Time-domain diagnostic figures for the theta-restricted analysis, including learning curves, learning-rate schedules, and held-out test confusion matrices for the selected baseline and attention TCN runs. Confusion-matrix cells are raw held-out test-trial counts, not percentages; colour intensity increases with the number of trials in the cell.

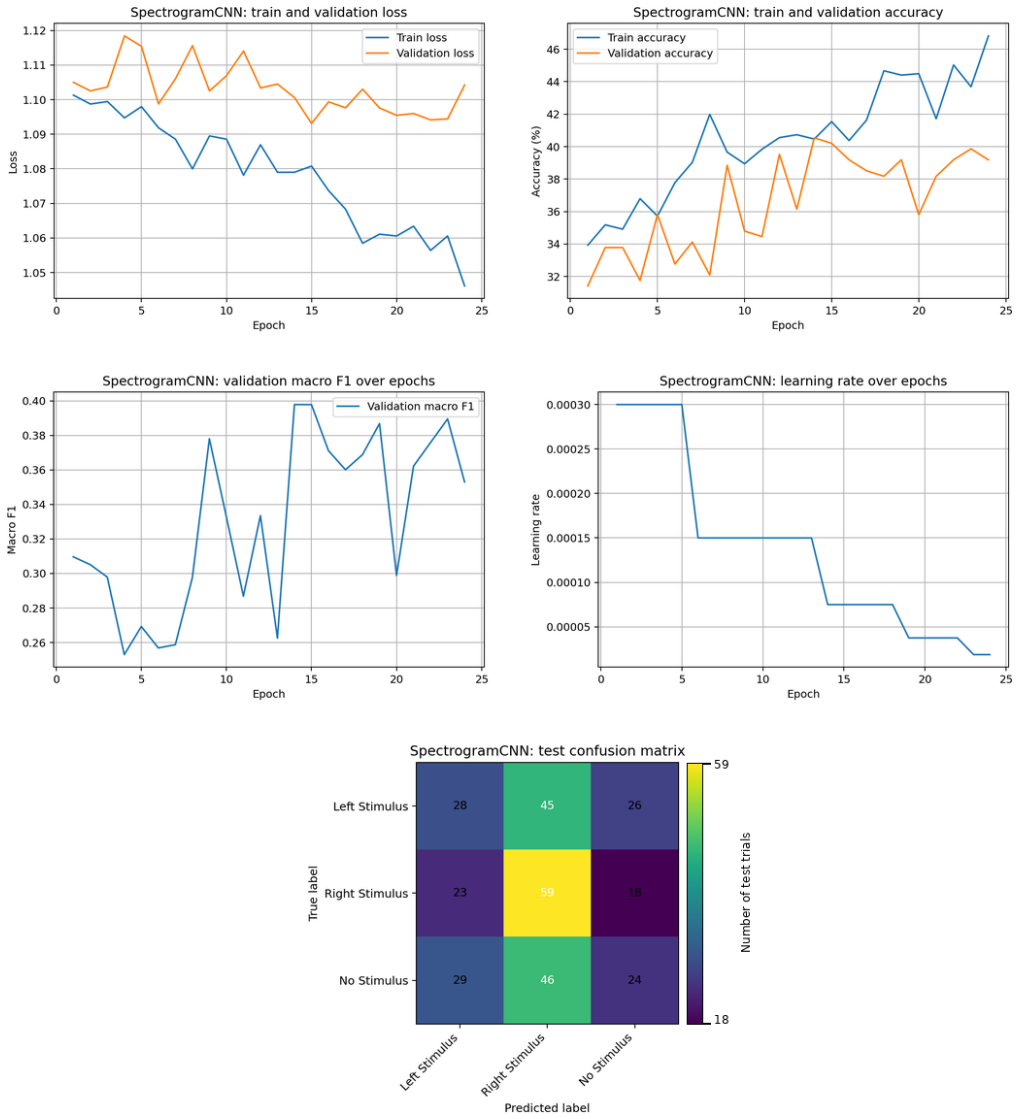

Figure D.6: Time-frequency diagnostic figures for the theta-restricted analysis, including learning curves, learning-rate schedule, and held-out test confusion matrix for the selected spectrogram CNN. Confusion-matrix cells are raw held-out test-trial counts, not percentages; colour intensity increases with the number of trials in the cell.

D.3 C3/C4-restricted analysis

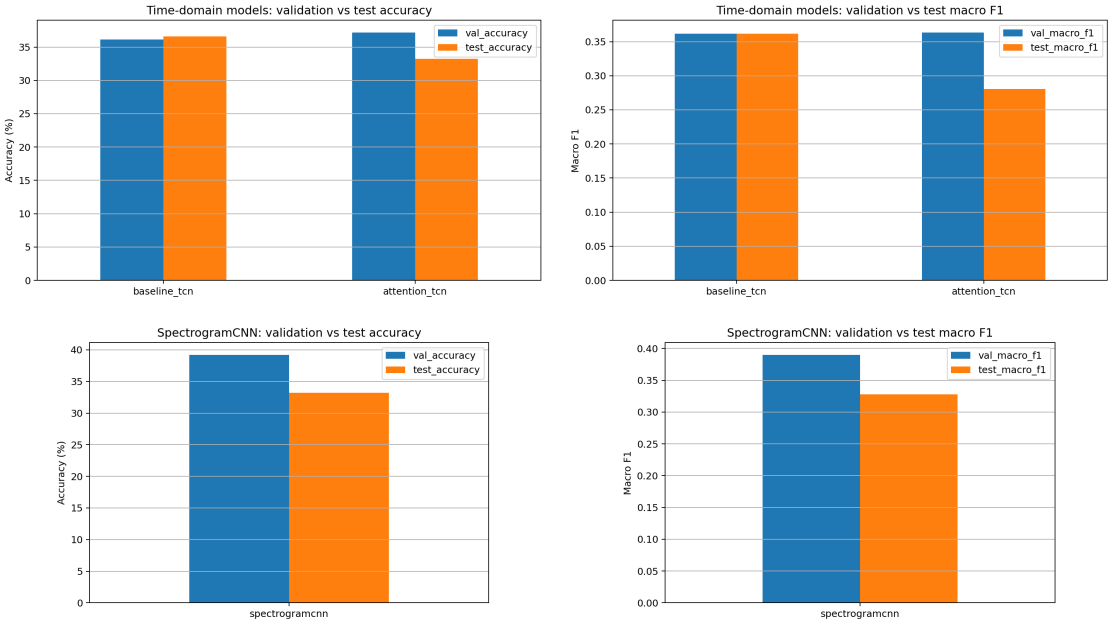

Figure D.7: Summary figures for the C3/C4-restricted analysis. The reduced-input results are shown in the same reporting format used for the other analyses.

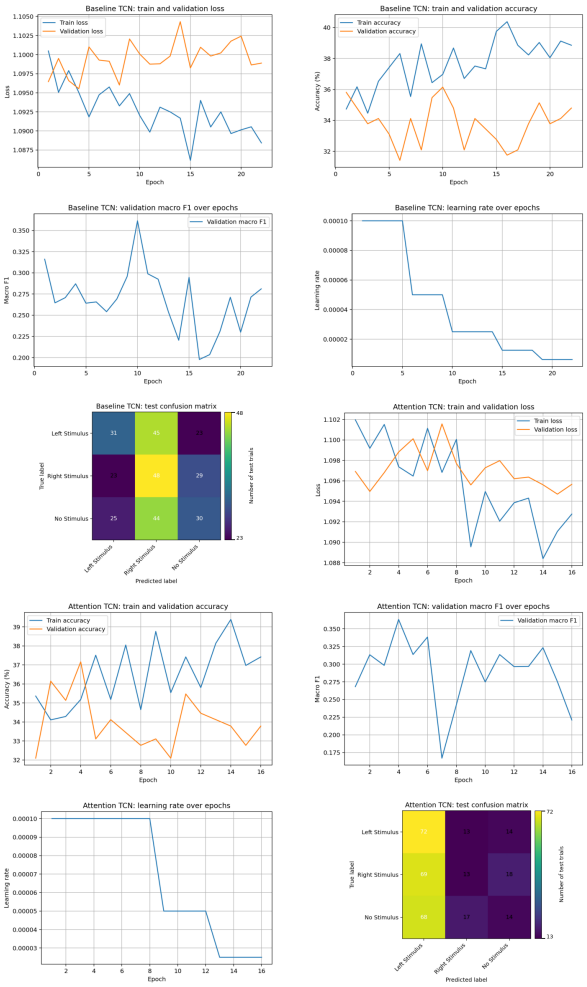

Figure D.8: Time-domain diagnostic figures for the C3/C4-restricted analysis, including learning curves, learning-rate schedules, and held-out test confusion matrices for the selected baseline and attention TCN runs. Confusion-matrix cells are raw held-out test-trial counts, not percentages; colour intensity increases with the number of trials in the cell.

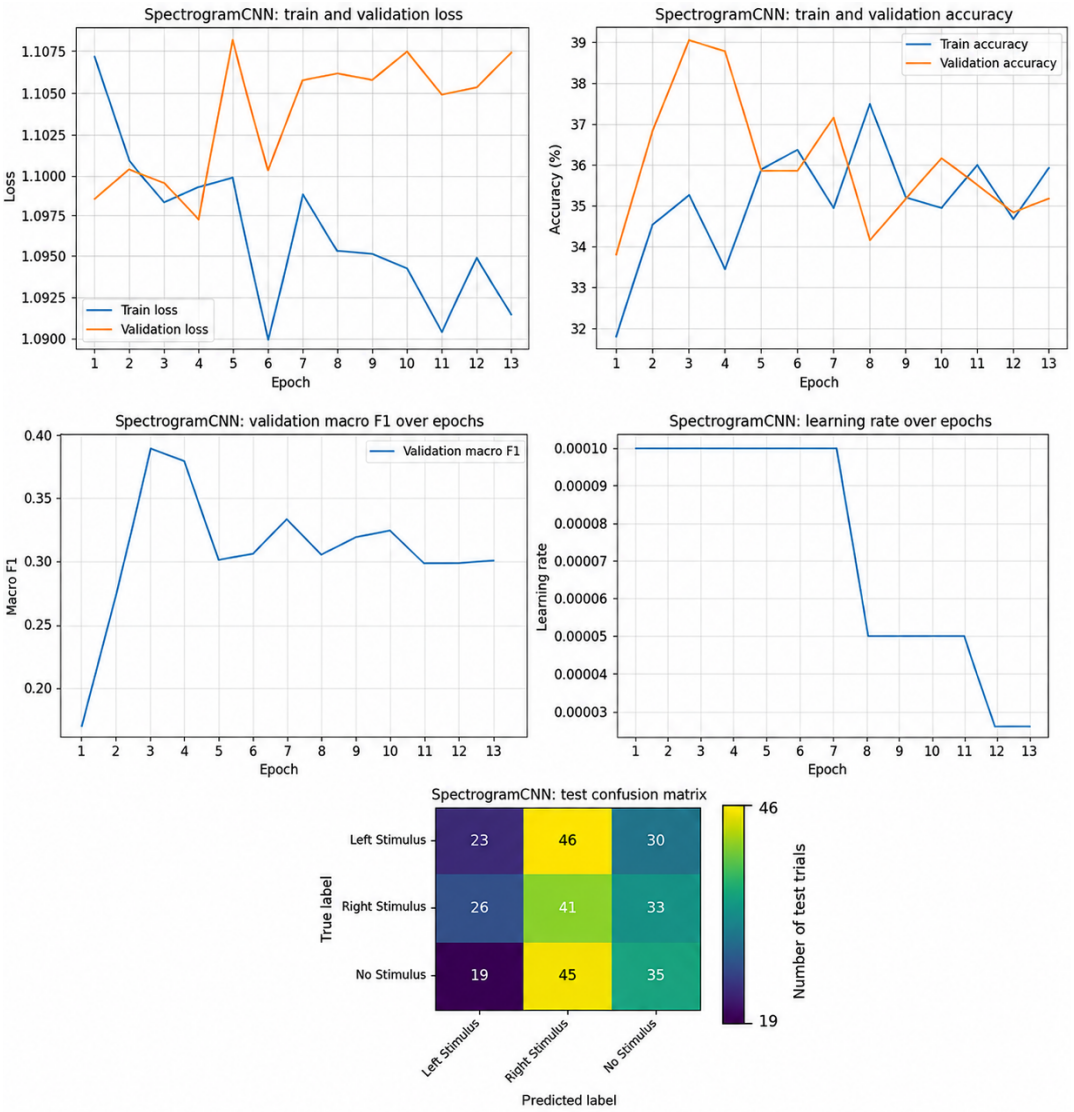

Figure D.9: Time-frequency diagnostic figures for the C3/C4-restricted analysis, including learning curves, learning-rate schedule, and held-out test confusion matrix for the selected spectrogram CNN. Confusion-matrix cells are raw held-out test-trial counts, not percentages; colour intensity increases with the number of trials in the cell.

Appendix E

Supplementary Analytical Tables and Notes

E.1 Comparative overview of analytical variants

The analytical variants are easier to compare when the data scope and the strongest held-out results are separated. Table E.1 records the exported representation shapes, and Table E.2 records the strongest held-out result in each setting.

Table E.1: Comparative overview of data scope and exported representations across analytical variants.

Analysis setting	Run label or scope	Band-pass	Retained channels	Exported shapes
Main full-band	Block 1 confirmatory run	0.5–100 Hz	All EEG channels	Time: (1711, 64, 1125); STFT: (1711, 64, 481, 8)
Alpha-only	Block 1 alpha-restricted run	8–12 Hz	All EEG channels	Time: (1711, 64, 1125); STFT: (1711, 64, 481, 8)
Theta-only	Block 1 theta-restricted run	4–8 Hz	All EEG channels	Time: (1711, 64, 1125); STFT: (1711, 64, 481, 8)
C3/C4-only	Block 1 channel-restricted run	0.5–100 Hz	C3 and C4 only	Time: (1711, 2, 1125); STFT: (1711, 2, 481, 8)

Table E.2: Strongest held-out result in each analytical setting. Accuracy is reported first, followed by macro-F1.

Analysis setting	Strongest held-out result(s)
Main full-band	Attention TCN: 39.93% / 0.3867; baseline TCN: 36.58% / 0.3637; spectrogram CNN: 35.23% / 0.3535
Alpha-only	Attention TCN: 40.27% / 0.3975; baseline TCN: 35.23% / 0.3448; spectrogram CNN: 30.87% / 0.3047
Theta-only	Spectrogram CNN: 37.25% / 0.3574; attention TCN: 36.91% / 0.3678; baseline TCN: 31.54% / 0.2980
C3/C4-only	Baseline TCN: 36.58% / 0.3616; attention TCN: 33.22% / 0.2808; spectrogram CNN: 33.22% / 0.3281

E.2 Supplementary note on the two-block run

A jointly trained Block 1+Block 2 analysis is reported only as supplementary context and is not used as a main empirical result in this thesis. The reason is conceptual rather than technical: reversal changes cue meanings and introduces updating-related processes, so its labels are less directly comparable to the stable acquisition labels in Block 1. For this reason, the main conclusion is restricted to the Block 1 acquisition analysis.

Bibliography

[1] N. C. Rogasch, R. A. Sullivan, J. J. Thomson, et al., "Analysing concurrent transcranial magnetic stimulation and electroencephalographic data: A review and introduction to the open-source TESA software," *NeuroImage*, 147:934–951, 2017.

[2] A. Brancaccio, D. Tabarelli, A. Zazio, et al., "Towards the definition of a standard in TMS–EEG data preprocessing," *NeuroImage*, 301:120874, 2024.

[3] M. Ploner, C. Sorg, and J. Gross, "Brain Rhythms of Pain," *Trends in Cognitive Sciences*, 21(2):100–110, 2017.

[4] C. Babiloni, A. Brancucci, C. Del Percio, P. Capotosto, L. Arendt-Nielsen, and A. C. N. Chen, "Anticipatory electroencephalography alpha rhythm predicts subjective perception of pain intensity," *The Journal of Pain*, 7(10):709–717, 2006.

[5] C. Babiloni, A. Brancucci, M. Del Percio, et al., "Cortical alpha rhythms are related to the anticipation of sensorimotor interaction between painful stimuli and movements: A high-resolution EEG study," *The Journal of Pain*, 9(10):902–911, 2008.

[6] C. Babiloni, M. Del Percio, F. Infarinato, et al., "Cortical EEG alpha rhythms reflect task-specific somatosensory and motor interactions in humans," *Clinical Neurophysiology*, 125(10):1936–1945, 2014.

[7] W. Peng, X. Huang, Y. Liu, and F. Cui, "Predictability modulates the anticipation and perception of pain in both self and others," *Social Cognitive and Affective Neuroscience*, 14(7):747–757, 2019.

[8] Y. Roy, H. J. Banville, I. Albuquerque, A. Gramfort, T. H. Falk, and J. Faubert, "Deep learning-based electroencephalography analysis: A systematic review," *Journal of Neural Engineering*, 16(5):051001, 2019.

[9] A. Craik, Y. He, and J. L. Contreras-Vidal, "Deep learning for electroencephalogram (EEG) classification tasks: A review," *Journal of Neural Engineering*, 16(3):031001, 2019.

[10] G. Varoquaux, "Cross-validation failure: Small sample sizes lead to large error bars," *NeuroImage*, 180(A):68–77, 2018.

[11] L. Y. Atlas and T. D. Wager, "How expectations shape pain," *Neuroscience Letters*, 520(2):140–148, 2012.

[12] K. Wiech, "Deconstructing the sensation of pain: The influence of cognitive processes on pain perception," *Science*, 354(6312):584–587, 2016.

[13] C. Büchel, "The role of expectations, control and reward in the development of pain persistence based on a unified model," *eLife*, 12:e81795, 2023.

[14] J. A. Greco and I. Liberzon, "Neuroimaging of fear-associated learning," *Neuropsychopharmacology*, 41:320–334, 2016.

[15] M. A. Fullana, B. J. Harrison, C. Soriano-Mas, et al., "Neural signatures of human fear conditioning: An updated and extended meta-analysis of fMRI studies," *Molecular Psychiatry*, 21(4):500–508, 2016.

[16] D. Schiller, I. Levy, Y. Niv, J. E. LeDoux, and E. A. Phelps, "From fear to safety and back: Reversal of fear in the human brain," *The Journal of Neuroscience*, 28(45):11517–11525, 2008.

[17] H. S. Savage, C. G. Davey, M. A. Fullana, and B. J. Harrison, "Clarifying the neural substrates of threat and safety reversal learning in humans," *NeuroImage*, 207:116427, 2020.

[18] G. Brookshire, J. Kasper, N. M. Blauch, et al., "Data leakage in deep learning studies of translational EEG," *Frontiers in Neuroscience*, 18:1373515, 2024.

[19] L. Y. Atlas, "How Instructions, Learning, and Expectations Shape Pain and Neurobiological Responses," *Annual Review of Neuroscience*, 46:167–189, 2023.

[20] A. Meulders, "Fear in the context of pain: Lessons learned from 100 years of fear conditioning research," *Behaviour Research and Therapy*, 131:103635, 2020.

[21] C. Vanden Bulcke, S. Van Damme, W. Durnez, and G. Crombez, "The anticipation of pain at a specific location of the body prioritizes tactile stimuli at that location," *Pain*, 154(8):1464–1468, 2013.

[22] C. Vanden Bulcke, G. Crombez, C. Spence, and S. Van Damme, "Are the spatial features of bodily threat limited to the exact location where pain is expected?," *Acta Psychologica*, 153:113–119, 2014.

[23] S. Haegens, L. Luther, and O. Jensen, "Somatosensory anticipatory alpha activity increases to suppress distracting input," *Journal of Cognitive Neuroscience*, 24(3):677–685, 2012.

[24] J. F. Cavanagh and M. J. Frank, "Frontal theta as a mechanism for cognitive control," *Trends in Cognitive Sciences*, 18(8):414–421, 2014.

[25] Y. Tu, X. Tan, A. Bai, et al., "Decoding Subjective Intensity of Nociceptive Pain from Pre-stimulus and Post-stimulus Brain Activities," *Frontiers in Computational Neuroscience*, 10:32, 2016.

[26] P. Taesler and M. Rose, "Multivariate prediction of pain perception based on pre-stimulus activity," *Scientific Reports*, 12:3199, 2022.

[27] O. Jensen and A. Mazaheri, "Shaping Functional Architecture by Oscillatory Alpha Activity: Gating by Inhibition," *Frontiers in Human Neuroscience*, 4:186, 2010.

[28] T. B. Lonsdorf, M. M. Menz, M. Andreatta, et al., "Don't fear 'fear conditioning': Methodological considerations for the design and analysis of studies on human fear acquisition, extinction, and return of fear," *Neuroscience and Biobehavioral Reviews*, 77:247–285, 2017.