LaBraM at Scale: Scaling Data and Models for EEG Pretraining

Jun-Yu Pan[1,†] **Ruicheng Yin**[2] **Bao-Liang Lu**[1] **Dongsheng Li**[3]
Wei-Long Zheng[1,*] **Yansen Wang**[3,*]

[1]Shanghai Jiao Tong University, Shanghai, China [2]Fudan University, Shanghai, China
[3]Microsoft Research Asia, Shanghai, China

{panjunyu,bllu,weilong}@sjtu.edu.cn rcyin23@m.fudan.edu.cn {dongsli,yansenwang}@microsoft.com

Abstract

LaBraM, an early EEG foundation model, demonstrated that large-scale EEG pretraining can yield strong representations for diverse downstream tasks. Building on this foundation, we expand pretraining data and model capacity while retaining its core encoder and objective. We train ten encoders with 5.8 million to 1.97 billion parameters on three EEG collections spanning 25,067–79,910 hours, adapting signal preparation and training to heterogeneous recordings. Across 23 tasks, Atom-L* reaches 59.05% mean balanced accuracy at its best evaluated full fine-tuning (Full-FT) checkpoint, 3.61 percentage points above the strongest baseline EEG foundation model. Quark-G leads linear probing (LP) at 54.33%, a 1.24-point margin. Both also lead average task rank. At the last evaluated checkpoints, larger encoders improve Full-FT more consistently than LP, and stronger LP does not guarantee stronger Full-FT. Preserving patch structure improves LP over average pooling. Data-fraction experiments suggest capacity-dependent saturation, while checkpoint trajectories show that downstream performance can decline as pretraining loss falls. Joint classification with a shared Qwen3 readout reaches 50.50% with Molecule-H and Qwen3-8B; increasing language-model size provides no consistent gain. These results demonstrate the value of scaling LaBraM and the importance of matching pretraining scale to downstream adaptation.



Figure 1. Full-FT and LP performance of ten extension encoders and three selected-checkpoint references (*) on 23 tasks. Stars mark checkpoints selected by mean Full-FT score; LP uses the same checkpoints. Dashed lines mark the strongest baseline in each setting. Figures 4 and 5 show the complete comparisons.

*Corresponding authors.
†Work done during Jun-Yu Pan's internship at Microsoft Research Asia.

1 Introduction

LaBraM, an early EEG foundation model, demonstrated that pretraining on large EEG collections can learn representations that transfer effectively across diverse downstream tasks [1]. Subsequent EEG foundation models have provided further evidence for the effectiveness of large-scale pretraining [2–10]. Building on LaBraM, we further investigate the downstream gains from expanding pretraining data and model capacity while retaining its core encoder design and pretraining objective.

To study this expansion, we assemble three EEG collections—Quark, Atom, and Molecule—spanning 25,067–79,910 hours, and train ten encoders ranging from 5.8 million to 1.97 billion parameters. Combining recordings from diverse sources requires handling differences in channel layout, signal scale, and recording length. Extended signal preparation and batching accommodate these inputs, while streamed data access, node-local storage, and distributed training support the larger collections and encoders.

The resulting encoders lead the comparisons across 23 classification tasks under both full fine-tuning (Full-FT) and linear probing (LP). With 46.3 million parameters, Atom-L* reaches 59.05% mean balanced accuracy at its best evaluated Full-FT checkpoint, exceeding REVE Large by 3.61 percentage points. Quark-G leads LP at 54.33%, and both models lead average task rank in their respective settings. The gains also extend to comparisons with small supervised models: Atom-L and Molecule-H each exceed the stronger of EEGNet and EEGConformer on 16 tasks under Full-FT, compared with at most 12 for a baseline EEG foundation model (Section 5).

Further analysis examines how downstream gains vary with encoder size, pretraining data, and training duration. At the last evaluated checkpoints, larger encoders improve Full-FT more consistently than LP. The two evaluation settings also produce different model orderings: stronger LP does not guarantee stronger Full-FT. Retaining patch structure improves LP over average pooling, showing that the readout affects how well frozen features can be used. Together, these results motivate assessing both access to frozen features and the gains from task-specific encoder adaptation (Section 5.3).

The data experiments show why recording hours alone do not determine transfer performance. Reduced-data Atom runs outperform Quark at matched encoder sizes despite using fewer hours, suggesting that broader source coverage may improve data efficiency. Within a fixed collection, additional data can eventually bring diminishing or negative returns. Fraction sweeps motivate an empirical capacity-dependent saturation relation, with fitted Full-FT peaks near 22,000 hours for Base and 40,000 hours for Large. Even within a single training run, downstream performance can decline while pretraining loss falls. Judging the benefit of further pretraining therefore requires downstream evaluation (Section 6).

Beyond traditional evaluation, we further explore our models' performance in a multitask setting, drawing inspiration from NeuroLM [5], UniMind [11], and GVG [12]. Frozen EEG encoders are paired with a shared Qwen3 readout for joint classification across all 23 tasks. Larger EEG collections and encoders show complementary gains, with Molecule-H and Qwen3-8B reaching 50.50% mean balanced accuracy. Increasing language-model size provides no consistent improvement. Together with the LP analysis, these results suggest that preserving EEG information and adapting it to the readout are important for realizing the benefits of encoder scale (Section 7).

2 Pretraining Data

2.1 From Quark to Atom and Molecule

Our pretraining experiments use three data versions (Table 1): Quark comprises 9 datasets, 25,067 hours, and 4,200,222 training windows; Atom comprises 521 datasets, 49,628 hours, and 24,723,392 windows; and Molecule comprises 505 datasets, 79,910 hours, and 35,207,959 windows.

Table 1. Pretraining data summary.

Data version	Datasets	Hours	Training windows
Quark	9	25,067	4,200,222
Atom	521	49,628	24,723,392
Molecule	505	79,910	35,207,959

Quark forms the initial collection and is dominated by clinical recordings: TUH contributes 23,653 hours. The remaining sources are HMC, Alljoined, THINGS-EEG2, a self-collected dataset, SEED, SEED-IV, SEED-V, and SEED-VII.

To explore the limits of data scaling, we assembled Atom and Molecule by collecting publicly available EEG datasets as comprehensively as possible and screening them for pretraining. Atom broadens this clinical starting point to include more recording settings and tasks. Added sources include I-CARE, CHB-MIT, Siena, ISRUC S3, FACED, EEGMMIDB, SparrKULee, and THINGS-EEG1, together with additional collections from OpenNeuro, Zenodo, and MOABB.

Molecule builds on this broader source pool with revised signal preparation and source selection. For this expansion, we collected approximately 92,032 hours of EEG from 514 datasets and constructed 41,637,575 windows before quality screening. Table 2 compares recording hours for selected sources across all three versions; Appendix A provides the complete inventory with publication and dataset links.

Table 2. Selected pretraining datasets, sources, and recording hours. The full inventory appears in Table 5.

Dataset	Source	Quark (h)	Atom (h)	Molecule (h)
TUH	[13]	23,653	22,554	22,158
I-CARE	[14]	–	1,657	28,583
CHB-MIT	[15]	–	991	1,964
HMC	[16]	741	17	1,142
ON_ds004395	[17]	–	729	6,827
ON_ds005555	[18]	–	2,002	1,947
Alljoined	[19]	371	–	461
THINGS-EEG2	[20]	184	–	228
self-collected dataset		28	28	22
SEED	[21]	16	9	–
SEED-IV	[22]	18	9	18
SEED-V	[23]	13	9	13
SEED-VII	[24]	43	9	43

A dash denotes absence from a collection.

This source-level comparison shows that most of Molecule's additional recording time comes from ICU data. I-CARE increases from 1,657 hours in Atom to 28,583 hours in Molecule, accounting for about 89% of the net increase between the two versions. Molecule therefore combines a larger ICU component with cognitive, sleep, and other experimental recordings from the broader source pool (Figure 2). Appendix B gives the category totals behind this change in composition.

2.2 Signal preparation and window construction

Public EEG datasets use different sampling rates and electrode naming conventions. We resample recordings to 200 Hz, apply band-pass filtering with a 0.5 Hz lower cutoff, and map electrode names to a shared vocabulary for the encoder.

Recordings can also contain unusable channels and gaps. Channels that are constant or contain at least 1% non-finite samples are removed; the remaining signals are split into finite segments to retain usable EEG

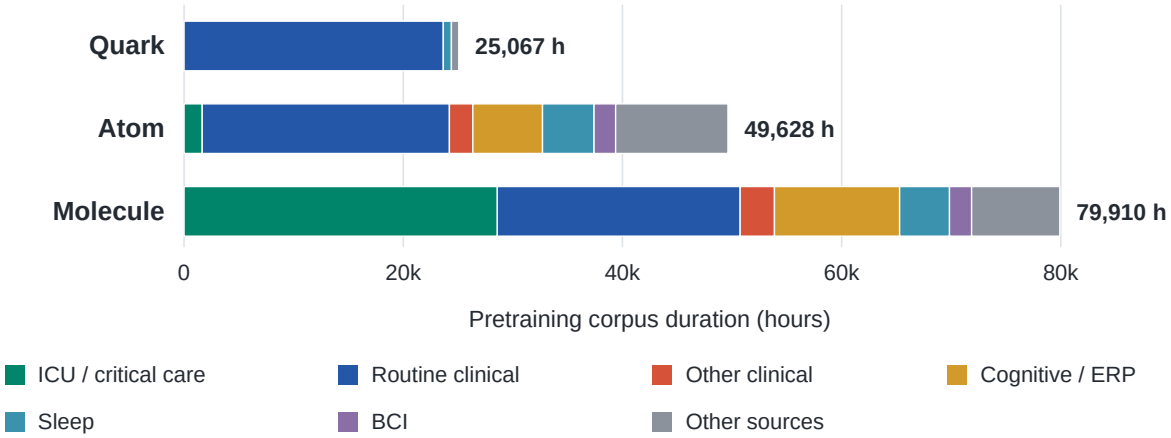


Figure 2. Pretraining data composition. Categories group sources by their predominant recording setting or task and are approximate. Molecule's increase in duration is concentrated in ICU recordings.

around gaps. Per-channel z-scoring within each recording or trial reduces amplitude scale differences across sources. Values are then clipped to ± 15 and stored in float16. Because normalization is performed during preparation, we use an input scale of 1.0; Quark and Atom retain 0.01 (Appendix B.3).

With a fixed window duration, recordings with more channels produce more tokens and require more attention memory. We therefore use non-overlapping windows of 4–16 seconds, adjusting duration to target 256 channel-by-time tokens. A 16-channel recording uses 16-second windows, while 64 channels use four seconds. Alljoined and THINGS-EEG2 retain their one-second trials to preserve trial boundaries. Tails of at least one second are zero-padded and retained with validity masks, preserving available EEG while marking padded positions.

2.3 Quality control and sampling

After preprocessing, windows still differ in signal quality and channel coverage. We remove windows with severe signal abnormalities or excessive padding to limit training on unreliable inputs, and apply the subject exclusion list separately. For Molecule, subject exclusions remove 687,187 windows and severe-quality screening removes 5,742,429, leaving 35,207,959 of the 41,637,575 constructed windows eligible for pretraining (Figure 3).

We retain windows with mild quality flags or few channels because they still contain usable EEG, using lower sampling weights to favor the main pool. The retained pool comprises 26,102,536 main windows with at least eight mapped channels, 8,180,196 mildly flagged windows, and 925,227 windows with fewer than eight mapped channels.

3 Scaling the Training Pipeline

3.1 Encoder sizes

We retain LaBraM's neural spectrum tokenizer architecture and masked-code prediction objective [1] while scaling the encoder from 5.8 million to 1.97 billion parameters (Table 3). Base, Large, and Huge retain the original LaBraM definitions. Giant extends the family to width 1,600 and 64 transformer layers. Model names combine the data version with B, L, H, or G for the corresponding encoder size. The training-pipeline extensions below support heterogeneous inputs and execution at larger data and model scales.

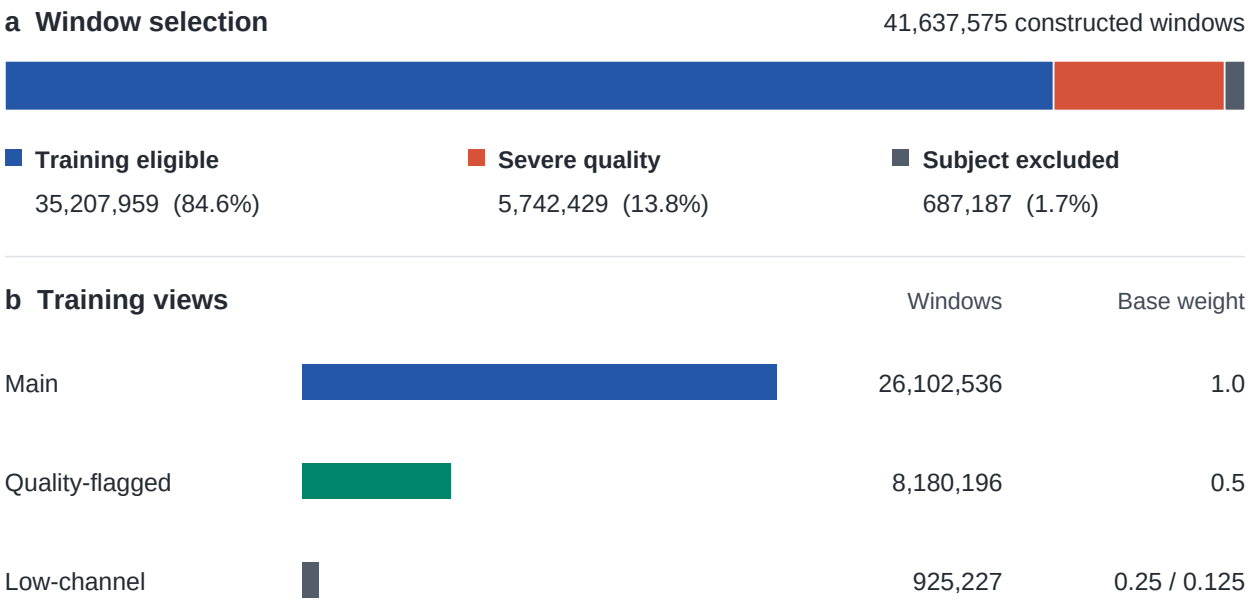


Figure 3. Molecule window screening and sampling. Screening excludes 15.4% of constructed windows; counts are shown before weighted sampling. Low-channel windows use weight 0.25, or 0.125 when they also carry quality flags. Source and provenance factors apply separately. Subject exclusions follow the earlier evaluation pool; see Appendix D for remaining overlap.

Table 3. Encoder configurations.

Size	Width	Depth	Heads	Parameters
Base	200	12	10	5.820M
Large	400	24	16	46.259M
Huge	800	48	16	369.236M
Giant	1,600	64	20	1.968B

3.2 Batching across datasets

Recordings from different datasets use different electrode sets and channel orders. LaBraM assigns spatial embeddings by electrode identity, so two windows with the same tensor shape can require different embedding indices. We therefore group windows in the Atom and Molecule loaders by channel count, patch count, channel signature, and ordered channel list. Samples in each batch then share the same dimensions and electrode ordering. For Molecule, we also reject unmapped pretraining channel names before batching.

Recording tails can end partway through a patch, so the resulting windows contain different amounts of observed EEG. In the Molecule encoder, sample-validity masks carry this information from window construction into attention and prediction. A fully padded patch contains no recorded EEG: we exclude it from attention keys and prediction masks and zero its output after the final normalization. Partially padded patches retain recorded samples, so we keep them in the sequence and weight their prediction loss in proportion to the valid sample counts. Appendix B describes the tokenizer's treatment of padded reconstruction targets.

3.3 Streaming and node-local storage

With tens of millions of windows, opening individual sample files would incur repeated file-access overhead. We therefore package Atom and Molecule in WebDataset tar shards so workers can read many windows from each archive as a stream. The two collections use 2,885 and 291,489 shards, respectively.

Because training revisits shards across epochs, each node caches downloaded shards, allowing workers to reuse local copies on subsequent reads. A least-recently-used policy evicts shards when space is needed. SQLite tracks the cache contents, and file locks coordinate workers requesting the same shard.

Persistent workers reduce process-startup overhead between epochs. Prefetching prepares upcoming batches while the current batch is processed, and pinned host memory supports transfer of the prepared batches to the accelerator.

3.4 Distributed training and precision

The training stack supports distributed data parallelism and fully sharded data parallelism (FSDP), with optional activation checkpointing to reduce activation memory. Molecule-H uses full parameter sharding; Quark-G was trained on MI300X GPUs. Table 4 lists GPU counts and global batches.

We replace the explicit attention path with PyTorch scaled dot-product attention, which selects kernels supported by the device, precision, and tensor layout. The training loop supports bf16 without loss scaling. Quark uses fp16 with GradScaler; Molecule extracts frozen-tokenizer targets in bf16.

4 Pretraining Configuration

The three collections differ in source composition and input preparation. We therefore use independently trained tokenizer weights for each collection and freeze them during encoder pretraining, so that target codes remain fixed as the encoder is optimized. Table 8 in the appendix lists the tokenizer settings and checkpoints used for target generation.

With these targets fixed, encoder optimization follows a common recipe: AdamW with weight decay 0.05, cosine learning-rate decay to 10^{-5}, drop path 0.1, and 50% masking. Within this recipe, Atom uses ten warmup epochs and gradient clipping at 1.0; Molecule uses five warmup epochs without gradient clipping. Table 4 gives the GPU allocation, global batch size, peak learning rate, and epoch schedule for each encoder.

Table 4. Encoder training configurations.

Model	GPUs	Global batch	Peak LR	Epochs
Quark-B	32	2,048	5×10^{-4}	50
Quark-L	32	512	3×10^{-4}	50
Quark-H	24	96	1.5×10^{-4}	50
Quark-G	64	1,024	1×10^{-4}	30
Atom-B	8	1,024	1.5×10^{-4}	50
Atom-L	16	1,024	1.5×10^{-4}	50
Atom-H	64	1,024	1.5×10^{-4}	50
Molecule-B	32	2,048	5×10^{-4}	50
Molecule-L	32	1,024	3×10^{-4}	30
Molecule-H	32	512	1×10^{-4}	50

The data-fraction experiments use this setup to examine how downstream transfer changes with data volume at a fixed encoder size. These runs retain 10%, 20%, or 50% of the corresponding window streams under a 50-epoch schedule. Larger fractions therefore increase both the number of training windows and the number of optimization updates. Appendices B and C give the subsampling and evaluation settings.

5 Evaluation

5.1 Evaluation setup

Following the evaluation framework of OmniEEG-Bench [25], we evaluate transfer on a subset of 23 classification tasks covering clinical EEG, motor imagery, evoked responses, emotion, sleep staging, and subject

traits. Full fine-tuning (Full-FT) measures performance after updating the encoder and task head; linear probing (LP) evaluates a readout from frozen features. All downstream evaluations use at most 40 training windows per subject–class pair, retaining all validation and test windows. Scores are averaged over seeds 42, 10, and 5. We compare overall performance by mean balanced accuracy, then use within-task ranks and individual task results to examine the breadth of the advantage. Appendix C lists task inputs and splits; Appendix C gives optimization settings.

The comparison includes 16 pretrained baseline encoders. The released LaBraM Base [1] provides a reference for scaling the original approach. The other baselines are BENDR [26], BIOT [27], CBraMod [2], BrainOmni [8], NeuroGPT [28], NeuroLM [5], EEGMamba [7], FEMBA Base [6], CSBrain [9], CodeBrain [10], REVE Base and Large [4], and LUNA Base, Large, and Huge [3]. Full-FT also includes EEGNet [29], EEGConformer [30], and a randomly initialized LaBraM Base as references for learning from scratch.

Unstarred extension comparisons use the last evaluated full-data checkpoints. Within each downstream fit, validation balanced accuracy selects the epoch. Seed42 test performance selects hyperparameters, which are then reused for seeds 10 and 5. For each model and task, LP reports the available Pooled or No-pool readout with the higher three-seed test mean, preferring Pooled on ties; SleepEDF uses Pooled.

5.2 Full fine-tuning

The scaled LaBraM encoders lead both mean balanced accuracy and average task rank under Full-FT (Figure 4). Atom-L* reaches 59.05%, exceeding REVE Large, the strongest baseline EEG foundation model, by 3.61 percentage points. The advantage also holds at the last evaluated checkpoints: Molecule-H reaches 58.02%, while Atom-L achieves 57.07% with 46.3 million parameters. Strong Full-FT performance therefore extends to the Large encoder, alongside the larger Huge models.

The task rankings show how this improvement extends across the benchmark. Extension configurations occupy the first eight positions by average task rank. Atom-L* leads at 8.74, and Molecule-H ranks second at 9.30; the strongest baseline by this measure, REVE Base, has an average rank of 14.72. Agreement between the two rankings supports a broad transfer advantage, with the remaining task-level differences examined in Section 5.4.

5.3 Linear probing

The gains carry over to frozen-feature transfer, with a different ordering among the leading encoders (Figure 5). Quark-G leads LP at 54.33% mean balanced accuracy, 1.24 percentage points above CodeBrain, the strongest baseline at 53.08% [10]. It also has the best average task rank at 8.63. Atom-H* and Atom-L* remain close in mean score at 54.25% and 54.09%. These checkpoints improve LP over their last evaluated counterparts, although they were selected using Full-FT scores.

Stronger LP performance does not necessarily imply stronger Full-FT performance. CodeBrain, for example, exceeds REVE Large in LP (53.08% versus 52.26%) and trails it in Full-FT (53.21% versus 55.44%). Across the 26 distinct pretrained encoders, 80 of 325 model pairs (24.6%) reverse order between the two settings (Appendix E.1). Full-FT also reaches a higher observed performance ceiling on this benchmark: the best evaluated configuration scores 59.05%, compared with 54.33% under LP.

Average pooling can suppress task-relevant information by collapsing patch features across channels and time. No-pool avoids this aggregation by retaining the ordered patches for a larger linear classifier, raising mean balanced accuracy from 44.74% to 49.61% (+4.87 points) across 26 models evaluated on the same 22 tasks. We further hypothesize that larger, stronger encoders learn more general representations whose alignment with a specific downstream task is more demanding. These representations may require more substantial task-specific reorganization to realize their potential. Full-FT can make such adjustments by updating the encoder, whereas LP relies on a fixed representation.

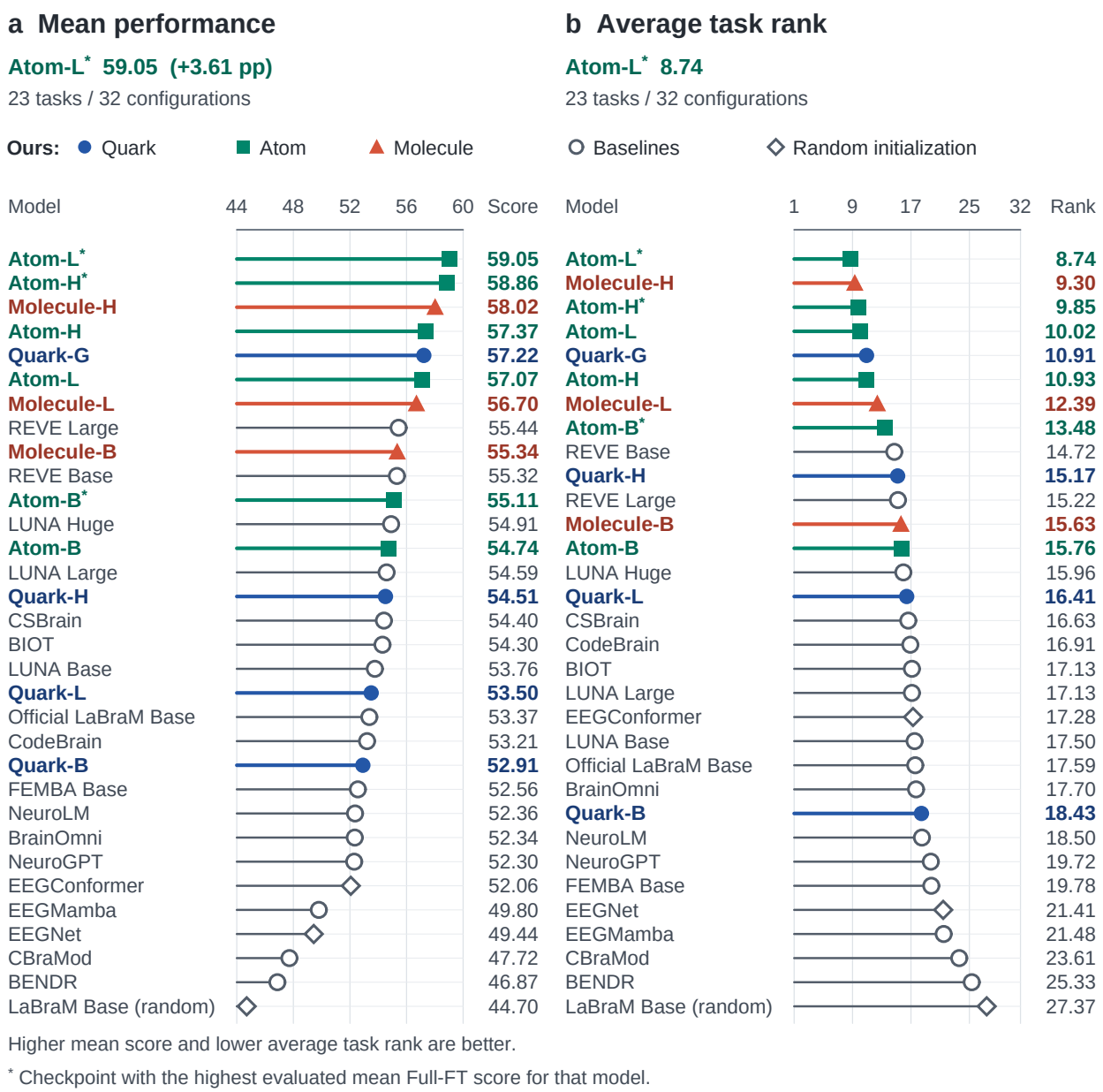


Figure 4. Full-FT ranked by mean balanced accuracy (left) and average task rank (right) across 32 configurations and 23 tasks. Stars mark checkpoints selected by mean Full-FT score: Atom-B* (ep19), Atom-L* (ep29), and Atom-H* (ep19). Ties receive average ranks. The strongest baseline by mean score is REVE Large [4].

5.4 Task-level differences

The individual task results show both the breadth and the concentration of the gains (Figure 6). Atom-L* exceeds REVE Large on 17 of 23 Full-FT tasks, with its largest gains on TUEV (+24.17 points), EEGDenoiseNet (+19.39), and HMC (+12.79). It trails on six tasks. For LP, Quark-G exceeds CodeBrain on 13 tasks and trails on ten. Its largest gains are on BETA SSVEP (+19.73 points), ISRUC S3 (+13.65), and Monitoring ErrP (+11.55), while TUSL and SEED remain weaknesses.

To assess the advantage over small models trained from scratch, we use the higher three-seed mean balanced accuracy of EEGNet and EEGConformer as the Full-FT reference for each task (Table 29). The released LaBraM Base exceeds it on 10 of 23 tasks. Among the 16 baseline EEG foundation models, REVE Base and Large lead with 12 wins each. Atom-L and Molecule-H each win on 16 tasks; Atom-L ties on one and trails on six, while Molecule-H trails on seven. Molecule-H retains all 12 wins of REVE Large and adds AD65, SEED-IV, HMC, and EEGDenoiseNet.

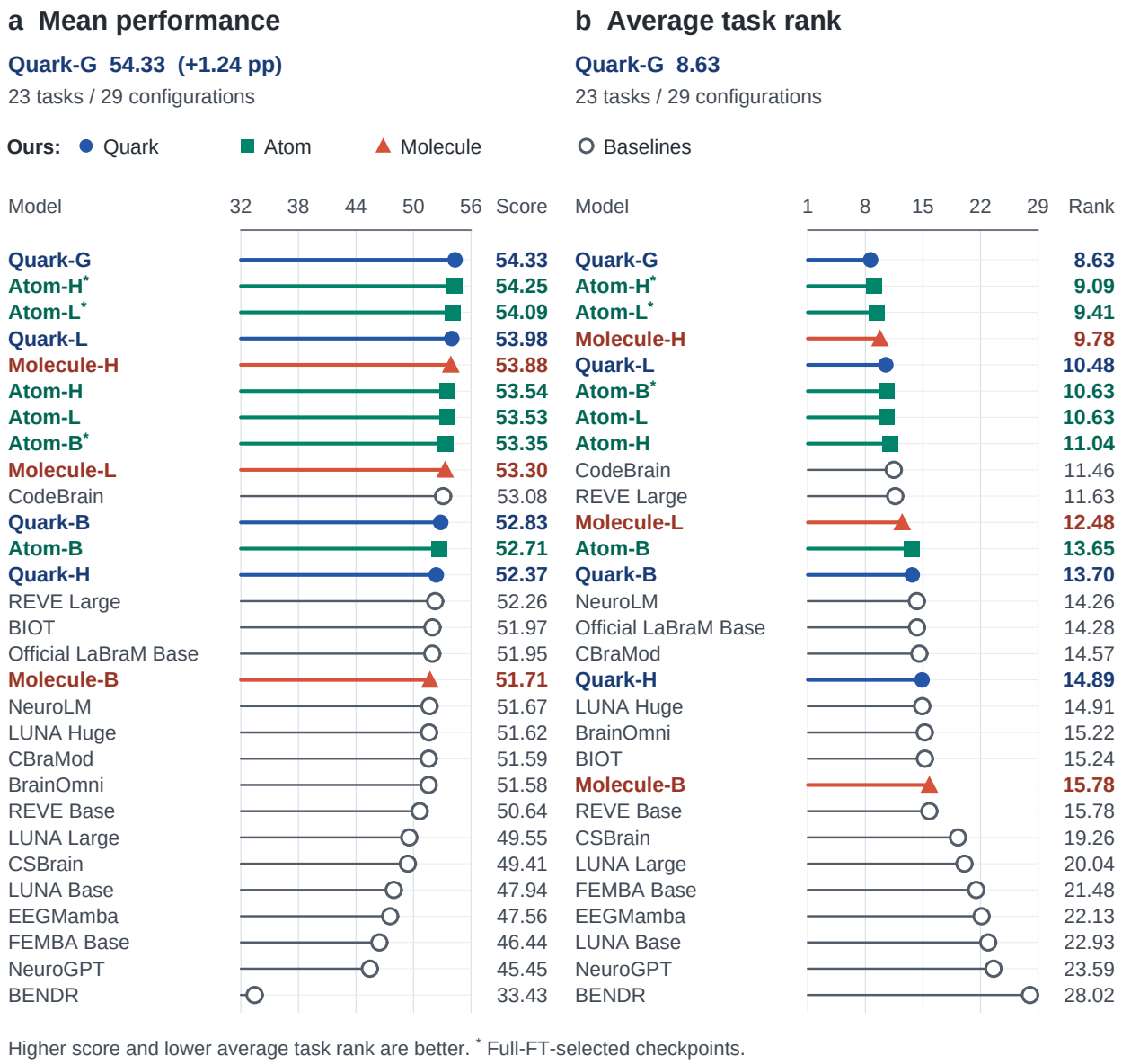


Figure 5. LP ranked by mean balanced accuracy (left) and average task rank (right) across 29 configurations of 26 pretrained models and 23 tasks. Stars use the Full-FT-selected EEG checkpoints. Each task uses the higher three-seed mean of the available Pooled and No-pool readouts.

5.5 Pretraining data, encoder size, and downstream performance

The model comparisons can also be viewed in terms of encoder size and pretraining hours (Figure 7). Atom-L* reaches 59.05% in Full-FT with 46.3M parameters and 49,628 hours, compared with 55.44% for REVE Large at 390M parameters and 61,415 hours. The last evaluated Atom-L checkpoint reaches 57.07%. Both comparisons favor Atom-L with a smaller encoder and fewer pretraining hours.

LP places the leading models much closer together despite large differences in size. Quark-G reaches 54.33% with 1.97B parameters and 25,067 pretraining hours; Quark-L reaches 53.98% with 46.3M parameters on the same collection. CodeBrain reaches 53.08% with 13.6M parameters and approximately 9,246 hours. Quark-L thus approaches Quark-G's LP score with a much smaller encoder, while CodeBrain remains competitive with a smaller model and corpus.

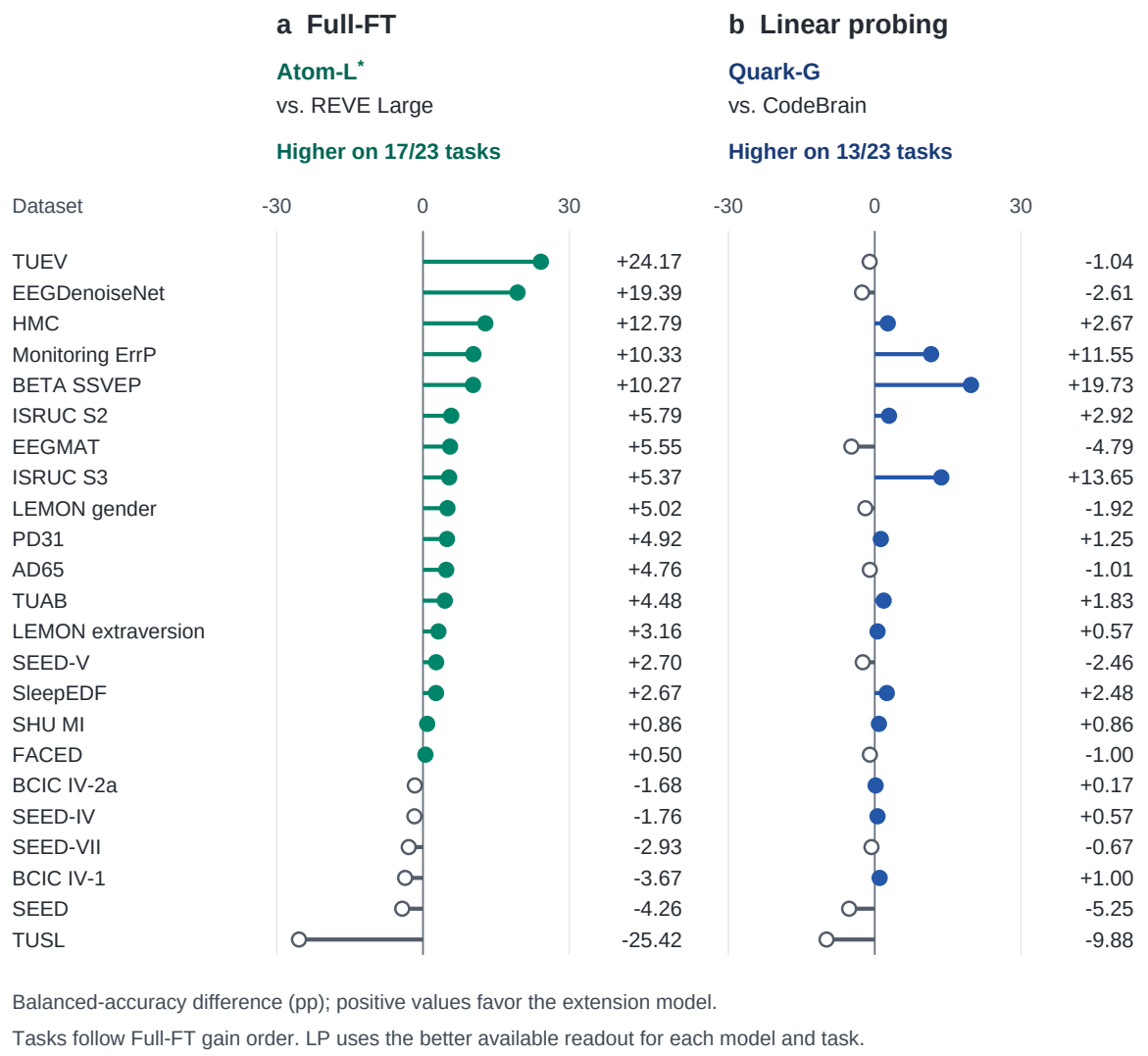


Figure 6. Task-level score differences between the leading extension configuration and baseline by mean score in each setting. Positive values favor the extension. The Full-FT panel uses Atom-L*; both panels cover 23 tasks.

6 Discussion

Scaling encoder capacity produces consistent Full-FT gains across the three data versions. The benefit of additional pretraining hours is more dependent on capacity, with fraction sweeps suggesting a saturation region for each model size. Training trajectories reveal a further limit: continued improvement on the pretraining objective can accompany declining downstream performance.

6.1 Model size

Increasing encoder capacity improves Full-FT within each pretraining collection (Figure 8). At the last evaluated checkpoints, Full-FT increases from Base to Huge by 1.60 points for Quark, 2.63 for Atom, and 2.68 for Molecule. Quark-G extends this trend to 57.22%, 2.71 points above Quark-H.

The same size comparisons give less consistent gains under LP. Quark falls from 53.98% at Large to 52.37% at Huge, Atom-L and Atom-H are nearly tied at 53.53% and 53.54%, and Molecule improves from Base to Huge. Capacity therefore benefits Full-FT more consistently than frozen-feature transfer in these runs. This pattern is consistent with the distinction in Section 5.3: increasing capacity can improve performance after task-specific adaptation without a corresponding gain in the linear accessibility of frozen features.

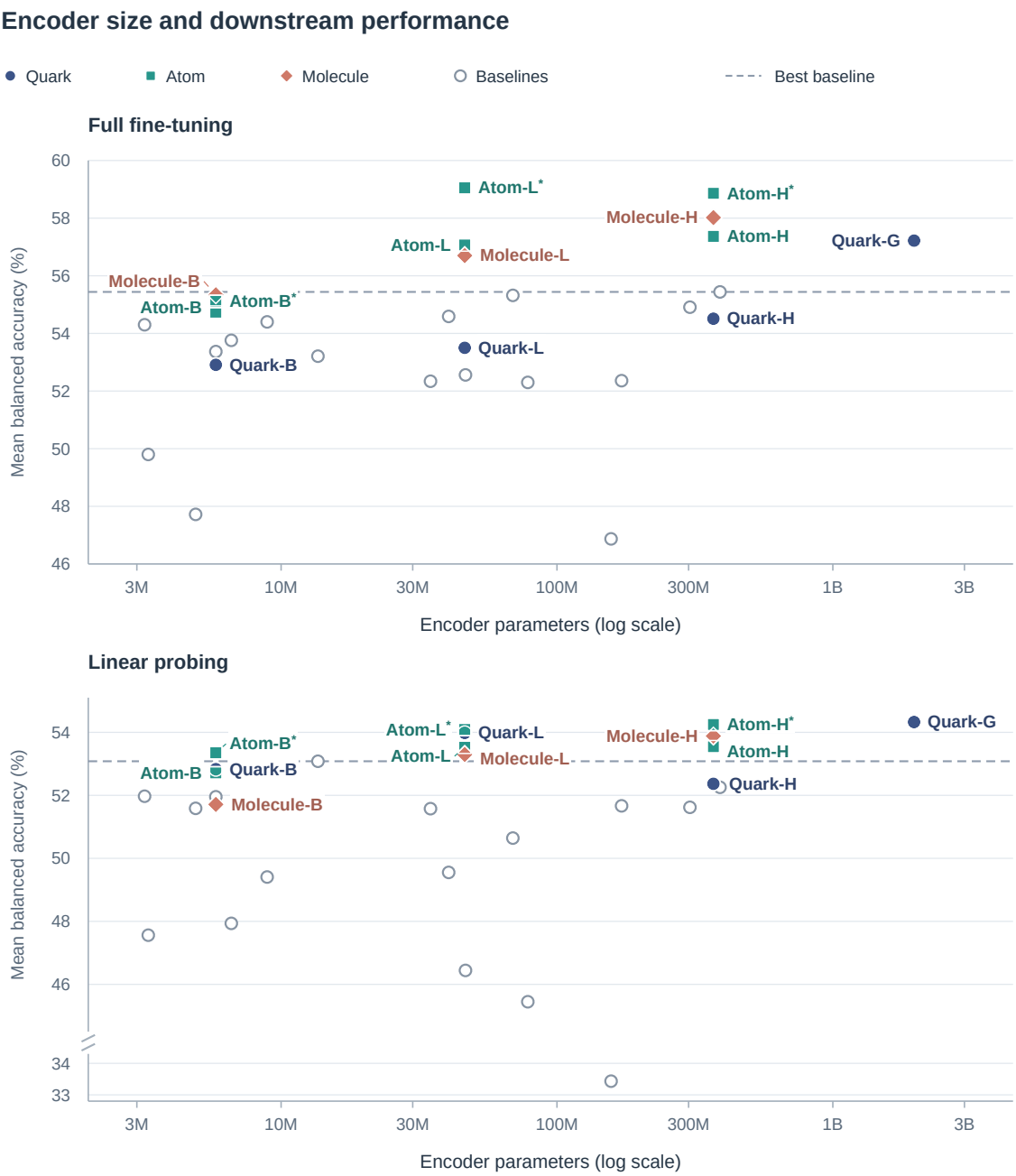


Figure 7. Encoder size versus Full-FT (top) and LP (bottom) performance, including the three Full-FT-selected Atom references (*). Dashed lines mark the highest baseline score; LP uses a broken vertical axis.

6.2 Pretraining data duration

At a fixed encoder size, the data versions provide a first comparison of pretraining corpus duration. Atom and Molecule exceed Quark in Full-FT at each shared size. Molecule improves over Atom at Base and Huge, and falls slightly below it at Large (56.70% versus 57.07%).

At Base size, random initialization and the released LaBraM provide two further references for these recipes. Quark-B, Atom-B, and Molecule-B reach 52.91%, 54.74%, and 55.34% Full-FT, all above the randomly initialized Base encoder at 44.70%. Against the released LaBraM Base at 53.37%, Atom-B and Molecule-B improve by 1.37 and 1.97 points, while Quark-B is 0.46 points lower.

The reduced-data Atom runs also outperform Quark at the same encoder sizes. With 10%, 20%, and 50% of Atom, corresponding to approximately 4,963, 9,926, and 24,814 hours, Atom-B reaches 54.54%, 53.40%, and 56.00% Full-FT, while Atom-L reaches 54.98%, 55.26%, and 57.15%. All six scores exceed their Quark

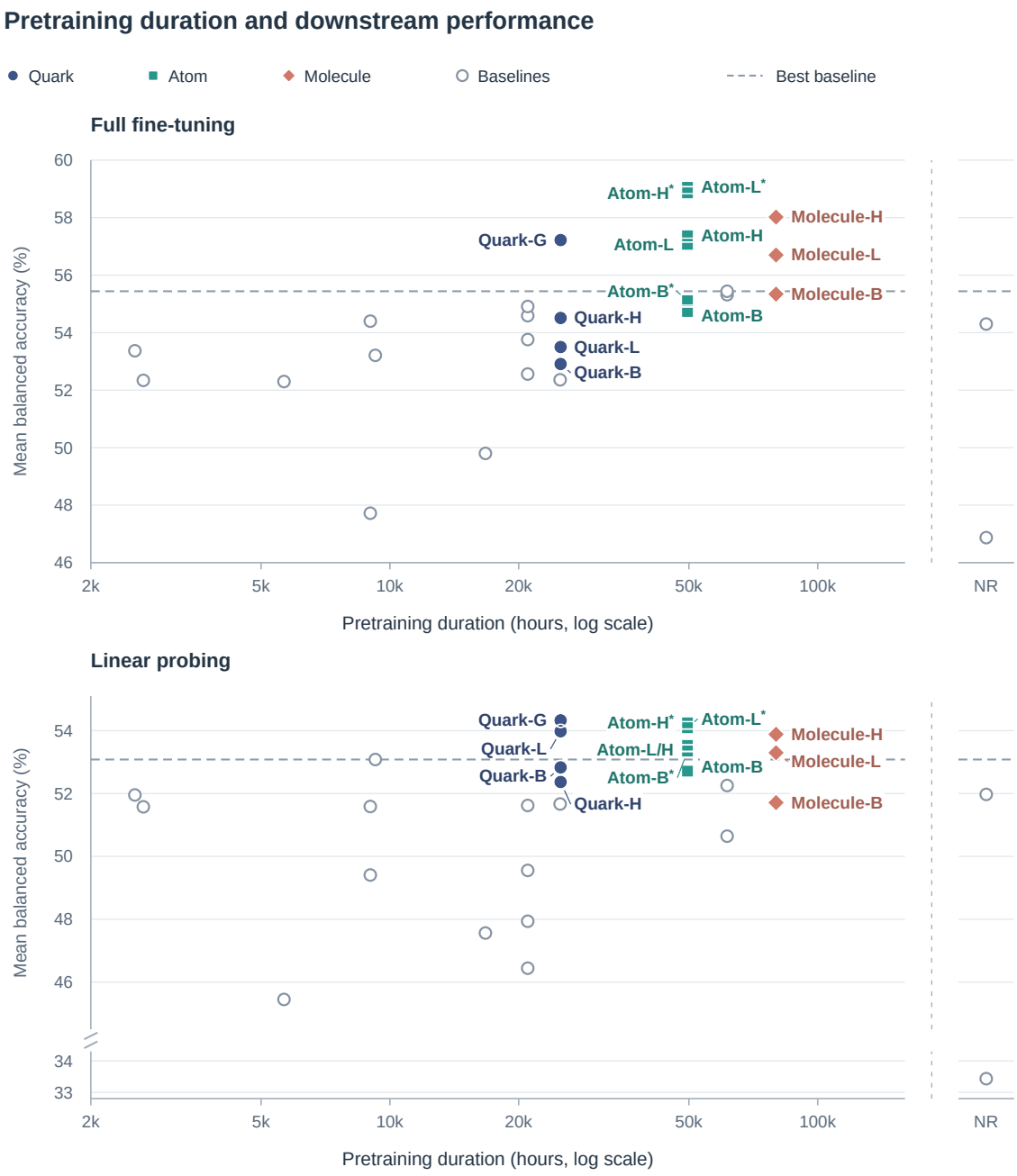


Figure 7. Pretraining duration versus Full-FT (top) and LP (bottom) performance. Published durations include estimates and lower bounds (EEG and MEG for BrainOmni). NR: not reported; Atom-L/H labels two nearly coincident LP points.

counterparts trained on 25,067 hours (52.91% for Base and 53.50% for Large). Atom-L approaches REVE with 10–20% of Atom; at 50%, both Atom sizes exceed REVE Base and Large (55.32% and 55.44%), pretrained on 61,415 hours [4].

A possible explanation is the breadth of Atom's pretraining sources. Its pool spans 521 datasets, compared with 9 for Quark and 92 for REVE [4]. Sampling windows across this broader pool may expose the encoder to more varied tasks and recording conditions, improving transfer even with fewer recording hours. This interpretation is consistent with OmniEEG-Bench, which reports an association between more pretraining datasets and stronger downstream rankings under LP [25].

Within each collection, the fraction sweeps also show that gains from additional hours can level off or reverse (Figure 9). Molecule-B is strongest at 20% data, approximately 16,000 hours: its Full-FT score falls from 56.35% there to 55.34% on the full collection, and LP falls from 53.09% to 51.71%. Atom-B peaks

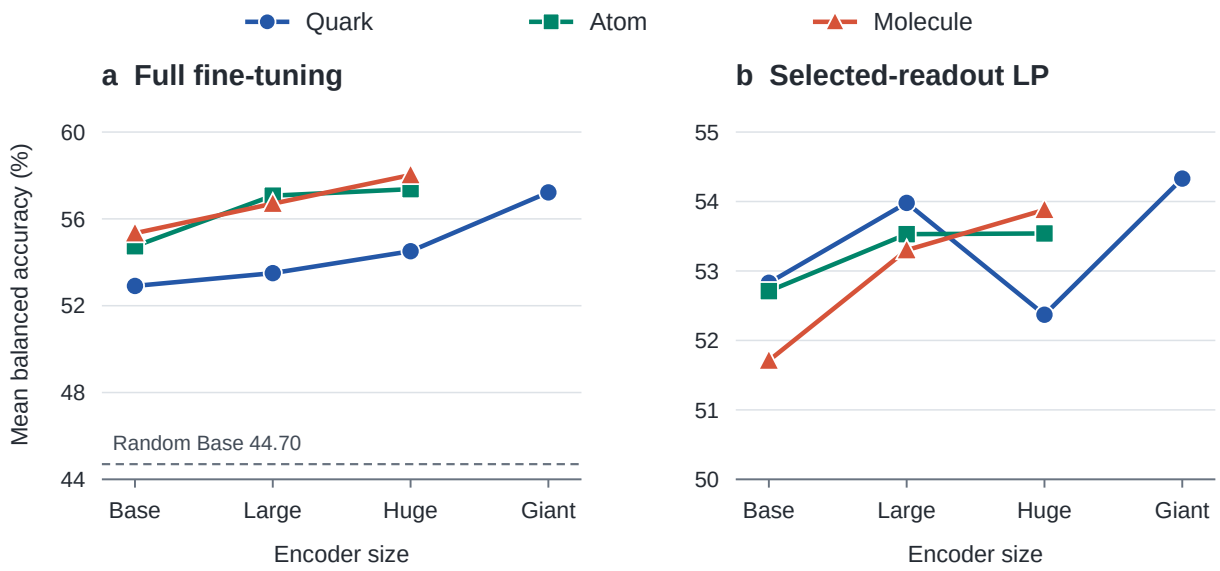


Figure 8. Performance by data version and encoder size at the last evaluated checkpoints. Full-FT improves with capacity within each version; selected-readout LP varies less consistently. The dashed line marks the randomly initialized Base Full-FT result.

later, at 50% data (approximately 24,800 hours), before both scores decline. Atom-L's Full-FT is nearly unchanged between 50% and 100% data, approximately 24,800–49,600 hours (57.15% versus 57.07%), although LP improves over this interval.

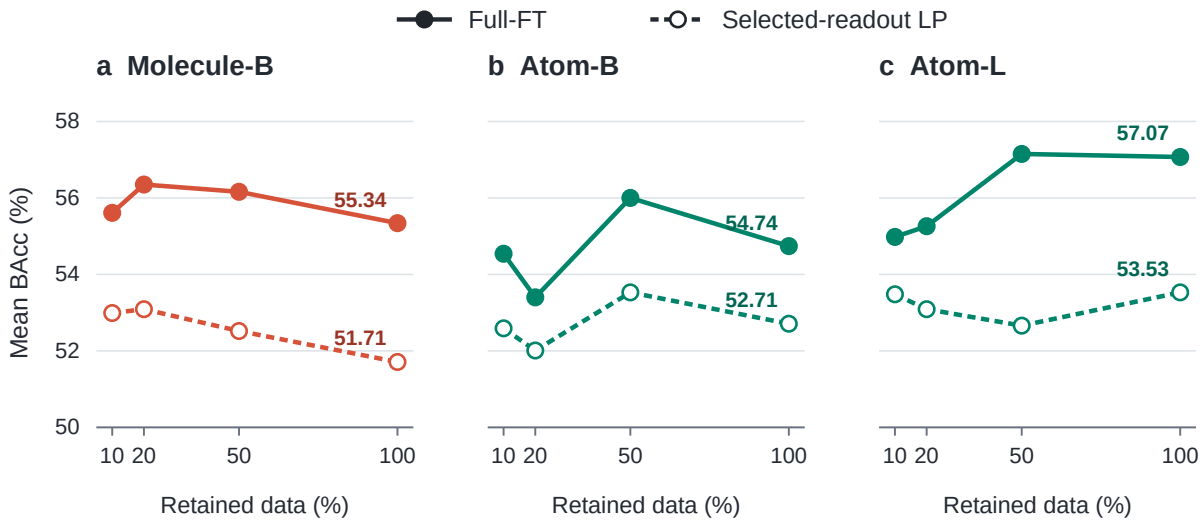


Figure 9. Full-FT and selected-readout LP across retained data fractions. Each series uses a 50-epoch schedule, so the fraction also changes the number of training updates. Full-data references use the same checkpoints as the main comparisons.

These sweeps use 50 epochs, so increasing the data fraction also increases optimizer updates. Hours are approximated by multiplying each collection's duration by the retained fraction. The curves therefore describe the combined effect of corpus size and training exposure under this schedule. Their turning regions motivate a closer look at how the useful amount of data depends on encoder capacity.

6.3 Capacity-dependent data saturation

At the sampled fractions, Base shows a decline where Large maintains similar Full-FT performance. This suggests that a fixed-size encoder has a saturation region and that larger encoders can benefit from more pretraining data. The broad spacing between fractions leaves the peak durations uncertain. A shared Full-FT response provides an estimate between the observed points, with duration normalized as $u = H/H_*(P)$:

$$\widehat{S}_i(H) = A_i + B\,\frac{2u}{1+u^2}, \qquad u = \frac{H}{H_*(P_i)}, \qquad H_*(P) = K\left(\frac{P}{5.8\times 10^6}\right)^{\eta}. \tag{1}$$

Here H is corpus duration without epoch repetition, P counts encoder parameters, and i identifies a model–collection series. The three series share amplitude $B \geq 0$ and response shape, with separate levels A_i. The response rises below $u = 1$ and falls above it, balancing two fixed power terms in u and u^{-1}.

Fitting the 12 Full-FT observations gives the following rounded peak-duration relation and its inverse:

$$\begin{aligned} H_*(P) &\approx 22{,}100\left(\frac{P}{5.8\times 10^6}\right)^{0.292}\ \text{hours}, \\ P_{\text{ref}}(H) &\approx 5.8\times 10^6\left(\frac{H}{22{,}100\ \text{hours}}\right)^{1/0.292}. \end{aligned} \tag{2}$$

The fit places the Base peak near 22,000 hours and the Large peak near 40,000 hours. This empirical scaling hypothesis links Base and Large; its inverse maps duration to encoder size along the same fitted peak relation. Atom-B's 20% point receives weight 0.25, and the Large peak is restricted to 24,814–49,628 hours. Appendix I gives the fitting assumptions and sensitivity checks.

6.4 The gap between pretraining and downstream evaluation

Lower pretraining loss does not imply stronger downstream performance. Full-FT can peak before the last evaluated checkpoint (Figure 10). Atom-L illustrates this directly: between ep29 and ep49, its loss decreases from 6.214 to 6.079 while Full-FT falls from 59.05% to 57.07%. Molecule-L shows a similar loss of about two Full-FT points during continued loss reduction, with declines on 16 of 23 tasks. The deterioration therefore extends across much of its downstream task profile.

The gap also appears in frozen features. Between ep9 and ep49, Molecule-B's loss falls from 8.322 to 7.949, while selected-readout LP declines from 52.98% to 51.71%. The aligned FT and LP trajectories in Figure 10 show that this mismatch occurs both when the encoder is frozen and when it is fully fine-tuned.

Further optimization can produce lower pretraining loss and worse downstream performance. Training loss alone is therefore insufficient for choosing a training duration or a transfer checkpoint; both choices require downstream evaluation. Appendix G provides the complete recorded trajectories.

7 A Shared Language-Model Readout for EEG Tasks

NeuroLM [5], UniMind [11], and GVG [12] explore unified EEG decoding across multiple tasks using language or multimodal language models. Inspired by these studies, we examine whether the gains from scaling LaBraM extend to a shared model trained jointly on all 23 tasks. This evaluation compares pretraining collections, EEG encoder capacities, and language-model sizes.

7.1 Connecting frozen EEG features to Qwen3

A shared Qwen3 readout [31] classifies all 23 tasks from frozen EEG features. Task prompts specify the problem and candidate labels. Channel-averaged features are standardized with training-split statistics, pooled or interpolated to 32 tokens, and mapped to Qwen3's input dimension by a two-layer MLP. Inference selects the highest-scoring textual class candidate.

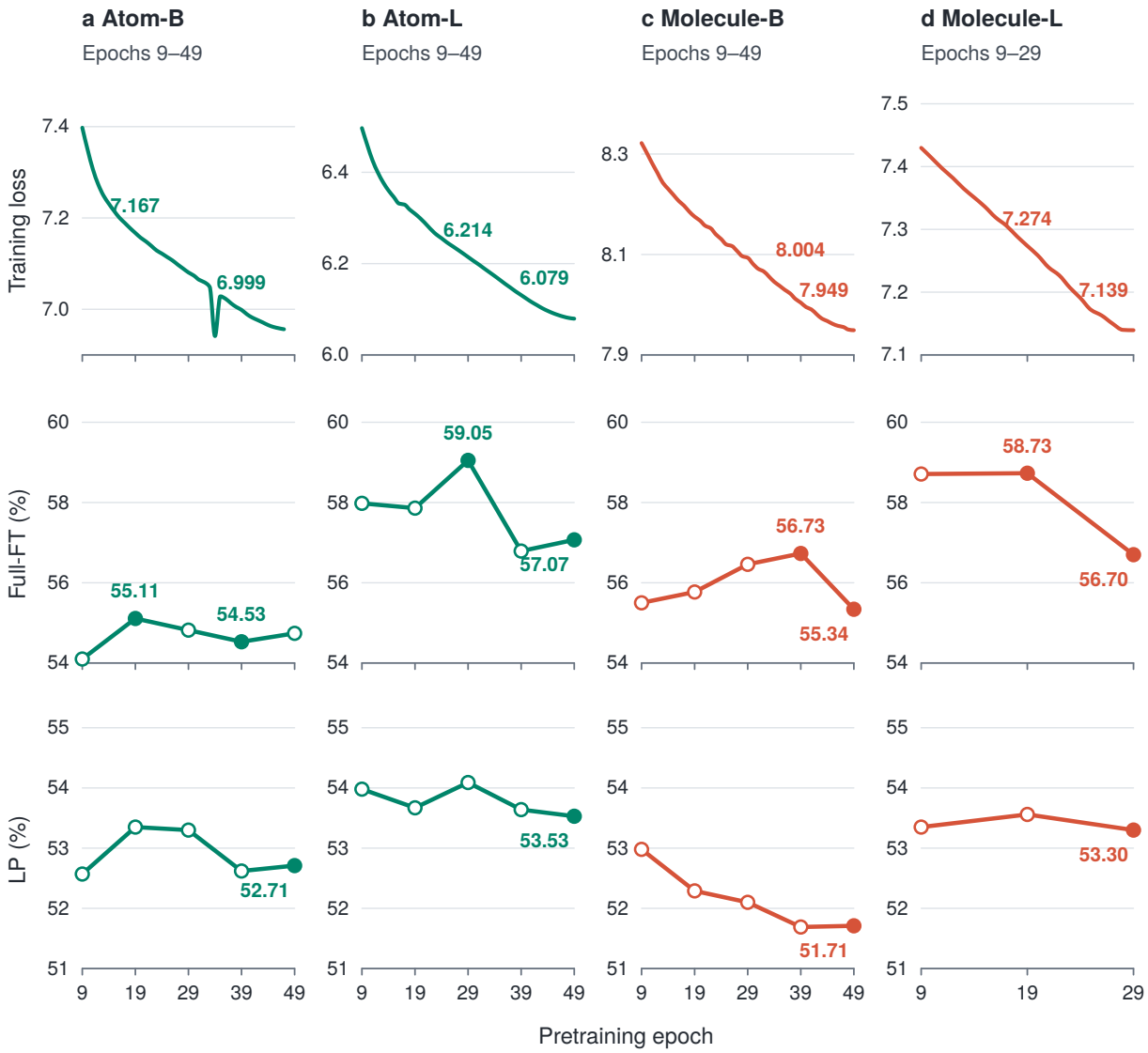


Figure 10. Pretraining loss (top), Full-FT (middle), and selected-readout LP (bottom) along four pretraining runs. Atom-B's loss ends at ep47; its downstream results extend to ep49.

Joint training updates the projector and LoRA adapters [32], keeping the EEG encoder and Qwen3 base weights frozen. Training uses the same 40-window subject–class cap as the single-task evaluations. Mean validation balanced accuracy selects one shared checkpoint per seed for full-test evaluation. Appendix H gives settings, splits, sample counts, and per-task scores with seed variability.

7.2 Data, encoder, and language-model scale

We pair ten EEG encoders with Qwen3-0.6B, 1.7B, 4B, 8B, 14B, and 32B: 60 configurations with three seeds each (Figure 11).

Pretraining data and EEG encoder capacity show complementary trends in joint classification. Moving from Quark's 25,067 hours to Atom's 49,628 hours improves performance in all 18 comparisons at matched EEG and Qwen3 sizes. With Molecule's 79,910 hours, further gains over Atom occur in five of six Huge comparisons and three of six Large comparisons, while Base performs worse throughout. The encoder comparisons reinforce this pattern: at a fixed pretraining collection and Qwen3 size, Huge (369M) outperforms Base (5.8M) in 16 of 18 comparisons. For Molecule, performance increases from Base through Large (46.3M) to Huge at all six Qwen3 sizes, with an average Base-to-Huge gain of 2.44 points.

Encoder orderings can change when moving from linear probing to joint EEG–LLM decoding. Quark-G



Figure 11. Joint classification across 23 EEG tasks with ten encoders and six Qwen3 sizes. Panels compare pretraining collections; curves distinguish encoder sizes. Scores are three-seed means on a shared vertical scale.

(1.97B) scores higher than Quark-H under Pooled LP (51.11% versus 48.05%), yet exceeds it in only one of the six Qwen3 settings. Larger language-model readouts also give inconsistent gains: every EEG encoder reaches its highest measured mean with Qwen3-8B or smaller, and Qwen3-32B exceeds 0.6B for only two of the ten encoders. Molecule-H with Qwen3-8B achieves the highest mean, $50.50 \pm 2.17\%$.

The Pooled/No-pool results in Section 5.3 offer a possible explanation for the limited return from a larger readout. Retaining individual patches improves LP over average pooling. Here, channel features are also averaged, although a temporal sequence of 32 tokens is retained. Qwen3 therefore lacks direct access to the individual channel tokens, and the frozen EEG encoder cannot adjust its features during joint training. Information retention and feature adaptation may thus limit how much additional language-model capacity helps.

The clearest gains in joint decoding come from combining larger EEG pretraining collections with greater encoder capacity. The LP comparisons suggest that preserving EEG structure and improving its adaptation to the shared readout are important for realizing these gains; increasing LLM parameters alone provides no consistent improvement.

8 Conclusion

Scaling pretraining data and encoder capacity while retaining LaBraM's core encoder and objective yields strong downstream representations. Across 23 tasks, Atom-L* leads Full-FT at 59.05% and Quark-G leads LP at 54.33%; both also lead average task rank. Against the better of EEGNet and EEGConformer on each task, Atom-L and Molecule-H each win on 16 of 23 Full-FT tasks, compared with at most 12 for a baseline EEG foundation model.

The benefits of capacity depend on downstream adaptation. Increasing encoder size improves Full-FT more consistently than LP, and higher LP scores do not guarantee higher Full-FT performance. No-pool's gains over average pooling suggest that retaining patch structure helps make frozen features useful to downstream tasks. The FT/LP ordering differences further support evaluating both linear accessibility and task-specific adaptation when judging the value of a pretrained encoder.

Data breadth and duration also shape the gains from pretraining. The reduced-data Atom results suggest that source diversity may improve data efficiency, while the fraction sweeps indicate a saturation region at fixed encoder capacity. The Full-FT fit for Base and Large gives $H_* \propto P^{0.292}$ and places the peaks near 22,000 and 40,000 hours, respectively, providing an empirical hypothesis for matching data duration to capacity (Equation 2). Continued optimization introduces a further limit: Full-FT and LP can decline while pretraining loss keeps falling. Training duration and checkpoint selection therefore require downstream evaluation.

Joint EEG–LLM decoding extends this dependence on adaptation to a shared readout. Larger EEG collections and encoders show complementary gains, while increasing Qwen3 size provides no consistent improvement. All ten encoders reach their highest mean with 8B or smaller, and Molecule-H with Qwen3-8B leads at 50.50%. The ordering changes between Pooled LP and joint decoding, together with the pooling analysis, suggest that information preservation and feature adaptation may limit the benefit of additional language-model capacity.

LaBraM's performance supports expanding an established EEG pretraining design as a route to stronger foundation models. Further progress is likely to come from matching data breadth and duration to encoder capacity, and improving how the resulting representations are preserved and adapted for downstream tasks.

References

[1] Wei-Bang Jiang, Li-Ming Zhao, and Bao-Liang Lu. Large brain model for learning generic representations with tremendous EEG data in BCI. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://arxiv.org/abs/2405.18765.

[2] Jiquan Wang, Sha Zhao, Zhiling Luo, Yangxuan Zhou, Haiteng Jiang, Shijian Li, Tao Li, and Gang Pan. CBraMod: A criss-cross brain foundation model for EEG decoding. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://arxiv.org/abs/2412.07236.

[3] Berkay Döner, Thorir Mar Ingolfsson, Luca Benini, and Yawei Li. LUNA: Efficient and topology-agnostic foundation model for EEG signal analysis. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems (NeurIPS)*, 2025. URL https://openreview.net/forum?id=uazfjnFL0G.

[4] Yassine El Ouahidi, Jonathan Lys, Philipp Thölke, Nicolas Farrugia, Bastien Pasdeloup, Vincent Gripon, Karim Jerbi, and Giulia Lioi. REVE: A foundation model for EEG – adapting to any setup with large-scale pretraining on 25,000 subjects. *arXiv preprint arXiv:2510.21585*, 2025.

[5] Wei-Bang Jiang, Yansen Wang, Bao-Liang Lu, and Dongsheng Li. NeuroLM: A universal multi-task foundation model for bridging the gap between language and EEG signals. In *The Thirteenth International Conference on Learning Representations*, 2025. URL https://arxiv.org/abs/2409.00101.

[6] Anna Tegon, Thorir Mar Ingolfsson, Xiaying Wang, Luca Benini, and Yawei Li. FEMBA: Efficient and scalable EEG analysis with a bidirectional Mamba foundation model. *arXiv preprint arXiv:2502.06438*, 2025.

[7] Jiquan Wang, Sha Zhao, Zhiling Luo, Yangxuan Zhou, Shijian Li, and Gang Pan. EEGMamba: An EEG foundation model with Mamba. *Neural Networks*, 192:107816, 2025. doi: 10.1016/j.neunet.2025.107816.

[8] Qinfan Xiao, Ziyun Cui, Chi Zhang, Siqi Chen, Wen Wu, Andrew Thwaites, Alexandra Woolgar, Bowen Zhou, and Chao Zhang. BrainOmni: A brain foundation model for unified EEG and MEG signals. *arXiv preprint arXiv:2505.18185*, 2025.

[9] Yuchen Zhou, Jiamin Wu, Zichen Ren, Zhouheng Yao, Weiheng Lu, Kunyu Peng, Qihao Zheng, Chunfeng Song, Wanli Ouyang, and Chao Gou. CSBrain: A cross-scale spatiotemporal brain foundation model for EEG decoding. In *Advances in Neural Information Processing Systems 38 (NeurIPS 2025)*, pages 96859–96904, 2025. doi: 10.52202/085713-2915.

[10] Jingying Ma, Feng Wu, Qika Lin, Yucheng Xing, Chenyu Liu, Ziyu Jia, and Mengling Feng. CodeBrain: Bridging decoupled tokenizer and multi-scale architecture for EEG foundation model. In *The Fourteenth International Conference on Learning Representations (ICLR)*, 2026. doi: 10.48550/arXiv.2506.09110.

[11] Weiheng Lu, Zhouheng Yao, Jiamin Wu, Pengyu Zhu, Yuchen Zhou, Weijian Mai, Qihao Zheng, Wanli Ouyang, and Chunfeng Song. UniMind: Unleashing the power of LLMs for unified multi-task brain decoding. *arXiv preprint arXiv:2506.18962*, 2025. URL https://arxiv.org/abs/2506.18962.

[12] Jun-Yu Pan, Yansen Wang, Enze Zhang, Bao-Liang Lu, Wei-Long Zheng, and Dongsheng Li. Visualizing the invisible: Generative visual grounding empowers universal EEG understanding in MLLMs. *arXiv preprint arXiv:2605.18172*, 2026. URL https://arxiv.org/abs/2605.18172.

[13] Iyad Obeid and Joseph Picone. The Temple University Hospital EEG Data Corpus. Frontiers in Neuroscience, 2016. URL https://doi.org/10.3389/fnins.2016.00196.

[14] Edilberto Amorim, Wei-Long Zheng, Mohammad M. Ghassemi, Mahsa Aghaeeaval, Pravinkumar Kandhare, Vishnu Karukonda, Jong Woo Lee, Susan T. Herman, Adithya Sivaraju, Nicolas Gaspard, Jeannette Hofmeijer, Michel J. A. M. van Putten, Reza Sameni, Matthew A. Reyna, Gari D. Clifford, and M. Brandon Westover. The International Cardiac Arrest Research Consortium Electroencephalography Database. Critical Care Medicine, 2023. URL https://doi.org/10.1097/CCM.0000000000006074.

[15] Ali Shoeb and John Guttag. Application of Machine Learning to Epileptic Seizure Detection. Proceedings of the 27th International Conference on Machine Learning, 2010. URL https://physionet.org/files/chbmit/1.0.0/shoeb-icml-2010.pdf.

[16] Diego Alvarez-Estevez and Roselyne M. Rijsman. Inter-database validation of a deep learning approach for automatic sleep scoring. PLOS ONE, 2021. URL https://doi.org/10.1371/journal.pone.0256111.

[17] Michael J. Kahana, Lynn J. Lohnas, M. Karl Healey, Ada Aka, Adam W. Broitman, Patrick Crutchley, Elizabeth Crutchley, Kylie H. Alm, Brandon S. Katerman, Nicole E. Miller, Joel R. Kuhn, Yuxuan Li, Nicole M. Long, Jonathan Miller, Madison D. Paron, Jesse K. Pazdera, Isaac Pedisich, Joseph H. Rudoler, and Christoph T. Weidemann. The Penn Electrophysiology of Encoding and Retrieval Study. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 50(9):1421–1443, 2024. doi: 10.1037/xlm0001319. URL https://doi.org/10.1037/xlm0001319.

[18] Eduardo López-Larraz, María Sierra-Torralba, Sergio Clemente, Galit Fierro, David Oriol, Javier Minguez, Luis Montesano, and Jens G. Klinzing. The Bitbrain Open Access Sleep (BOAS) dataset. OpenNeuro dataset, 2025. URL https://openneuro.org/datasets/ds005555.

[19] Jonathan Xu, Ugo Bruzadin Nunes, Wangshu Jiang, Samuel Ryther, Jordan Pringle, Paul S. Scotti, Arnaud Delorme, and Reese Kneeland. Alljoined-1.6M: A Million-Trial EEG-Image Dataset for Evaluating Affordable Brain-Computer Interfaces. arXiv preprint arXiv:2508.18571, 2025. URL https://arxiv.org/abs/2508.18571.

[20] Alessandro T. Gifford, Kshitij Dwivedi, Gemma Roig, and Radoslaw M. Cichy. A large and rich EEG dataset for modeling human visual object recognition. NeuroImage, 2022. URL https://doi.org/10.1016/j.neuroimage.2022.119754.

[21] Wei-Long Zheng and Bao-Liang Lu. Investigating Critical Frequency Bands and Channels for EEG-Based Emotion Recognition With Deep Neural Networks. IEEE Transactions on Autonomous Mental Development, 2015. URL https://doi.org/10.1109/TAMD.2015.2431497.

[22] Wei-Long Zheng, Wei Liu, Yifei Lu, Bao-Liang Lu, and Andrzej Cichocki. EmotionMeter: A Multimodal Framework for Recognizing Human Emotions. IEEE Transactions on Cybernetics, 2019. URL https://doi.org/10.1109/TCYB.2018.2797176.

[23] Wei Liu, Jie-Lin Qiu, Wei-Long Zheng, and Bao-Liang Lu. Comparing Recognition Performance and Robustness of Multimodal Deep Learning Models for Multimodal Emotion Recognition. IEEE Transactions on Cognitive and Developmental Systems, 2022. URL https://doi.org/10.1109/TCDS.2021.3071170.

[24] Wei-Bang Jiang, Xuan-Hao Liu, Wei-Long Zheng, and Bao-Liang Lu. SEED-VII: A Multimodal Dataset of Six Basic Emotions With Continuous Labels for Emotion Recognition. IEEE Transactions on Affective Computing, 2025. URL https://doi.org/10.1109/TAFFC.2024.3485057.

[25] Ziling Lu, Zongsheng Li, Xinke Shen, Kexin Lou, Yingyue Xin, Xiaoqi Chen, Shinan Wang, Xiang Chen, Jiahao Fan, Chenyu Huang, Xin Xu, Zhoujie Hou, Chen Wei, and Quanying Liu. OmniEEG-Bench: A standardized evaluation benchmark for EEG foundation models. *arXiv preprint arXiv:2606.00815*, 2026. URL https://arxiv.org/abs/2606.00815.

[26] Demetres Kostas, Stephane Aroca-Ouellette, and Frank Rudzicz. BENDR: Using transformers and a contrastive self-supervised learning task to learn from massive amounts of EEG data. *Frontiers in Human Neuroscience*, 15, 2021. URL https://arxiv.org/abs/2101.12037.

[27] Chaoqi Yang, M. Brandon Westover, and Jimeng Sun. BIOT: Cross-data biosignal learning in the wild. In *Advances in Neural Information Processing Systems*, 2023. URL https://arxiv.org/abs/2305.10351.

[28] Wenhui Cui, Woojae Jeong, Philipp Thölke, Takfarinas Medani, Karim Jerbi, Anand A. Joshi, and Richard M. Leahy. Neuro-GPT: Towards a foundation model for EEG. *arXiv preprint arXiv:2311.03764*, 2023.

[29] Vernon J. Lawhern, Amelia J. Solon, Nicholas R. Waytowich, Stephen M. Gordon, Chou P. Hung, and Brent J. Lance. EEGNet: A compact convolutional network for EEG-based brain-computer interfaces. *Journal of Neural Engineering*, 15(5):056013, 2018.

[30] Yonghao Song, Qingqing Zheng, Bingchuan Liu, and Xiaorong Gao. EEG Conformer: Convolutional transformer for EEG decoding and visualization. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 31:710–719, 2023. doi: 10.1109/TNSRE.2022.3230250.

[31] An Yang, Anfeng Li, Baosong Yang, Beichen Zhang, Binyuan Hui, Bo Zheng, Bowen Yu, Chang Gao, Chengen Huang, Chenxu Lv, Chujie Zheng, Dayiheng Liu, Fan Zhou, Fei Huang, Feng Hu, Hao Ge, Haoran Wei, Huan Lin, Jialong Tang, Jian Yang, Jianhong Tu, Jianwei Zhang, Jianxin Yang, Jiaxi Yang, Jing Zhou, Jingren Zhou, Junyang Lin, Kai Dang, Keqin Bao, Kexin Yang, Le Yu, Lianghao Deng, Mei Li, Mingfeng Xue, Mingze Li, Pei Zhang, Peng Wang, Qin Zhu, Rui Men, Ruize Gao, Shixuan Liu, Shuang Luo, Tianhao Li, Tianyi Tang, Wenbiao Yin, Xingzhang Ren, Xinyu Wang, Xinyu Zhang, Xuancheng Ren, Yang Fan, Yang Su, Yichang Zhang, Yinger Zhang, Yu Wan, Yuqiong Liu, Zekun Wang, Zeyu Cui, Zhenru Zhang, Zhipeng Zhou, and Zihan Qiu. Qwen3 technical report. *arXiv preprint arXiv:2505.09388*, 2025.

[32] Edward J Hu, Yelong Shen, Phillip Wallis, Zeyuan Allen-Zhu, Yuanzhi Li, Shean Wang, Lu Wang, and Weizhu Chen. LoRA: Low-rank adaptation of large language models. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=nZeVKeeFYf9.

[33] Michael Tangermann et al. Review of the BCI Competition IV. Frontiers in Neuroscience, 2012. URL https://doi.org/10.3389/fnins.2012.00055.

[34] Ji-Hoon Jeong, Jeong-Hyun Cho, Young-Eun Lee, Seo-Hyun Lee, Gi-Hwan Shin, Young-Seok Kweon, José del R. Millán, Klaus-Robert Müller, and Seong-Whan Lee. 2020 International brain–computer interface competition: A review. Frontiers in Human Neuroscience, 2022. URL https://doi.org/10.3389/fnhum.2022.898300.

[35] Igor Zyma, Sergii Tukaev, Ivan Seleznov, Ken Kiyono, Anton Popov, Mariia Chernykh, and Oleksii Shpenkov. Electroencephalograms during Mental Arithmetic Task Performance. Data, 2019. URL https://doi.org/10.3390/data4010014.

[36] Gerwin Schalk, Dennis J. McFarland, Thilo Hinterberger, Niels Birbaumer, and Jonathan R. Wolpaw. BCI2000: A General-Purpose Brain-Computer Interface (BCI) System. IEEE Transactions on Biomedical Engineering, 2004. URL https://doi.org/10.1109/TBME.2004.827072.

[37] Jingjing Chen, Xiaobin Wang, Chen Huang, Xin Hu, Xinke Shen, and Dan Zhang. A Large Finer-grained Affective Computing EEG Dataset. *Scientific Data*, 10(1):740, 2023. doi: 10.1038/s41597-023-02650-w. URL https://doi.org/10.1038/s41597-023-02650-w.

[38] Alireza Chamanzar, Sarah M Haigh, Pulkit Grover, and Marlene Behrmann. Abnormalities in cortical pattern of coherence in migraine detected using ultra high-density EEG. *Brain Communications*, 3(2):fcab061, 2021. doi: 10.1093/braincomms/fcab061. URL https://doi.org/10.1093/braincomms/fcab061.

[39] S. A. Nuiten, A. Canales-Johnson, L. Beerendonk, N. Nanuashvili, J. J. Fahrenfort, T. Bekinschtein, and S. van Gaal. Preserved sensory processing but hampered conflict detection when stimulus input is task-irrelevant. *Elife*, 10:e64431, 2021. doi: 10.7554/elife.64431. URL https://doi.org/10.7554/elife.64431.

[40] Stefan Wiens, Rasmus Eklund, Malina Szychowska, Alexander Miloff, Danielle Cosme, Stephen Pierzchajlo, and Per Carlbring. Electrophysiological correlates of in vivo and virtual reality exposure therapy in spider phobia. *Psychophysiology*, 59(12):e14117, 2022. doi: 10.1111/psyp.14117. URL https://doi.org/10.1111/psyp.14117.

[41] Abigail Dickinson, Shafali Jeste, and Elizabeth Milne. Electrophysiological signatures of brain aging in autism spectrum disorder. *Cortex*, 148:139–151, 2022. doi: 10.1016/j.cortex.2021.09.022. URL https://doi.org/10.1016/j.cortex.2021.09.022.

[42] Jeehyun Kim, Xiyuan Jiang, Dylan Forenzo, Yixuan Liu, Nancy Anderson, Carol M. Greco, and Bin He. Immediate effects of short-term meditation on sensorimotor rhythm-based brain–computer interface performance. *Frontiers in Human Neuroscience*, 16:1019279, 2022. doi: 10.3389/fnhum.2022.1019279. URL https://doi.org/10.3389/fnhum.2022.1019279.

[43] J. E. Gonzalez, N. Nieto, P. Brusco, A. Gravano, and J. E. Kamienkowski. Speech-induced suppression during natural dialogues. *Commun Biol*, 7(1):291, 2024. doi: 10.1038/s42003-024-05945-9. URL https://doi.org/10.1038/s42003-024-05945-9.

[44] M. Gu, W. Pei, X. Gao, and Y. Wang. An open dataset for human SSVEPs in the frequency range of 1-60 Hz. *Sci Data*, 11(1):196, 2024. doi: 10.1038/s41597-024-03023-7. URL https://doi.org/10.1038/

s41597-024-03023-7.

[45] Dylan Forenzo, Yixuan Liu, Jeehyun Kim, Yidan Ding, Taehyung Yoon, and Bin He. Integrating Simultaneous Motor Imagery and Spatial Attention for EEG-BCI Control. *IEEE Transactions on Biomedical Engineering*, 71(1):282–294, 2024. doi: 10.1109/tbme.2023.3298957. URL https://doi.org/10.1109/tbme.2023.3298957.

[46] Y. Chen, S. Fazli, and C. Wallraven. An EEG Dataset of Neural Signatures in a Competitive Two-Player Game Encouraging Deceptive Behavior. *Sci Data*, 11(1):389, 2024. doi: 10.1038/s41597-024-03234-y. URL https://doi.org/10.1038/s41597-024-03234-y.

[47] Dylan Forenzo, Hao Zhu, Jenn Shanahan, Jaehyun Lim, and Bin He. Continuous tracking using deep learning-based decoding for noninvasive brain–computer interface. *PNAS Nexus*, 3(4):pgae145, 2024. doi: 10.1093/pnasnexus/pgae145. URL https://doi.org/10.1093/pnasnexus/pgae145.

[48] Joshua Kosnoff, Kai Yu, Chang Liu, and Bin He. Transcranial Focused Ultrasound to V5 Enhances Human Visual Motion Brain-Computer Interface by Modulating Feature-Based Attention. bioRxiv preprint, 2023. URL https://doi.org/10.1101/2023.09.04.556252.

[49] Sam Watson, Torsten Dau, and Jens Hjortkjær. Task-dependent Modulation Masking of 4 Hz Envelope Following Responses. *Journal of Cognitive Neuroscience*, 37(7):1189–1201, 2025. doi: 10.1162/jocn_a_02309. URL https://doi.org/10.1162/jocn_a_02309.

[50] Dylan Forenzo, Yisha Zhang, George F. Wittenberg, and Bin He. Continuous Reaching and Grasping with a BCI Controlled Robotic Arm in Healthy and Stroke-Affected Individuals. medRxiv preprint, 2025. URL https://doi.org/10.1101/2025.04.16.25325551.

[51] W. Chang, W. Kong, G. Yan, R. Lv, K. Du, M. T. Sadiq, B. Guo, R. Yin, and X. Liu. A multi-paradigm EEG dataset for studying upper limb rehabilitation exercises. *Sci Data*, 12(1):1877, 2025. doi: 10.1038/s41597-025-06147-6. URL https://doi.org/10.1038/s41597-025-06147-6.

[52] Elena Bondi, Yidan Ding, Yisha Zhang, Eleonora Maggioni, and Bin He. Investigating the neurovascular coupling across multiple motor execution and imagery conditions: a whole-brain EEG-informed fMRI analysis. *NeuroImage*, 317:121311, 2025. doi: 10.1016/j.neuroimage.2025.121311. URL https://doi.org/10.1016/j.neuroimage.2025.121311.

[53] A. Malekmohammadi and G. Cheng. Music Familiarization Elicits Functional Connectivity Between Right Frontal/Temporal and Parietal Areas in the Theta and Alpha Bands. *Brain Topogr*, 38(1):2, 2024. doi: 10.1007/s10548-024-01081-z. URL https://doi.org/10.1007/s10548-024-01081-z.

[54] Rasmus Eklund and Stefan Wiens. Visual awareness negativity is an early neural correlate of awareness: A preregistered study with two Gabor sizes. *Cognitive, Affective, & Behavioral Neuroscience*, 18(1):176–188, 2018. doi: 10.3758/s13415-018-0562-z. URL https://doi.org/10.3758/s13415-018-0562-z.

[55] Stefan Wiens, Malina Szychowska, Rasmus Eklund, and Erik van Berlekom. Cascade and no-repetition rules are comparable controls for the auditory frequency mismatch negativity in oddball tasks. *Psychophysiology*, 56(1):e13280, 2019. doi: 10.1111/psyp.13280. URL https://doi.org/10.1111/psyp.13280.

[56] Z. Cao, C. H. Chuang, J. K. King, and C. T. Lin. Multi-channel EEG recordings during a sustained-attention driving task. *Sci Data*, 6(1):19, 2019. doi: 10.1038/s41597-019-0027-4. URL https://doi.org/10.1038/s41597-019-0027-4.

[57] Stefan Wiens, Erik van Berlekom, Malina Szychowska, and Rasmus Eklund. Visual Perceptual Load Does Not Affect the Frequency Mismatch Negativity. *Frontiers in Psychology*, 10:1970, 2019. doi: 10.3389/fpsyg.2019.01970. URL https://doi.org/10.3389/fpsyg.2019.01970.

[58] Stefan Wiens, Malina Szychowska, Rasmus Eklund, and Mats E. Nilsson. Data on the auditory duration mismatch negativity for different sound pressure levels and visual perceptual loads. *Data in Brief*, 11:159–164, 2017. doi: 10.1016/j.dib.2017.02.007. URL https://doi.org/10.1016/j.dib.2017.02.007.

[59] Malina Szychowska, Rasmus Eklund, Mats E. Nilsson, and Stefan Wiens. Effects of sound pressure level and visual perceptual load on the auditory mismatch negativity. *Neuroscience Letters*, 640:37–41, 2017. doi: 10.1016/j.neulet.2017.01.001. URL https://doi.org/10.1016/j.neulet.2017.01.001.

[60] Sirvan Khalighi, Teresa Sousa, José Moutinho Santos, and Urbano Nunes. ISRUC-Sleep: A comprehensive public dataset for sleep researchers. Computer Methods and Programs in Biomedicine, 2016. URL https://doi.org/10.1016/j.cmpb.2015.10.013.

[61] Alexandre Barachant. *Commande robuste d'un effecteur par une interface cerveau machine EEG asynchrone*. PhD thesis, Université de Grenoble, 2012. URL https://theses.hal.science/tel-01196752.

[62] L. Korczowski, E. Ostaschenko, A. Andreev, G. Cattan, P. L. C. Rodrigues, V. Gautheret, and M. Congedo.

Brain Invaders calibration-less P300-based BCI using dry EEG electrodes Dataset (BI2014a). HAL, report hal-02171575, 2019. URL https://hal.science/hal-02171575.

[63] Robert Leeb, Felix Lee, Claudia Keinrath, Reinhold Scherer, Horst Bischof, and Gert Pfurtscheller. Brain–Computer Communication: Motivation, Aim, and Impact of Exploring a Virtual Apartment. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 15(4):473–482, 2007. doi: 10.1109/TNSRE.2007.906956. URL https://doi.org/10.1109/TNSRE.2007.906956.

[64] A. Riccio, L. Simione, F. Schettini, A. Pizzimenti, M. Inghilleri, M. O. Belardinelli, and D. Mattia. Attention and P300-based BCI performance in people with amyotrophic lateral sclerosis. *Frontiers in Human Neuroscience*, 7:732, 2013. doi: 10.3389/fnhum.2013.00732. URL https://doi.org/10.3389/fnhum.2013.00732.

[65] P. Aricò, F. Aloise, F. Schettini, S. Salinari, D. Mattia, and F. Cincotti. Influence of P300 latency jitter on event related potential-based brain–computer interface performance. *Journal of Neural Engineering*, 11(3):035008, 2014. doi: 10.1088/1741-2560/11/3/035008. URL https://doi.org/10.1088/1741-2560/11/3/035008.

[66] J. Faller, C. Vidaurre, T. Solis-Escalante, C. Neuper, and R. Scherer. Autocalibration and Recurrent Adaptation: Towards a Plug and Play Online ERD-BCI. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 20(3):313–319, 2012. doi: 10.1109/TNSRE.2012.2189584. URL https://doi.org/10.1109/TNSRE.2012.2189584.

[67] C. Guger, S. Daban, E. Sellers, C. Holzner, G. Krausz, R. Carabalona, F. Gramatica, and G. Edlinger. How many people are able to control a P300-based brain–computer interface (BCI)? *Neuroscience Letters*, 462(1):94–98, 2009. doi: 10.1016/j.neulet.2009.06.045. URL https://doi.org/10.1016/j.neulet.2009.06.045.

[68] R. Scherer, J. Faller, E. V. C. Friedrich, E. Opisso, U. Costa, A. Kübler, and G. R. Müller-Putz. Individually Adapted Imagery Improves Brain-Computer Interface Performance in End-Users with Disability. *PLOS ONE*, 10(5):e0123727, 2015. doi: 10.1371/journal.pone.0123727. URL https://doi.org/10.1371/journal.pone.0123727.

[69] H. Cho, M. Ahn, S. Ahn, M. Kwon, and S. C. Jun. EEG datasets for motor imagery brain computer interface. *GigaScience*, 2017. doi: 10.1093/gigascience/gix034. URL https://doi.org/10.1093/gigascience/gix034.

[70] Pauline Dreyer, Aline Roc, Léa Pillette, Sébastien Rimbert, and Fabien Lotte. A large EEG database with users' profile information for motor imagery brain-computer interface research. *Scientific Data*, 10(1):580, 2023. doi: 10.1038/s41597-023-02445-z. URL https://doi.org/10.1038/s41597-023-02445-z.

[71] U. Hoffmann, J.-M. Vesin, T. Ebrahimi, and K. Diserens. An efficient P300-based brain–computer interface for disabled subjects. *Journal of Neuroscience Methods*, 2008. doi: 10.1016/j.jneumeth.2007.03.005. URL https://doi.org/10.1016/j.jneumeth.2007.03.005.

[72] Min-Ho Lee, O-Yeon Kwon, Yong-Jeong Kim, Hong-Kyung Kim, Young-Eun Lee, John Williamson, Siamac Fazli, and Seong-Whan Lee. EEG dataset and OpenBMI toolbox for three BCI paradigms: an investigation into BCI illiteracy. *GigaScience*, 8(5):giz002, 2019. doi: 10.1093/gigascience/giz002. URL https://doi.org/10.1093/gigascience/giz002.

[73] Haijie Liu, P. Wei, H. Wang, et al. An EEG motor imagery dataset for brain computer interface in acute stroke patients. *Scientific Data*, 11:131, 2024. doi: 10.1038/s41597-023-02787-8. URL https://doi.org/10.1038/s41597-023-02787-8.

[74] V. P. Oikonomou, G. Liaros, K. Georgiadis, E. Chatzilari, K. Adam, S. Nikolopoulos, and I. Kompatsiaris. Comparative evaluation of state-of-the-art algorithms for SSVEP-based BCIs. arXiv preprint arXiv:1602.00904, 2016. URL https://doi.org/10.48550/arXiv.1602.00904.

[75] Masaki Nakanishi, Yijun Wang, Yu-Te Wang, and Tzyy-Ping Jung. A Comparison Study of Canonical Correlation Analysis Based Methods for Detecting Steady-State Visual Evoked Potentials. *PLOS ONE*, 10(10):e0140703, 2015. doi: 10.1371/journal.pone.0140703. URL https://doi.org/10.1371/journal.pone.0140703.

[76] Patrick Ofner, Andreas Schwarz, Joana Pereira, and Gernot R. Müller-Putz. Upper limb movements can be decoded from the time-domain of low-frequency EEG. *PLOS ONE*, 12(8):e0182578, 2017. doi: 10.1371/journal.pone.0182578. URL https://doi.org/10.1371/journal.pone.0182578.

[77] Robin Tibor Schirrmeister et al. Deep learning with convolutional neural networks for EEG decoding and visualization. *Human Brain Mapping*, 38(11):5391–5420, 2017. doi: 10.1002/hbm.23730. URL https://doi.org/10.1002/hbm.23730.

[78] J. Shin, A. von Lühmann, B. Blankertz, D. W. Kim, J. Jeong, H. J. Hwang, and K. R. Müller. Open Access Dataset for EEG+NIRS Single-Trial Classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(10):1735–1745, 2017. doi: 10.1109/TNSRE.2016.2628057. URL https://doi.org/10.1109/TNSRE.2016.2628057.

[79] Yijun Wang, Xiaogang Chen, Xiaorong Gao, and Shangkai Gao. A Benchmark Dataset for SSVEP-Based Brain–Computer Interfaces. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 25(10): 1746–1752, 2017. doi: 10.1109/TNSRE.2016.2627556. URL https://doi.org/10.1109/TNSRE.2016.2627556.

[80] Weibo Yi, Shuang Qiu, Kun Wang, Hongzhi Qi, Lixin Zhang, Peng Zhou, Feng He, and Dong Ming. Evaluation of EEG Oscillatory Patterns and Cognitive Process during Simple and Compound Limb Motor Imagery. *PLOS ONE*, 9(12):e0114853, 2014. doi: 10.1371/journal.pone.0114853. URL https://doi.org/10.1371/journal.pone.0114853.

[81] Wajid Mumtaz, Likun Xia, Mohd Azhar Mohd Yasin, Syed Saad Azhar Ali, and Aamir Saeed Malik. A wavelet-based technique to predict treatment outcome for Major Depressive Disorder. PLOS ONE, 2017. URL https://doi.org/10.1371/journal.pone.0171409.

[82] Rui Li, Le-Dian Liu, and Bao-Liang Lu. Discrimination of Decision Confidence Levels from EEG Signals. In *2021 10th International IEEE/EMBS Conference on Neural Engineering (NER)*, pages 946–949, 2021. doi: 10.1109/NER49283.2021.9441086. URL https://doi.org/10.1109/NER49283.2021.9441086.

[83] Wei-Bang Jiang, Li-Ming Zhao, Ping Guo, and Bao-Liang Lu. Discriminating Surprise and Anger from EEG and Eye Movements with a Graph Network. In *2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM)*, pages 1353–1357, 2021. doi: 10.1109/BIBM52615.2021.9669637. URL https://bcmi.sjtu.edu.cn/~blu/papers/2021/2021-13.pdf.

[84] Benjamin Blankertz, Guido Dornhege, Matthias Krauledat, Klaus-Robert Müller, and Gabriel Curio. The non-invasive Berlin Brain–Computer Interface: Fast acquisition of effective performance in untrained subjects. *NeuroImage*, 37(2):539–550, 2007. doi: 10.1016/j.neuroimage.2007.01.051. URL https://doi.org/10.1016/j.neuroimage.2007.01.051.

[85] Rui Li, Yiting Wang, and Bao-Liang Lu. Measuring Decision Confidence Levels from EEG Using a Spectral-Spatial-Temporal Adaptive Graph Convolutional Neural Network. In *Neural Information Processing*, pages 395–406. Springer Nature Singapore, 2023. doi: 10.1007/978-981-99-1642-9_34. URL https://bcmi.sjtu.edu.cn/~blu/papers/2022/2022-14.pdf.

[86] Arman Savran, Koray Ciftci, Guillaume Chanel, Javier Mota, Luong Hong Viet, Blent Sankur, Lale Akarun, Alice Caplier, and Michele Rombaut. Emotion Detection in the Loop from Brain Signals and Facial Images. In *Proceedings of the eNTERFACE 2006 Workshop*, 2006. URL https://archive-ouverte.unige.ch/unige:47926.

[87] Matthew D Luciw, Ewa Jarocka, and Benoni B Edin. Multi-channel EEG recordings during 3,936 grasp and lift trials with varying weight and friction. *Scientific Data*, 1(1):140047, 2014. doi: 10.1038/sdata.2014.47. URL https://doi.org/10.1038/sdata.2014.47.

[88] Perrin Margaux, Maby Emmanuel, Daligault Sébastien, Bertrand Olivier, and Mattout Jérémie. Objective and Subjective Evaluation of Online Error Correction during P300-Based Spelling. *Advances in Human-Computer Interaction*, 2012:1–13, 2012. doi: 10.1155/2012/578295. URL https://doi.org/10.1155/2012/578295.

[89] Shuai Luo, Yu-Ting Lan, Dan Peng, Ziyi Li, Wei-Long Zheng, and Bao-Liang Lu. Multimodal Emotion Recognition in Response to Oil Paintings. In *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 2022. doi: 10.1109/EMBC48229.2022.9871630. URL https://bcmi.sjtu.edu.cn/~blu/papers/2022/2022-10.pdf.

[90] Logan T. Trujillo. Mental Effort and Information-Processing Costs Are Inversely Related to Global Brain Free Energy During Visual Categorization. *Frontiers in Neuroscience*, 13:1292, 2019. doi: 10.3389/fnins.2019.01292. URL https://doi.org/10.3389/fnins.2019.01292.

[91] Logan T. Trujillo, Ph.D., Candice T. Stanfield, and Ruben D. Vela. The Effect of Electroencephalogram (EEG) Reference Choice on Information-Theoretic Measures of the Complexity and Integration of EEG Signals. *Frontiers in Neuroscience*, 11:425, 2017. doi: 10.3389/fnins.2017.00425. URL https://doi.org/10.3389/fnins.2017.00425.

[92] M. Torkamani-Azar, S. D. Kanik, S. Aydin, and M. Cetin. Prediction of Reaction Time and Vigilance Variability From Spatio-Spectral Features of Resting-State EEG in a Long Sustained Attention Task. *IEEE Journal of Biomedical and Health Informatics*, 24(9):2550–2558, 2020. doi: 10.1109/JBHI.2020.2980056. URL https:

//doi.org/10.1109/JBHI.2020.2980056.

[93] Pia Schoknecht, Himanshu Yadav, and Shravan Vasishth. Do syntactic and semantic similarity lead to interference effects? Evidence from self-paced reading and event-related potentials using German. *Journal of Memory and Language*, 141:104599, 2025. doi: 10.1016/j.jml.2024.104599. URL https://doi.org/10.1016/j.jml.2024.104599.

[94] S. Zhao, S. Hui, M. Chen, C. Feng, M. Zhai, and W. Feng. Semantic-based audiovisual integration enhances active storage in visual working memory. *Neuroimage*, 323:121577, 2025. doi: 10.1016/j.neuroimage.2025.121577. URL https://doi.org/10.1016/j.neuroimage.2025.121577.

[95] Rhiânan E. Ellis, Elizabeth Milne, and Liat Levita. Reduced visual cortical plasticity in autism spectrum disorder. *Brain Research Bulletin*, 170:11–21, 2021. doi: 10.1016/j.brainresbull.2021.01.019. URL https://doi.org/10.1016/j.brainresbull.2021.01.019.

[96] Xuan-Hao Liu, Yan-Kai Liu, Yansen Wang, Kan Ren, Hanwen Shi, Zilong Wang, Dongsheng Li, Bao-Liang Lu, and Wei-Long Zheng. EEG2Video: Towards Decoding Dynamic Visual Perception from EEG Signals. Advances in Neural Information Processing Systems, 2024. URL https://proceedings.neurips.cc/paper_files/paper/2024/hash/84bad835faaf48f24d990072bb5b80ee-Abstract.html.

[97] Ziyi Li, Le-Yan Tao, Rui-Xiao Ma, Wei-Long Zheng, and Bao-Liang Lu. Investigating the Effects of Sleep Conditions on Emotion Responses with EEG Signals and Eye Movements. IEEE Transactions on Affective Computing, 2025. URL https://doi.org/10.1109/TAFFC.2025.3572504.

[98] Wei-Long Zheng and Bao-Liang Lu. A multimodal approach to estimating vigilance using EEG and forehead EOG. Journal of Neural Engineering, 2017. URL https://doi.org/10.1088/1741-2552/aa5a98.

[99] Wei Liu, Wei-Long Zheng, Ziyi Li, Si-Yuan Wu, Lu Gan, and Bao-Liang Lu. Identifying similarities and differences in emotion recognition with EEG and eye movements among Chinese, German, and French people. Journal of Neural Engineering, 2022. URL https://doi.org/10.1088/1741-2552/ac5c8d.

[100] Jun Ma, Banghua Yang, Wenzheng Qiu, Yunzhe Li, Shouwei Gao, and Xinxing Xia. A large EEG dataset for studying cross-session variability in motor imagery brain-computer interface. Scientific Data, 2022. URL https://doi.org/10.1038/s41597-022-01647-1.

[101] Paolo Detti, Giovanni Vatti, and G. Zabalo Manrique de Lara. EEG Synchronization Analysis for Seizure Prediction: A Study on Data of Noninvasive Recordings. Processes, 2020. URL https://doi.org/10.3390/pr8070846.

[102] Lies Bollens, Bernd Accou, Hugo Van hamme, and Tom Francart. SparrKULee: A Speech-Evoked Auditory Response Repository from KU Leuven, Containing the EEG of 85 Participants. Data, 2024. URL https://doi.org/10.3390/data9080094.

[103] Tijl Grootswagers, Ivy Zhou, Amanda K. Robinson, Martin N. Hebart, and Thomas A. Carlson. Human EEG recordings for 1,854 concepts presented in rapid serial visual presentation streams. Scientific Data, 2022. URL https://doi.org/10.1038/s41597-021-01102-7.

[104] Christoph Schneider, Michael Pereira, Luca Tonin, and José del R. Millán. Real-time EEG Feedback on Alpha Power Lateralization Leads to Behavioral Improvements in a Covert Attention Task. *Brain Topography*, 33(1):48–59, 2020. doi: 10.1007/s10548-019-00725-9. URL https://doi.org/10.1007/s10548-019-00725-9.

[105] Michael Pereira, Nathan Faivre, Iñaki Iturrate, Marco Wirthlin, Luana Serafini, Stéphanie Martin, Arnaud Desvachez, Olaf Blanke, Dimitri Van De Ville, and José del R. Millán. Disentangling the origins of confidence in speeded perceptual judgments through multimodal imaging. *Proceedings of the National Academy of Sciences*, 117(15):8382–8390, 2020. doi: 10.1073/pnas.1918335117. URL https://doi.org/10.1073/pnas.1918335117.

[106] Giulia Lioi, Claire Cury, Lorraine Perronnet, Marsel Mano, Elise Bannier, Anatole Lécuyer, and Christian Barillot. Simultaneous MRI-EEG during a motor imagery neurofeedback task: an open access brain imaging dataset for multi-modal data integration. bioRxiv preprint, 2019. URL https://doi.org/10.1101/862375.

[107] Lorraine Perronnet, Anatole Lécuyer, Marsel Mano, Mathis Fleury, Giulia Lioi, Claire Cury, Maureen Clerc, Fabien Lotte, and Christian Barillot. Learning 2-in-1: Towards Integrated EEG-fMRI-Neurofeedback. bioRxiv preprint, 2018. URL https://doi.org/10.1101/397729.

[108] Daniel G Wakeman and Richard N Henson. A multi-subject, multi-modal human neuroimaging dataset. *Scientific Data*, 2(1):150001, 2015. doi: 10.1038/sdata.2015.1. URL https://doi.org/10.1038/sdata.2015.1.

[109] Duncan Williams, Alexis Kirke, Eduardo Miranda, Ian Daly, Faustina Hwang, James Weaver, and Slawomir

Nasuto. Affective Calibration of Musical Feature Sets in an Emotionally Intelligent Music Composition System. *ACM Transactions on Applied Perception*, 14(3):1–13, 2017. doi: 10.1145/3059005. URL https://doi.org/10.1145/3059005.

[110] Ian Daly, Asad Malik, Faustina Hwang, Etienne Roesch, James Weaver, Alexis Kirke, Duncan Williams, Eduardo Miranda, and Slawomir J. Nasuto. Neural correlates of emotional responses to music: An EEG study. *Neuroscience Letters*, 573:52–57, 2014. doi: 10.1016/j.neulet.2014.05.003. URL https://doi.org/10.1016/j.neulet.2014.05.003.

[111] Ian Daly, Duncan Williams, Alexis Kirke, James Weaver, Asad Malik, Faustina Hwang, Eduardo Miranda, and Slawomir J Nasuto. Affective brain–computer music interfacing. *Journal of Neural Engineering*, 13(4): 046022, 2016. doi: 10.1088/1741-2560/13/4/046022. URL https://doi.org/10.1088/1741-2560/13/4/046022.

[112] Ian Daly, Duncan Williams, Faustina Hwang, Alexis Kirke, Eduardo R. Miranda, and Slawomir J. Nasuto. Electroencephalography reflects the activity of sub-cortical brain regions during approach-withdrawal behaviour while listening to music. *Scientific Reports*, 9(1):9415, 2019. doi: 10.1038/s41598-019-45105-2. URL https://doi.org/10.1038/s41598-019-45105-2.

[113] Jobi S. George, Jon Strunk, Rachel Mak-McCully, Melissa Houser, Howard Poizner, and Adam R. Aron. Dopaminergic therapy in Parkinson's disease decreases cortical beta band coherence in the resting state and increases cortical beta band power during executive control. *NeuroImage: Clinical*, 3:261–270, 2013. doi: 10.1016/j.nicl.2013.07.013. URL https://doi.org/10.1016/j.nicl.2013.07.013.

[114] Christian Andreas Kothe, Scott Makeig, and Julie Anne Onton. Emotion Recognition from EEG during Self-Paced Emotional Imagery. In *2013 Humaine Association Conference on Affective Computing and Intelligent Interaction*, pages 855–858, 2013. doi: 10.1109/acii.2013.160. URL https://doi.org/10.1109/acii.2013.160.

[115] Nadine Svenja Josée Jacobsen, Sarah Blum, Karsten Witt, and Stefan Debener. A walk in the park? Characterizing gait-related artifacts in mobile EEG recordings. *European Journal of Neuroscience*, 54(12):8421–8440, 2021. doi: 10.1111/ejn.14965. URL https://doi.org/10.1111/ejn.14965.

[116] James F. Cavanagh, Andrew W. Bismark, Michael J. Frank, and John J. B. Allen. Multiple Dissociations Between Comorbid Depression and Anxiety on Reward and Punishment Processing: Evidence From Computationally Informed EEG. *Computational Psychiatry*, 3(0):1, 2019. doi: 10.1162/cpsy_a_00024. URL https://doi.org/10.1162/cpsy_a_00024.

[117] James F. Cavanagh, Andrea A. Mueller, Darin R. Brown, Jacqueline R. Janowich, Jacqueline H. Story-Remer, Ashley Wegele, and Sarah Pirio Richardson. Cognitive states influence dopamine-driven aberrant learning in Parkinson's disease. *Cortex*, 90:115–124, 2017. doi: 10.1016/j.cortex.2017.02.021. URL https://doi.org/10.1016/j.cortex.2017.02.021.

[118] Arun Singh, Sarah Pirio Richardson, Nandakumar Narayanan, and James F. Cavanagh. Mid-frontal theta activity is diminished during cognitive control in Parkinson's disease. *Neuropsychologia*, 117:113–122, 2018. doi: 10.1016/j.neuropsychologia.2018.05.020. URL https://doi.org/10.1016/j.neuropsychologia.2018.05.020.

[119] Björn Holtze, Manuela Jaeger, Stefan Debener, Kamil Adiloğlu, and Bojana Mirkovic. Are They Calling My Name? Attention Capture Is Reflected in the Neural Tracking of Attended and Ignored Speech. *Frontiers in Neuroscience*, 15:643705, 2021. doi: 10.3389/fnins.2021.643705. URL https://doi.org/10.3389/fnins.2021.643705.

[120] James F. Cavanagh and Joel Castellanos. Identification of canonical neural events during continuous gameplay of an 8-bit style video game. *NeuroImage*, 133:1–13, 2016. doi: 10.1016/j.neuroimage.2016.02.075. URL https://doi.org/10.1016/j.neuroimage.2016.02.075.

[121] James F. Cavanagh, Sean E. Masters, Kevin Bath, and Michael J. Frank. Conflict acts as an implicit cost in reinforcement learning. *Nature Communications*, 5(1):5394, 2014. doi: 10.1038/ncomms6394. URL https://doi.org/10.1038/ncomms6394.

[122] James M. Broadway, Michael J. Frank, and James F. Cavanagh. Dopamine D2 agonist affects visuospatial working memory distractor interference depending on individual differences in baseline working memory span. *Cognitive, Affective, & Behavioral Neuroscience*, 18(3):509–520, 2018. doi: 10.3758/s13415-018-0584-6. URL https://doi.org/10.3758/s13415-018-0584-6.

[123] James F. Cavanagh, J. Kevin Wilson, Rebecca E. Rieger, Darbi Gill, James M. Broadway, Jacqueline Hope Story Remer, Violet Fratzke, Andrew R. Mayer, and Davin K. Quinn. ERPs predict symptomatic distress

and recovery in sub-acute mild traumatic brain injury. *Neuropsychologia*, 132:107125, 2019. doi: 10.1016/j.neuropsychologia.2019.107125. URL https://doi.org/10.1016/j.neuropsychologia.2019.107125.

[124] Andrew Goldman, Tyreek Jackson, and Paul Sajda. Improvisation experience predicts how musicians categorize musical structures. *Psychology of Music*, 48(1):18–34, 2020. doi: 10.1177/0305735618779444. URL https://doi.org/10.1177/0305735618779444.

[125] Magnus Liebherr, Andrew W. Corcoran, Phillip M. Alday, Scott Coussens, Valeria Bellan, Caitlin A. Howlett, Maarten A. Immink, Mark Kohler, Matthias Schlesewsky, and Ina Bornkessel-Schlesewsky. EEG and behavioral correlates of attentional processing while walking and navigating naturalistic environments. *Scientific Reports*, 11(1):22325, 2021. doi: 10.1038/s41598-021-01772-8. URL https://doi.org/10.1038/s41598-021-01772-8.

[126] Nicolás Nieto, Victoria Peterson, Hugo Leonardo Rufiner, Juan Kamienkoski, and Ruben Spies. "Thinking out loud": an open-access EEG-based BCI dataset for inner speech recognition. bioRxiv preprint, 2021. URL https://doi.org/10.1101/2021.04.19.440473.

[127] Maria J. Ribeiro and Miguel Castelo-Branco. Age-related differences in event-related potentials and pupillary responses in cued reaction time tasks. *Neurobiology of Aging*, 73:177–189, 2019. doi: 10.1016/j.neurobiolaging.2018.09.028. URL https://doi.org/10.1016/j.neurobiolaging.2018.09.028.

[128] Samantha Gregory, Hongfang Wang, and Klaus Kessler. EEG alpha and theta signatures of socially and non-socially cued working memory in virtual reality. Center for Open Science preprint, 2021. URL https://doi.org/10.31234/osf.io/mkags.

[129] Meghan H. Puglia, Kathleen M. Krol, Manuela Missana, Cabell L. Williams, Travis S. Lillard, James P. Morris, Jessica J. Connelly, and Tobias Grossmann. Epigenetic tuning of brain signal entropy in emergent human social behavior. *BMC Medicine*, 18(1):244, 2020. doi: 10.1186/s12916-020-01683-x. URL https://doi.org/10.1186/s12916-020-01683-x.

[130] Steven M. Peterson and Daniel P. Ferris. Group-level cortical and muscular connectivity during perturbations to walking and standing balance. *NeuroImage*, 198:93–103, 2019. doi: 10.1016/j.neuroimage.2019.05.038. URL https://doi.org/10.1016/j.neuroimage.2019.05.038.

[131] Kun Chen, Ruien Wang, Jiamin Huang, Fei Gao, Zhen Yuan, Yanyan Qi, and Haiyan Wu. A resource for assessing dynamic binary choices in the adult brain using EEG and mouse-tracking. *Scientific Data*, 9(1):416, 2022. doi: 10.1038/s41597-022-01538-5. URL https://doi.org/10.1038/s41597-022-01538-5.

[132] Christoffer Hatlestad-Hall, Trine Waage Rygvold, and Stein Andersson. BIDS-structured resting-state electroencephalography (EEG) data extracted from an experimental paradigm. *Data in Brief*, 45:108647, 2022. doi: 10.1016/j.dib.2022.108647. URL https://doi.org/10.1016/j.dib.2022.108647.

[133] Lukas Gehrke, Sezen Akman, Pedro Lopes, Albert Chen, Avinash Kumar Singh, Hsiang-Ting Chen, Chin-Teng Lin, and Klaus Gramann. Detecting Visuo-Haptic Mismatches in Virtual Reality using the Prediction Error Negativity of Event-Related Brain Potentials. In *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems*, pages 1–11, 2019. doi: 10.1145/3290605.3300657. URL https://doi.org/10.1145/3290605.3300657.

[134] Sophia M. Shatek, Amanda K Robinson, Tijl Grootswagers, and Thomas A. Carlson. Capacity for movement is an organisational principle in object representations. Center for Open Science preprint, 2021. URL https://doi.org/10.31234/osf.io/3x2qh.

[135] Claire Braboszcz, B. Rael Cahn, Jonathan Levy, Manuel Fernandez, and Arnaud Delorme. Increased Gamma Brainwave Amplitude Compared to Control in Three Different Meditation Traditions. *PLOS ONE*, 12(1):e0170647, 2017. doi: 10.1371/journal.pone.0170647. URL https://doi.org/10.1371/journal.pone.0170647.

[136] Leonhard Waschke, Thomas Donoghue, Lorenz Fiedler, Sydney Smith, Douglas D Garrett, Bradley Voytek, and Jonas Obleser. Modality-specific tracking of attention and sensory statistics in the human electrophysiological spectral exponent. *eLife*, 10:e70068, 2021. doi: 10.7554/elife.70068. URL https://doi.org/10.7554/elife.70068.

[137] Tijl Grootswagers, Amanda K. Robinson, and Thomas A. Carlson. The representational dynamics of visual objects in rapid serial visual processing streams. *NeuroImage*, 188:668–679, 2019. doi: 10.1016/j.neuroimage.2018.12.046. URL https://doi.org/10.1016/j.neuroimage.2018.12.046.

[138] Denise Moerel, Tijl Grootswagers, Amanda K. Robinson, Sophia M. Shatek, Alexandra Woolgar, Thomas A. Carlson, and Anina N. Rich. Undivided attention: The temporal effects of attention dissociated from decision, memory, and expectation. bioRxiv preprint, 2021. URL https://doi.org/10.1101/2021.05.24.445376.

[139] Javier O. Garcia, Justin Brooks, Scott Kerick, Tony Johnson, Tim R. Mullen, and Jean M. Vettel. Estimating direction in brain-behavior interactions: Proactive and reactive brain states in driving. *NeuroImage*, 150:239–249, 2017. doi: 10.1016/j.neuroimage.2017.02.057. URL https://doi.org/10.1016/j.neuroimage.2017.02.057.

[140] Jon Touryan, Brent J. Lance, Scott E. Kerick, Anthony J. Ries, and Kaleb McDowell. Common EEG features for behavioral estimation in disparate, real-world tasks. *Biological Psychology*, 114:93–107, 2016. doi: 10.1016/j.biopsycho.2015.12.009. URL https://doi.org/10.1016/j.biopsycho.2015.12.009.

[141] Jon Touryan, Gregory Apker, Brent J. Lance, Scott E. Kerick, Anthony J. Ries, and Kaleb McDowell. Estimating endogenous changes in task performance from EEG. *Frontiers in Neuroscience*, 8, 2014. doi: 10.3389/fnins.2014.00155. URL https://doi.org/10.3389/fnins.2014.00155.

[142] Cameron D. Hassall, Laurence T. Hunt, and Clay B. Holroyd. Task-Level Value Affects Trial-Level Reward Processing. bioRxiv preprint, 2021. URL https://doi.org/10.1101/2021.09.16.460600.

[143] Graciela C. Alatorre-Cruz, Heather Downs, Darcy Hagood, Seth T. Sorensen, D. Keith Williams, and Linda Larson-Prior. Effect of obesity on inhibitory control in preadolescents during stop-signal task. An event-related potentials study. *International Journal of Psychophysiology*, 165:56–67, 2021. doi: 10.1016/j.ijpsycho.2021.04.003. URL https://doi.org/10.1016/j.ijpsycho.2021.04.003.

[144] Foteini Simistira Liwicki, Vibha Gupta, Rajkumar Saini, Kanjar De, Nosheen Abid, Sumit Rakesh, Scott Wellington, Holly Wilson, Marcus Liwicki, and Johan Eriksson. Bimodal electroencephalography-functional magnetic resonance imaging dataset for inner-speech recognition. bioRxiv preprint, 2022. URL https://doi.org/10.1101/2022.05.24.492109.

[145] Cameron D. Hassall, Jack Harley, Nils Kolling, and Laurence T. Hunt. Temporal scaling of human scalp-recorded potentials. bioRxiv preprint, 2020. URL https://doi.org/10.1101/2020.12.11.421180.

[146] Denise Moerel, Tijl Grootswagers, Amanda K Robinson, Patrick Engeler, Alex O. Holcombe, and Thomas A. Carlson. Rotation-tolerant representations elucidate the time-course of high-level object processing. Center for Open Science preprint, 2022. URL https://doi.org/10.31234/osf.io/wp73u.

[147] Alexis D.J. Makin, Moon M Wilton, Anna Pecchinenda, and Marco Bertamini. Symmetry perception and affective responses: A combined EEG/EMG study. *Neuropsychologia*, 50(14):3250–3261, 2012. doi: 10.1016/j.neuropsychologia.2012.10.003. URL https://doi.org/10.1016/j.neuropsychologia.2012.10.003.

[148] Kaare B. Mikkelsen, David Bové Villadsen, Marit Otto, and Preben Kidmose. Automatic sleep staging using ear-EEG. *BioMedical Engineering OnLine*, 16(1):111, 2017. doi: 10.1186/s12938-017-0400-5. URL https://doi.org/10.1186/s12938-017-0400-5.

[149] Tracy Brandmeyer and Arnaud Delorme. Closed-Loop Frontal Midlineθ Neurofeedback: A Novel Approach for Training Focused-Attention Meditation. *Frontiers in Human Neuroscience*, 14:246, 2020. doi: 10.3389/fnhum.2020.00246. URL https://doi.org/10.3389/fnhum.2020.00246.

[150] Tong Shan, Madeline S. Cappelloni, and Ross K. Maddox. Subcortical responses to music and speech are alike while cortical responses diverge. *Scientific Reports*, 14(1):789, 2024. doi: 10.1038/s41598-023-50438-0. URL https://doi.org/10.1038/s41598-023-50438-0.

[151] Tijl Grootswagers, Amanda K. Robinson, Sophia M. Shatek, and Thomas A. Carlson. Mapping the dynamics of visual feature coding: Insights into perception and integration. *PLOS Computational Biology*, 20(1):e1011760, 2024. doi: 10.1371/journal.pcbi.1011760. URL https://doi.org/10.1371/journal.pcbi.1011760.

[152] Vasileios Dimakopoulos, Giorgio Selmin, Luca Regli, and Johannes Sarnthein. Optimization of signal-to-noise ratio in short-duration SEP recordings by variation of stimulation rate. *Clinical Neurophysiology*, 150:89–97, 2023. doi: 10.1016/j.clinph.2023.03.008. URL https://doi.org/10.1016/j.clinph.2023.03.008.

[153] Seitaro Iwama, Masumi Morishige, Midori Kodama, Yoshikazu Takahashi, Ryotaro Hirose, and Junichi Ushiba. High-density scalp electroencephalogram dataset during sensorimotor rhythm-based brain-computer interfacing. *Scientific Data*, 10(1):385, 2023. doi: 10.1038/s41597-023-02260-6. URL https://doi.org/10.1038/s41597-023-02260-6.

[154] Noelle A. Jacobsen and Daniel P. Ferris. Electrocortical activity correlated with locomotor adaptation during split-belt treadmill walking. *The Journal of Physiology*, 601(17):3921–3944, 2023. doi: 10.1113/jp284505. URL https://doi.org/10.1113/jp284505.

[155] Andreas Miltiadous, Katerina D. Tzimourta, Theodora Afrantou, Panagiotis Ioannidis, Nikolaos Grigoriadis, Dimitrios G. Tsalikakis, Pantelis Angelidis, Markos G. Tsipouras, Euripidis Glavas, Nikolaos Giannakeas, and Alexandros T. Tzallas. A Dataset of Scalp EEG Recordings of Alzheimer's Disease, Frontotemporal Dementia and Healthy Subjects from Routine EEG. *Data*, 8(6):95, 2023. doi: 10.3390/data8060095. URL https:

//doi.org/10.3390/data8060095.

[156] Amanda Studnicki, Ryan J. Downey, and Daniel P. Ferris. Characterizing and Removing Artifacts Using Dual-Layer EEG during Table Tennis. *Sensors*, 22(15):5867, 2022. doi: 10.3390/s22155867. URL https://doi.org/10.3390/s22155867.

[157] Edward F. Ester and Asal Nouri. Internal selective attention is delayed by competition between endogenous and exogenous factors. bioRxiv preprint, 2022. URL https://doi.org/10.1101/2022.07.05.498906.

[158] Edward F. Ester and Paige Pytel. Changes in behavioral priority influence the accessibility but not the quality of working memory content. bioRxiv preprint, 2022. URL https://doi.org/10.1101/2022.08.12.503778.

[159] V. Volpert, B. Xu, A. Tchechmedjiev, S. Harispe, A. Aksenov, Q. Mesnildrey, and A. Beuter. Characterization of spatiotemporal dynamics in EEG data during picture naming with optical flow patterns. *Mathematical Biosciences and Engineering*, 20(6):11429–11463, 2023. doi: 10.3934/mbe.2023507. URL https://doi.org/10.3934/mbe.2023507.

[160] Zoltan Kekecs, Endre Csikos, Nguyen Dang Quy Minh, Yeganeh Farahzadi, Peter Simor, Balazs Nyiri, Pietro Rizzo, Jay A. Olson, and Gary Elkins. Can Any Procedure Be Hypnosis? Exploring the Effect of Framing on Hypnotic Depth and Electrophysiological Correlates of Hypnosis in a Balanced Placebo Design. *Psychophysiology*, 62(11):e70183, 2025. doi: 10.1111/psyp.70183. URL https://doi.org/10.1111/psyp.70183.

[161] Arun Singh, Rachel C Cole, Arturo I Espinoza, Jan R. Wessel, James F. Cavanagh, and Nandakumar S Narayanan. Evoked midfrontal activity predicts cognitive dysfunction in Parkinson's disease. medRxiv preprint, 2022. URL https://doi.org/10.1101/2022.07.26.22278079.

[162] G.A. Otero, T. Harmony, F.B. Pliego-Rivero, J. Ricardo-Garcell, J. Bosch-Bayard, R. Porcayo-Mercado, A. Fernández-Bouzas, L. Díaz-Comas, L. Galán, P. Vieyra-Reyes, and T. Fernández. QEEG norms for the first year of life. *Early Human Development*, 87(10):691–703, 2011. doi: 10.1016/j.earlhumdev.2011.05.010. URL https://doi.org/10.1016/j.earlhumdev.2011.05.010.

[163] Benjamin G. Lowe, Jonathan E. Robinson, Naohide Yamamoto, Hinze Hogendoorn, and Patrick Johnston. Same but different: The latency of a shared expectation signal interacts with stimulus attributes. *Cortex*, 168:143–156, 2023. doi: 10.1016/j.cortex.2023.08.004. URL https://doi.org/10.1016/j.cortex.2023.08.004.

[164] Patrycja Dzianok, Ingrida Antonova, Jakub Wojciechowski, Joanna Dreszer, and Ewa Kublik. The Nencki-Symfonia electroencephalography/event-related potential dataset: Multiple cognitive tasks and resting-state data collected in a sample of healthy adults. *GigaScience*, 11:giac015, 2022. doi: 10.1093/gigascience/giac015. URL https://doi.org/10.1093/gigascience/giac015.

[165] Szymon Mąka, Marta Chrustowicz, and Łukasz Okruszek. Can we dissociate hypervigilance to social threats from altered perceptual decision-making processes in lonely individuals? An exploration with Drift Diffusion Modeling and event-related potentials. *Psychophysiology*, 60(12):e14406, 2023. doi: 10.1111/psyp.14406. URL https://doi.org/10.1111/psyp.14406.

[166] Armen Bagdasarov, Denis Brunet, Christoph M. Michel, and Michael S. Gaffrey. Microstate Analysis of Continuous Infant EEG: Tutorial and Reliability. *Brain Topography*, 37(4):496–513, 2024. doi: 10.1007/s10548-024-01043-5. URL https://doi.org/10.1007/s10548-024-01043-5.

[167] Matthew R. Dougherty, Woohyeuk Chang, Joseph H. Rudoler, Brandon S. Katerman, David J. Halpern, James P. Bruska, Nicholas B. Diamond, and Michael J. Kahana. Neural correlates of memory in a naturalistic spatiotemporal context. bioRxiv preprint, 2022. URL https://doi.org/10.1101/2022.11.30.518606.

[168] Mohammad Momenian, Zhengwu Ma, Shuyi Wu, Chengcheng Wang, Jonathan Brennan, John Hale, Lars Meyer, and Jixing Li. Le Petit Prince Hong Kong (LPPHK): Naturalistic fMRI and EEG data from older Cantonese speakers. bioRxiv preprint, 2024. URL https://doi.org/10.1101/2024.04.24.590842.

[169] Ryan J. Downey and Daniel P. Ferris. iCanClean Removes Motion, Muscle, Eye, and Line-Noise Artifacts from Phantom EEG. *Sensors*, 23(19):8214, 2023. doi: 10.3390/s23198214. URL https://doi.org/10.3390/s23198214.

[170] Patrycja Dzianok and Ewa Kublik. PEARL-Neuro Database: EEG, fMRI, health and lifestyle data of middle-aged people at risk of dementia. *Scientific Data*, 11(1):276, 2024. doi: 10.1038/s41597-024-03106-5. URL https://doi.org/10.1038/s41597-024-03106-5.

[171] Tijl Grootswagers, Amanda K Robinson, Sophia M Shatek, and Thomas A Carlson. The neural dynamics underlying prioritisation of task-relevant information. *Neurons, Behavior, Data analysis, and Theory*, 5(1), 2021. doi: 10.51628/001c.21174. URL https://doi.org/10.51628/001c.21174.

[172] Chuqin Xiang, Xinrui Fan, Duo Bai, Ke Lv, and Xu Lei. A resting-state EEG dataset for sleep deprivation. *Scientific Data*, 11(1):427, 2024. doi: 10.1038/s41597-024-03268-2. URL https://doi.org/10.1038/

s41597-024-03268-2.

[173] Gabriela Valdebenito-Oyarzo, María Paz Martínez-Molina, Patricia Soto-Icaza, Francisco Zamorano, Alejandra Figueroa-Vargas, Josefina Larraín-Valenzuela, Ximena Stecher, César Salinas, Julien Bastin, Antoni Valero-Cabré, Rafael Polania, and Pablo Billeke. The parietal cortex has a causal role in ambiguity computations in humans. *PLOS Biology*, 22(1):e3002452, 2024. doi: 10.1371/journal.pbio.3002452. URL https://doi.org/10.1371/journal.pbio.3002452.

[174] Xinyu Mou, Cuilin He, Liwei Tan, Junjie Yu, Huadong Liang, Jianyu Zhang, Yan Tian, Yu-Fang Yang, Ting Xu, Qing Wang, Miao Cao, Zijiao Chen, Chuan-Peng Hu, Xindi Wang, Quanying Liu, and Haiyan Wu. ChineseEEG: A Chinese Linguistic Corpora EEG Dataset for Semantic Alignment and Neural Decoding. *Scientific Data*, 11(1):550, 2024. doi: 10.1038/s41597-024-03398-7. URL https://doi.org/10.1038/s41597-024-03398-7.

[175] Denise Moerel, James Psihoyos, and Thomas A. Carlson. The Time-Course of Food Representation in the Human Brain. bioRxiv preprint, 2023. URL https://doi.org/10.1101/2023.06.06.543985.

[176] Amanda K Robinson, Tijl Grootswagers, Sophia M. Shatek, Marlene Behrmann, and Thomas A. Carlson. Dynamics of visual object coding within and across the hemispheres: Objects in the periphery. Center for Open Science preprint, 2024. URL https://doi.org/10.31234/osf.io/a3bpd.

[177] Tijl Grootswagers, Genevieve Quek, Zhen Zeng, and Manuel Varlet. Human infant EEG recordings for 200 object images presented in rapid visual streams. Center for Open Science preprint, 2024. URL https://doi.org/10.31234/osf.io/dzrkq.

[178] James F. Cavanagh, Rebecca E. Rieger, J. Kevin Wilson, Darbi Gill, Lynne Fullerton, Emma Brandt, and Andrew R. Mayer. Joint analysis of frontal theta synchrony and white matter following mild traumatic brain injury. *Brain Imaging and Behavior*, 14(6):2210–2223, 2020. doi: 10.1007/s11682-019-00171-y. URL https://doi.org/10.1007/s11682-019-00171-y.

[179] Elizabeth M. Siefert, Sindhuja Uppuluri, Jianing Mu, Marlie C. Tandoc, James W. Antony, and Anna C. Schapiro. Memory Reactivation during Sleep Does Not Act Holistically on Object Memory. *The Journal of Neuroscience*, 44(24):e0022242024, 2024. doi: 10.1523/jneurosci.0022-24.2024. URL https://doi.org/10.1523/jneurosci.0022-24.2024.

[180] Quentin Chenot, Caroline Hamery, Moritz Truninger, Nicolas Langer, Xavier De boissezon, and Sébastien Scannella. Investigating the relationship between resting-state EEG microstates and executive functions: A null finding. *Cortex*, 178:1–17, 2024. doi: 10.1016/j.cortex.2024.05.019. URL https://doi.org/10.1016/j.cortex.2024.05.019.

[181] Melissa J. Polonenko and Ross K. Maddox. Fundamental frequency predominantly drives talker differences in auditory brainstem responses to continuous speech. *JASA Express Letters*, 4(11):114401, 2024. doi: 10.1121/10.0034329. URL https://doi.org/10.1121/10.0034329.

[182] Nayid Triana-Guzman, Alvaro D. Orjuela-Cañon, Andres L. Jutinico, Omar Mendoza-Montoya, and Javier M. Antelis. Decoding EEG rhythms offline and online during motor imagery for standing and sitting based on a brain-computer interface. *Frontiers in Neuroinformatics*, 16:961089, 2022. doi: 10.3389/fninf.2022.961089. URL https://doi.org/10.3389/fninf.2022.961089.

[183] Patrick D Gajewski, Stephan Getzmann, Peter Bröde, Michael Burke, Cristina Cadenas, Silvia Capellino, Maren Claus, Erhan Genç, Klaus Golka, Jan G Hengstler, Thomas Kleinsorge, Rosemarie Marchan, Michael A Nitsche, Jörg Reinders, Christoph van Thriel, Carsten Watzl, and Edmund Wascher. Impact of Biological and Lifestyle Factors on Cognitive Aging and Work Ability in the Dortmund Vital Study: Protocol of an Interdisciplinary, Cross-sectional, and Longitudinal Study. *JMIR Research Protocols*, 11(3):e32352, 2022. doi: 10.2196/32352. URL https://doi.org/10.2196/32352.

[184] Christopher Hilton, Lilian Befort, Ronja Brinkmann, Matthias Ballestrem, Joerg Fingerhut, and Klaus Gramann. Stairs as multifunctional spaces: Cortical responses to environmental affordances incorporate the intention to act. *Journal of Environmental Psychology*, 102:102528, 2025. doi: 10.1016/j.jenvp.2025.102528. URL https://doi.org/10.1016/j.jenvp.2025.102528.

[185] Yuri G. Pavlov, Nick S. Menger, Andreas Keil, and Boris Kotchoubey. Contingency awareness shapes neural responses in fear conditioning. bioRxiv preprint, 2024. URL https://doi.org/10.1101/2024.08.13.607803.

[186] Seyed Yahya Shirazi, Alexandre Franco, Maurício Scopel Hoffmann, Nathalia B. Esper, Dung Truong, Arnaud Delorme, Michael P. Milham, and Scott Makeig. HBN-EEG: The FAIR implementation of the Healthy Brain Network (HBN) electroencephalography dataset. bioRxiv preprint, 2024. URL https://doi.org/10.1101/

2024.10.03.615261.

[187] Martín Esparza-Iaizzo, María Sierra-Torralba, Jens G Klinzing, Javier Minguez, Luis Montesano, and Eduardo López-Larraz. Automatic sleep scoring for real-time monitoring and stimulation in individuals with and without sleep apnea. *Computers in Biology and Medicine*, 205:111560, 2026. doi: 10.1016/j.compbiomed.2026.111560. URL https://doi.org/10.1016/j.compbiomed.2026.111560.

[188] María Paz Martínez-Molina, Gabriela Valdebenito-Oyarzo, Patricia Soto-Icaza, Francisco Zamorano, Alejandra Figueroa-Vargas, Patricio Carvajal-Paredes, Ximena Stecher, César Salinas, Antoni Valero-Cabré, Rafael Polania, and Pablo Billeke. Lateral prefrontal theta oscillations causally drive a computational mechanism underlying conflict expectation and adaptation. *Nature Communications*, 15(1):9858, 2024. doi: 10.1038/s41467-024-54244-8. URL https://doi.org/10.1038/s41467-024-54244-8.

[189] Jack E. Taylor, Rasmus Sinn, Cosimo Iaia, and Christian J. Fiebach. Beyond Letters: Optimal Transport as a Model for Sub-Letter Orthographic Processing. bioRxiv preprint, 2024. URL https://doi.org/10.1101/2024.11.11.622929.

[190] Paula Pena, Ana F. Palenciano, Carlos González-García, and María Ruz. Novel Verbal Instructions Recruit Abstract Neural Patterns of Time-Variable Information Dimensionality. *The Journal of Neuroscience*, 45(17): e1964242025, 2025. doi: 10.1523/jneurosci.1964-24.2025. URL https://doi.org/10.1523/jneurosci.1964-24.2025.

[191] Leonardo Novelli, Devon Stoliker, Tamrin Barta, Matthew D. Greaves, Sidhant Chopra, James Jackson, Jessica Kwee, James C. Pang, Martin L. Williams, and Adeel Razi. PsiConnect: Multimodal Neuroimaging of Context-Dependent Brain and Behaviour Dynamics under Psilocybin. *Scientific Data*, 13(1):1407, 2026. doi: 10.1038/s41597-026-07312-1. URL https://doi.org/10.1038/s41597-026-07312-1.

[192] María Melcón, Enrique Stern, Dominique Kessel, Lydia Arana, Claudia Poch, Pablo Campo, and Almudena Capilla. Perception of near-threshold visual stimuli is influenced by prestimulus alpha-band amplitude but not by alpha phase. *Psychophysiology*, 61(5):e14525, 2024. doi: 10.1111/psyp.14525. URL https://doi.org/10.1111/psyp.14525.

[193] Andy Jeesu Kim, Santiago Morales, Joshua Senior, and Mara Mather. Electroencephalography, pupillometry, and behavioral evidence for locus coeruleus-noradrenaline system related tonic hyperactivity in older adults. bioRxiv preprint, 2025. URL https://doi.org/10.1101/2025.10.02.680040.

[194] Klaus Gramann, Thomas Töllner, and Hermann J. Müller. Dimension-based attention modulates early visual processing. *Psychophysiology*, 2010. doi: 10.1111/j.1469-8986.2010.00998.x. URL https://doi.org/10.1111/j.1469-8986.2010.00998.x.

[195] Julie A. Onton, Katherine C. Simon, Allison B. Morehouse, Alessandra E. Shuster, Jing Zhang, Andres A. Peña, and Sara C. Mednick. Validation of spectral sleep scoring with polysomnography using forehead EEG device. *Frontiers in Sleep*, 3:1349537, 2024. doi: 10.3389/frsle.2024.1349537. URL https://doi.org/10.3389/frsle.2024.1349537.

[196] Denise Moerel, Tijl Grootswagers, Jessica L L Chin, Francesca Ciardo, Patti Nijhuis, Genevieve L Quek, Sophie Smit, and Manuel Varlet. Neural decoding of competitive decision-making in Rock–Paper–Scissors. *Social Cognitive And Affective Neuroscience*, 20(1):nsaf101, 2025. doi: 10.1093/scan/nsaf101. URL https://doi.org/10.1093/scan/nsaf101.

[197] Denise Moerel, Tijl Grootswagers, Genevieve L. Quek, Sophie Smit, and Manuel Varlet. Information alignment between interacting brains. bioRxiv preprint, 2025. URL https://doi.org/10.1101/2025.01.07.631802.

[198] Şahcan Özdemir, Eren Günseli, and Daniel Schneider. Motor Control Processes Moderate Visual Working Memory Gating. *The Journal of Neuroscience*, 45(47):e0673252025, 2025. doi: 10.1523/jneurosci.0673-25.2025. URL https://doi.org/10.1523/jneurosci.0673-25.2025.

[199] Pablo Oyarzo, Johannes J.D. Singer, Kohitij Kar, Diego Vidaurre, and Radoslaw M. Cichy. Adaptive recruitment of cortex-wide recurrence for visual object recognition. bioRxiv preprint, 2025. URL https://doi.org/10.1101/2025.10.17.682937.

[200] Petunia Reinke, Lisa Deneke, and Sebastian Ocklenburg. Hemispheric asymmetries in the EEG: Is there an association between N1 lateralization and alpha asymmetry? *Laterality*, 31(1):40–89, 2026. doi: 10.1080/1357650x.2025.2591660. URL https://doi.org/10.1080/1357650x.2025.2591660.

[201] Verónica Henao Isaza, David Aguillon, Carlos Andrés Tobón-Quintero, Francisco Lopera, and John Fredy Ochoa-Gómez. Comprehensive methodology for sample enrichment in EEG biomarker studies for Alzheimer's risk classification. *PLOS One*, 21(3):e0343722, 2026. doi: 10.1371/journal.pone.0343722. URL https:

//doi.org/10.1371/journal.pone.0343722.

[202] Verónica Henao Isaza, Valeria Cadavid Castro, Luisa María Zapata-Saldarriaga, Yorguin-Jose Mantilla-Ramos, Jazmin Ximena Suarez-Revelo, Carlos Andrés Tobón-Quintero, and John Fredy Ochoa-Gómez. Longitudinal test-retest reliability of quantitative EEG in healthy individuals using an automated preprocessing approach. *Biomedical Signal Processing and Control*, 116:109484, 2026. doi: 10.1016/j.bspc.2026.109484. URL https://doi.org/10.1016/j.bspc.2026.109484.

[203] Maria Martzoukou, Nefeli K. Dimitriou, Binbin Xu, Malo Renaud-D'Ambra, Anastasia Nousia, Anne Beuter, and Grigorios Nasios. Transcranial Alternating Current Stimulation (tACS) for patients with Post-Stroke Anomia: Preliminary Data on Picture Naming Performance. *PLOS One*, 21(2):e0342191, 2026. doi: 10.1371/journal.pone.0342191. URL https://doi.org/10.1371/journal.pone.0342191.

[204] Daniel Vered, Zoya Katzir, Idan Daniel Grosbard, Avner Thaler, and Inbal Maidan. Electroencephalography-Based Clustering Reveals Robust Neurophysiological Subtypes in Parkinson's Disease. *Movement Disorders*, 41(8):2107–2116, 2026. doi: 10.1002/mds.70348. URL https://doi.org/10.1002/mds.70348.

[205] Tamari Shalamberidze, Kyle Nash, and Jeremy B. Caplan. Rhythmic activity in resting-state EEG predicts trait anxiety. *Imaging Neuroscience*, 3:IMAG.a.44, 2025. doi: 10.1162/imag.a.44. URL https://doi.org/10.1162/imag.a.44.

[206] Henrik Normannseth, Christoffer Hatlestad-Hall, Trine Waage Rygvold, Alena Hadzic, and Stein Andersson. Hormonal contraceptive use is associated with reduced central serotonergic activity indexed by the loudness dependence of auditory evoked potentials. *Frontiers in Human Neuroscience*, 19:1647425, 2025. doi: 10.3389/fnhum.2025.1647425. URL https://doi.org/10.3389/fnhum.2025.1647425.

[207] Haorui Sun, Alaina Birney, Niharika Singh, Ardyn Olszko, Peijin Chen, Jin Ke, Monica D. Rosenberg, and David C. Jangraw. ReMind: A Retrospective Self-Report Paradigm for Studying Mind-Wandering Onset During Reading. bioRxiv preprint, 2026. URL https://doi.org/10.64898/2026.05.14.725227.

[208] Marion Brickwedde, Rupali Limachya, Roksana Markiewicz, Emma Sutton, Christopher Postzich, Kimron Shapiro, Ole Jensen, and Ali Mazaheri. Cross-modal interaction of Alpha Activity does not reflect inhibition of early sensory processing: A frequency tagging study using EEG and MEG. eLife preprint, 2025. URL https://doi.org/10.7554/elife.106050.3.

[209] Ana Calzada-Reyes, Lidice Galán-García, Trinidad Virués-Alba, Lidia Charroó-Ruiz, Laura Perez-Mayo, Maria Luisa Bringas-Vega, Jorge Bosch-Bayard, Yanely Acosta-Ymas, Eduardo Aubert-Vázquez, Mayrim Vega-Hernández, Joel Gutiérrez-Gil, Antonio Caballero-Moreno, Annette Valdés-Virués, Mitchell Valdés-Sosa, Peng Ren, Dezhong Yao, Luo Cheng, Roberto Rodriguez-Labrada, and Pedro Valdés-Sosa. EEG Signatures of COVID-19 Survival compared to close contacts and the Cuban EEG normative database. *Clinical Neurophysiology Practice*, 11:174–186, 2026. doi: 10.1016/j.cnp.2026.02.007. URL https://doi.org/10.1016/j.cnp.2026.02.007.

[210] Helene Vitali, Claudio Campus, Sabrina Signorini, Valentina De Giorgis, Federica Morelli, Costanza Varesio, Ludovica Pasca, Alessia Sammartano, and Monica Gori. Blindness affects the developmental trajectory of the sleeping brain. *NeuroImage*, 286:120508, 2024. doi: 10.1016/j.neuroimage.2024.120508. URL https://doi.org/10.1016/j.neuroimage.2024.120508.

[211] Yuanyuan Yao, Axel Stebner, Tinne Tuytelaars, Simon Geirnaert, and Alexander Bertrand. Identifying temporal correlations between natural single-shot videos and EEG signals. *Journal of Neural Engineering*, 21(1):016018, 2024. doi: 10.1088/1741-2552/ad2333. URL https://doi.org/10.1088/1741-2552/ad2333.

[212] P. A. Grasso, I. Petrizzo, F. Coniglio, and R. Arrighi. Electrophysiological correlates of temporal numerosity adaptation. *Front Neurosci*, 18:1349540, 2024. doi: 10.3389/fnins.2024.1349540. URL https://doi.org/10.3389/fnins.2024.1349540.

[213] V. Weber, S. Ruch, N. H. Skieresz, N. Rothen, and T. P. Reber. Correlates of implicit semantic processing as revealed by representational similarity analysis applied to EEG. *iScience*, 27(11):111149, 2024. doi: 10.1016/j.isci.2024.111149. URL https://doi.org/10.1016/j.isci.2024.111149.

[214] Constantin Jehn, Adrian Kossmann, Niki Katerina Vavatzanidis, Anja Hahne, and Tobias Reichenbach. CNNs improve decoding of selective attention to speech in cochlear implant users. *Journal of Neural Engineering*, 22(3):036034, 2025. doi: 10.1088/1741-2552/addb7b. URL https://doi.org/10.1088/1741-2552/addb7b.

[215] Betina Korka and Max-Philipp Stenner. The P3 event-related potential increases when humans learn a strategy for motor adaptation. *European Journal of Neuroscience*, 60(6):5362–5380, 2024. doi: 10.1111/ejn.16476. URL https://doi.org/10.1111/ejn.16476.

[216] Rosa S. Gisladottir, Dorothee J. Chwilla, and Stephen C. Levinson. Conversation Electrified: ERP Correlates of Speech Act Recognition in Underspecified Utterances. *PLOS ONE*, 10(3):e0120068, 2015. doi: 10.1371/journal.pone.0120068. URL https://doi.org/10.1371/journal.pone.0120068.

[217] Roman Kessler and Michael A. Skeide. Electrophysiological decoding captures the temporal trajectory of face categorization in infants. bioRxiv preprint, 2024. URL https://doi.org/10.1101/2024.10.07.617144.

[218] Rose Nasrawi, Mika Mautner-Rohde, and Freek van Ede. Memory load influences our preparedness to act on visual representations in working memory without affecting their accessibility. bioRxiv preprint, 2024. URL https://doi.org/10.1101/2024.02.23.581707.

[219] E. Pierrieau, C. Dussard, A. Plantey-Veux, C. Guerrini, B. Lau, L. Pillette, N. George, and C. Jeunet-Kelway. Changes in cortical beta power predict motor control flexibility, not vigor. *Commun Biol*, 8(1):1041, 2025. doi: 10.1038/s42003-025-08465-2. URL https://doi.org/10.1038/s42003-025-08465-2.

[220] Maurice Rekrut, Abdulrahman Mohamed Selim, and Antonio Kruger. Improving Silent Speech BCI Training Procedures Through Transfer from Overt to Silent Speech. In *2022 IEEE International Conference on Systems, Man, and Cybernetics (SMC)*, pages 2650–2656, 2022. doi: 10.1109/smc53654.2022.9945447. URL https://doi.org/10.1109/smc53654.2022.9945447.

[221] Yuanyuan Yao, Wout De Swaef, Simon Geirnaert, and Alexander Bertrand. EEG-Based Decoding of Selective Visual Attention in Superimposed Videos. *IEEE Journal of Biomedical and Health Informatics*, 29(10):7248–7261, 2025. doi: 10.1109/jbhi.2025.3580261. URL https://doi.org/10.1109/jbhi.2025.3580261.

[222] Patrick Ofner, Andreas Schwarz, Joana Pereira, Daniela Wyss, Renate Wildburger, and Gernot R. Müller-Putz. Attempted Arm and Hand Movements can be Decoded from Low-Frequency EEG from Persons with Spinal Cord Injury. *Scientific Reports*, 9(1):7134, 2019. doi: 10.1038/s41598-019-43594-9. URL https://doi.org/10.1038/s41598-019-43594-9.

[223] K. Egger, T. Dau, and B. Epp. Supra-threshold perception and neural representation of tones presented in noise in conditions of masking release. *PLoS One*, 14(10):e0222804, 2019. doi: 10.1371/journal.pone.0222804. URL https://doi.org/10.1371/journal.pone.0222804.

[224] Katharina Egger, Torsten Dau, and Bastian Epp. Supra-threshold perception and neural representation of tones presented in noise in conditions of masking release. bioRxiv preprint, 2019. URL https://doi.org/10.1101/575720.

[225] M. A. Albrecht, J. A. Waltz, J. F. Cavanagh, M. J. Frank, and J. M. Gold. Reduction of Pavlovian Bias in Schizophrenia: Enhanced Effects in Clozapine-Administered Patients. *PLoS One*, 11(4):e0152781, 2016. doi: 10.1371/journal.pone.0152781. URL https://doi.org/10.1371/journal.pone.0152781.

[226] Thomas Zandonai, Darías Holgado, Luis F. Ciria, Mikel Zabala, James Hopker, Tristán Bekinschtein, and Daniel Sanabria. Novel evidence on the effect of tramadol on self-paced high-intensity cycling. *Journal of Sports Sciences*, 39(13):1452–1460, 2021. doi: 10.1080/02640414.2021.1877440. URL https://doi.org/10.1080/02640414.2021.1877440.

[227] Søren A. Fuglsang, Jonatan Märcher-Rørsted, Torsten Dau, and Jens Hjortkjær. Effects of Sensorineural Hearing Loss on Cortical Synchronization to Competing Speech during Selective Attention. *The Journal of Neuroscience*, 40(12):2562–2572, 2020. doi: 10.1523/jneurosci.1936-19.2020. URL https://doi.org/10.1523/jneurosci.1936-19.2020.

[228] A. Burroughs, N. Kazanina, and C. Houghton. Grammatical category and the neural processing of phrases. *Sci Rep*, 11(1):2446, 2021. doi: 10.1038/s41598-021-81901-5. URL https://doi.org/10.1038/s41598-021-81901-5.

[229] G. Gennari, S. Marti, M. Palu, A. Fló, and G. Dehaene-Lambertz. Orthogonal neural codes for speech in the infant brain. *Proc Natl Acad Sci U S A*, 118(31):e2020410118, 2021. doi: 10.1073/pnas.2020410118. URL https://doi.org/10.1073/pnas.2020410118.

[230] Melissa J Polonenko and Ross K Maddox. Exposing distinct subcortical components of the auditory brainstem response evoked by continuous naturalistic speech. *eLife*, 10:e62329, 2021. doi: 10.7554/elife.62329. URL https://doi.org/10.7554/elife.62329.

[231] Joel Frohlich, Damla Senturk, Vidya Saravanapandian, Peyman Golshani, Lawrence T. Reiter, Raman Sankar, Ronald L. Thibert, Charlotte DiStefano, Scott Huberty, Edwin H. Cook, and Shafali S. Jeste. A Quantitative Electrophysiological Biomarker of Duplication 15q11.2-q13.1 Syndrome. *PLOS ONE*, 11(12):e0167179, 2016. doi: 10.1371/journal.pone.0167179. URL https://doi.org/10.1371/journal.pone.0167179.

[232] Abhrajeet V. Roy, Keith W. Jamison, Sheng He, Stephen A. Engel, and Bin He. Deactivation in the posterior mid-cingulate cortex reflects perceptual transitions during binocular rivalry: Evidence from simultaneous EEG-

fMRI. *NeuroImage*, 152:1–11, 2017. doi: 10.1016/j.neuroimage.2017.02.041. URL https://doi.org/10.1016/j.neuroimage.2017.02.041.

[233] Assaf Breska and Leon Y. Deouell. Neural mechanisms of rhythm-based temporal prediction: Delta phase-locking reflects temporal predictability but not rhythmic entrainment. *PLOS Biology*, 15(2):e2001665, 2017. doi: 10.1371/journal.pbio.2001665. URL https://doi.org/10.1371/journal.pbio.2001665.

[234] M. Wischnewski, I. Alekseichuk, and D. J. L. G. Schutter. Behavioral and electrocortical effects of transcranial alternating current stimulation during advice-guided decision-making. *Neuroimage Rep*, 1(4):100052, 2021. doi: 10.1016/j.ynirp.2021.100052. URL https://doi.org/10.1016/j.ynirp.2021.100052.

[235] Justin Böhmer and Barbara Schmidt. Safety on demand: Post-hypnotic suggestions of safety reduce neural signals of reward sensitivity with long-term beneficial effects. *Psychophysiology*, 59(7):e14015, 2022. doi: 10.1111/psyp.14015. URL https://doi.org/10.1111/psyp.14015.

[236] Melissa J Polonenko and Ross K Maddox. Optimizing Parameters for Using the Parallel Auditory Brainstem Response (pABR) to Quickly Estimate Hearing Thresholds. bioRxiv preprint, 2021. URL https://doi.org/10.1101/2021.05.13.444069.

[237] S. Ladouce, L. Darmet, J. J. Torre Tresols, S. Velut, G. Ferraro, and F. Dehais. Improving user experience of SSVEP BCI through low amplitude depth and high frequency stimuli design. *Sci Rep*, 12(1):8865, 2022. doi: 10.1038/s41598-022-12733-0. URL https://doi.org/10.1038/s41598-022-12733-0.

[238] F. Doricchi, M. Pinto, M. Pellegrino, F. Marson, M. Aiello, S. Campana, F. Tomaiuolo, and S. Lasaponara. Deficits of hierarchical predictive coding in left spatial neglect. *Brain Commun*, 3(2):fcab111, 2021. doi: 10.1093/braincomms/fcab111. URL https://doi.org/10.1093/braincomms/fcab111.

[239] A. Castelnovo, J. Amacker, M. Maiolo, N. Amato, M. Pereno, S. Riccardi, A. Danani, S. Ulzega, and M. Manconi. High-density EEG power topography and connectivity during confusional arousal. *Cortex*, 155:62–74, 2022. doi: 10.1016/j.cortex.2022.05.021. URL https://doi.org/10.1016/j.cortex.2022.05.021.

[240] J. Nicolas, B. R. King, D. Levesque, L. Lazzouni, E. Coffey, S. Swinnen, J. Doyon, J. Carrier, and G. Albouy. Sigma oscillations protect or reinstate motor memory depending on their temporal coordination with slow waves. *Elife*, 11:e73930, 2022. doi: 10.7554/elife.73930. URL https://doi.org/10.7554/elife.73930.

[241] S. Yrttiaho, A. Kylliäinen, T. Parviainen, and M. J. Peltola. Neural specialization to human faces at the age of 7 months. *Sci Rep*, 12(1):12471, 2022. doi: 10.1038/s41598-022-16691-5. URL https://doi.org/10.1038/s41598-022-16691-5.

[242] L. Santamaria, I. Kashif, N. McGinley, and P. A. Lewis. Memory reactivation in slow wave sleep enhances relational learning in humans. *Commun Biol*, 7(1):288, 2024. doi: 10.1038/s42003-024-05947-7. URL https://doi.org/10.1038/s42003-024-05947-7.

[243] Lorena Santamaria, Ibad Kashif, Niall McGinley, and Penelope A. Lewis. Memory reactivation in slow wave sleep enhances relational learning. bioRxiv preprint, 2022. URL https://doi.org/10.1101/2022.03.29.486197.

[244] K. Stone, B. Nicenboim, S. Vasishth, and F. Rösler. Understanding the Effects of Constraint and Predictability in ERP. *Neurobiol Lang (Camb)*, 4(2):221–256, 2023. doi: 10.1162/nol_a_00094. URL https://doi.org/10.1162/nol_a_00094.

[245] E. V. Pedapati, L. M. Schmitt, L. E. Ethridge, M. Miyakoshi, J. A. Sweeney, R. Liu, E. Smith, R. C. Shaffer, K. C. Dominick, D. L. Gilbert, S. W. Wu, P. S. Horn, D. K. Binder, M. Lamy, M. Axford, and C. A. Erickson. Neocortical localization and thalamocortical modulation of neuronal hyperexcitability contribute to Fragile X Syndrome. *Commun Biol*, 5(1):442, 2022. doi: 10.1038/s42003-022-03395-9. URL https://doi.org/10.1038/s42003-022-03395-9.

[246] P. Manser, H. Poikonen, and E. D. de Bruin. Feasibility, usability, and acceptance of "Brain-IT"-A newly developed exergame-based training concept for the secondary prevention of mild neurocognitive disorder: a pilot randomized controlled trial. *Front Aging Neurosci*, 15:1163388, 2023. doi: 10.3389/fnagi.2023.1163388. URL https://doi.org/10.3389/fnagi.2023.1163388.

[247] C. Stinkeste, M. A. Vincent, L. Delrue, and A. Brunellière. Between alpha and gamma oscillations: Neural signatures of linguistic predictions and listener's attention to speaker's communication intention. *Biol Psychol*, 180:108583, 2023. doi: 10.1016/j.biopsycho.2023.108583. URL https://doi.org/10.1016/j.biopsycho.2023.108583.

[248] E. Villar Ortega, K. A. Buetler, E. A. Aksöz, and L. Marchal-Crespo. Enhancing touch sensibility with sensory electrical stimulation and sensory retraining. *J Neuroeng Rehabil*, 21(1):79, 2024. doi: 10.1186/s12984-024-01371-4. URL https://doi.org/10.1186/s12984-024-01371-4.

[249] R. Nasrawi, S. E. P. Boettcher, and F. van Ede. Prospection of Potential Actions during Visual Working Memory Starts Early, Is Flexible, and Predicts Behavior. *J Neurosci*, 43(49):8515–8524, 2023. doi: 10.1523/jneurosci.0709-23.2023. URL https://doi.org/10.1523/jneurosci.0709-23.2023.

[250] G. Y. T. Legendre, M. Moyne, J. Domínguez-Borràs, S. Kumar, V. Sterpenich, S. Schwartz, and L. H. Arnal. Scream's roughness grants privileged access to the brain during sleep. *Sci Rep*, 15(1):16686, 2025. doi: 10.1038/s41598-025-01560-8. URL https://doi.org/10.1038/s41598-025-01560-8.

[251] Alexander P. Rockhill, Nicko Jackson, Jobi George, Adam Aron, and Nicole C. Swann. UC San Diego Resting State EEG Data from Patients with Parkinson's Disease. OpenNeuro dataset, 2021. URL https://openneuro.org/datasets/ds002778.

[252] Bingchuan Liu, Xiaoshan Huang, Yijun Wang, Xiaogang Chen, and Xiaorong Gao. BETA: A Large Benchmark Database Toward SSVEP-BCI Application. *Frontiers in Neuroscience*, 14:627, 2020. doi: 10.3389/fnins.2020.00627. URL https://doi.org/10.3389/fnins.2020.00627.

[253] Ricardo Chavarriaga and José del R. Millan. Learning From EEG Error-Related Potentials in Noninvasive Brain-Computer Interfaces. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 18(4):381–388, 2010. doi: 10.1109/tnsre.2010.2053387. URL https://doi.org/10.1109/tnsre.2010.2053387.

[254] Anahit Babayan, Miray Erbey, Deniz Kumral, Janis D. Reinelt, Andrea M. F. Reiter, Josefin Röbbig, H. Lina Schaare, Marie Uhlig, Alfred Anwander, Pierre-Louis Bazin, Annette Horstmann, Leonie Lampe, Vadim V. Nikulin, Hadas Okon-Singer, Sven Preusser, André Pampel, Christiane S. Rohr, Julia Sacher, Angelika Thöne-Otto, Sabrina Trapp, Till Nierhaus, Denise Altmann, Katrin Arelin, Maria Blöchl, Edith Bongartz, Patric Breig, Elena Cesnaite, Sufang Chen, Roberto Cozatl, Saskia Czerwonatis, Gabriele Dambrauskaite, Maria Dreyer, Jessica Enders, Melina Engelhardt, Marie Michele Fischer, Norman Forschack, Johannes Golchert, Laura Golz, C. Alexandrina Guran, Susanna Hedrich, Nicole Hentschel, Daria I. Hoffmann, Julia M. Huntenburg, Rebecca Jost, Anna Kosatschek, Stella Kunzendorf, Hannah Lammers, Mark E. Lauckner, Keyvan Mahjoory, Ahmad S. Kanaan, Natacha Mendes, Ramona Menger, Enzo Morino, Karina Näthe, Jennifer Neubauer, Handan Noyan, Sabine Oligschläger, Patricia Panczyszyn-Trzewik, Dorothee Poehlchen, Nadine Putzke, Sabrina Roski, Marie-Catherine Schaller, Anja Schieferbein, Benito Schlaak, Robert Schmidt, Krzysztof J. Gorgolewski, Hanna Maria Schmidt, Anne Schrimpf, Sylvia Stasch, Maria Voss, Annett Wiedemann, Daniel S. Margulies, Michael Gaebler, and Arno Villringer. A mind-brain-body dataset of MRI, EEG, cognition, emotion, and peripheral physiology in young and old adults. *Scientific Data*, 6(1):180308, 2019. doi: 10.1038/sdata.2018.308. URL https://doi.org/10.1038/sdata.2018.308.

[255] B. Kemp, A.H. Zwinderman, B. Tuk, H.A.C. Kamphuisen, and J.J.L. Oberye. Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG. *IEEE Transactions on Biomedical Engineering*, 47(9):1185–1194, 2000. doi: 10.1109/10.867928. URL https://doi.org/10.1109/10.867928.

[256] S. Lopez, G. Suarez, D. Jungreis, I. Obeid, and J. Picone. Automated identification of abnormal adult EEGs. In *2015 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, pages 1–5, 2015. doi: 10.1109/spmb.2015.7405423. URL https://doi.org/10.1109/spmb.2015.7405423.

[257] A. Harati, M. Golmohammadi, S. Lopez, I. Obeid, and J. Picone. Improved EEG event classification using differential energy. In *2015 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, pages 1–4, 2015. doi: 10.1109/spmb.2015.7405421. URL https://doi.org/10.1109/spmb.2015.7405421.

[258] E. von Weltin, T. Ahsan, V. Shah, D. Jamshed, M. Golmohammadi, I. Obeid, and J. Picone. Electroencephalographic slowing: A primary source of error in automatic seizure detection. In *2017 IEEE Signal Processing in Medicine and Biology Symposium (SPMB)*, pages 1–5, 2017. doi: 10.1109/spmb.2017.8257018. URL https://doi.org/10.1109/spmb.2017.8257018.

[259] Haoming Zhang, Mingqi Zhao, Chen Wei, Dante Mantini, Zherui Li, and Quanying Liu. EEGdenoiseNet: a benchmark dataset for deep learning solutions of EEG denoising. *Journal of Neural Engineering*, 18(5):056057, 2021. doi: 10.1088/1741-2552/ac2bf8. URL https://doi.org/10.1088/1741-2552/ac2bf8.

A Pretraining Datasets

Table 5 lists all 523 dataset entries, with links to publications and data records. Quark, Atom, and Molecule contain 9, 521, and 505 entries, respectively. Durations are the recording hours used in each collection; a dash denotes absence. Accessions joined by a plus sign form one entry.

Numbered citations identify publications in the References; Data links point to repository records. A dagger marks a related publication whose exact correspondence to the listed dataset remains unconfirmed;

unresolved aliases are labeled explicitly. The accompanying CSV files contain full source metadata, recording hours, and training-window counts.

Table 5. Pretraining datasets, source links, and recording hours used in each collection.

Dataset or accession	Source	Quark (h)	Atom (h)	Molecule (h)
Named datasets and other repositories				
Alljoined	Data; [19]	370.96	–	461.03
BCI-IV-2a	Data; [33]	–	7.82	18.23
BCIC2020_3	Data; [34]	–	5.00	5.00
CHB-MIT	Data; [15]	–	991.45	1,963.72
EEGMAT	Data; [35]	–	1.44	2.37
EEGMMIDB	Data; [36]	–	16.14	22.52
EEGWorkload	Data 1; Data 2; [35]†	–	1.68	–
FACED	Data; [37]	–	17.01	30.33
figshare_12636731	Data; [38]	–	3.53	20.56
figshare_13067018	Data	–	17.96	14.79
figshare_14709420	Data; [39]	–	103.91	80.45
figshare_14891607	Data; [40]	–	20.54	20.45
figshare_16840351	Data; [41]	–	0.22	2.49
figshare_21644429	Data; [42]	–	1.34	76.45
figshare_22647313	Data; [43]	–	10.01	10.70
figshare_23641092	Code; [44]†	–	0.08	0.02
figshare_23677098	Data; [45]	–	43.45	31.55
figshare_24485689	Data	–	63.87	66.04
figshare_24631257	Data	–	5.29	14.37
figshare_24760827	Data; [46]	–	61.88	48.97
figshare_25360300	Data; [47]	–	75.69	176.46
figshare_25583334	Data; [48]	–	4.80	12.67
figshare_26943565	Data	–	47.64	46.70
figshare_28408334	Data; [49]	–	41.34	24.41
figshare_28452131	Data; [50]	–	55.98	37.59
figshare_28831730	Data; [51]	–	30.34	30.88
figshare_29264621	Data; [52]	–	13.14	13.12
figshare_29902325	Data; [53]	–	5.23	7.14
figshare_5418967	Data; [54]	–	34.58	32.00
figshare_5616304	Data; [55]	–	28.87	27.70
figshare_6427334	Data; [56]	–	11.20	80.00
figshare_7016369	Data; [57]	–	23.45	19.44
figshare_7658315	Data; [58, 59]	–	75.03	59.11
self-collected dataset		27.74	27.74	21.67
HMC	Data; [16]	740.88	17.35	1,141.80
I-CARE	Data; [14]	–	1,657.46	28,583.16
ISRUC_S3	Data; [60]	–	23.89	74.07
mendeley_j2v7btchdy	Data	–	2.96	2.98
mendeley_yj52xrfgtz	Data	–	5.99	5.47
MOABB_AlexMI	Data; [61]	–	0.54	1.10
MOABB_BI2014a	Data; [62]	–	7.47	11.87
MOABB_BNCI2014_004	Data; [63]	–	14.48	26.26
MOABB_BNCI2014_008	[64]	–	1.49	3.02
MOABB_BNCI2014_009	[65]	–	0.93	1.60
MOABB_BNCI2015_001	[66]	–	11.29	16.68
MOABB_BNCI2015_003	[67]	–	0.55	0.80
MOABB_BNCI2015_004	[68]	–	7.74	13.75
MOABB_Cho2017	[69]	–	12.60	19.99
MOABB_Dreyer2023	[70]	–	16.29	61.96
MOABB_EPFLP300	[71]	–	1.40	3.00
MOABB_Lee2019_MI	[72]	–	5.59	42.62
MOABB_Lee2019_SSVEP	[72]	–	4.53	40.24
MOABB_Liu2024	[73]	–	2.67	4.44
MOABB_MAMEM1	Data; [74]	–	5.50	5.68

Continued on the next page.

Table 5. Pretraining datasets, sources, and recording hours (continued).

Dataset or accession	Source	Quark (h)	Atom (h)	Molecule (h)
MOABB_MAMEM2	Data; [74]	–	2.74	4.97
MOABB_MAMEM3	Data; [74]	–	2.35	4.59
MOABB_Nakanishi2015	[75]	–	1.18	2.12
MOABB_Ofner2017	[76]	–	8.00	26.93
MOABB_Schirrmeister2017	[77]	–	16.20	28.65
MOABB_Shin2017A	[78]	–	8.43	14.55
MOABB_Wang2016	[79]	–	8.49	14.50
MOABB_Weibo2014	[80]	–	7.79	–
Mumtaz2016-MDD	Data; [81]	–	12.86	32.62
NeuroLM/confidence-figure	[82]†	–	25.60	25.24
NeuroLM_smallsets/angersurprise	[83]†	–	44.46	40.12
NeuroLM_smallsets/bcicompetitioniv1	Data; [84]	–	8.19	8.21
NeuroLM_smallsets/confidence-text	[85]†	–	46.48	43.28
NeuroLM_smallsets/emobrain	[86]	–	4.91	3.27
NeuroLM_smallsets/grasplifteeg	[87]	–	11.22	11.74
NeuroLM_smallsets/inriabcichallenge	Data; [88]	–	29.67	29.98
NeuroLM_smallsets/luoshuaiart	[89]†	–	23.76	22.63
NeuroLM_smallsets/raweegdata	Data; [90]	–	34.19	33.42
NeuroLM_smallsets/restingstate	Data; [91]	–	2.96	3.03
NeuroLM_smallsets/spisrestingstate	Data; [92]	–	0.80	0.84
OSF_4fpru	Data; [93]	–	0.76	0.76
OSF_epu3r	Data; [94]	–	62.44	–
OSF_q3u7j	Data; [95]	–	4.90	28.70
SEED	Data; [21]	15.52	8.84	–
SEED-DV	Data; [96]	–	21.23	–
SEED-IV	Data; [22]	18.17	8.94	18.17
SEED-SD	Data; [97]	–	200.58	–
SEED-V	Data; [23]	13.48	8.93	13.44
SEED-VIG	Data; [98]	–	45.23	–
SEED-VII	Data; [24]	43.26	8.85	43.26
SEED_FRA	Data; [99]	–	25.46	15.86
SEED_GER	Data; [99]	–	25.99	13.80
SEED_OIL	Unresolved	–	23.99	14.77
SHU_MI	Data; [100]	–	9.03	–
Siena	Data; [101]	–	47.53	165.57
SparrKULee	Data; [102]	–	154.93	154.27
THINGS-EEG1	Data; [103]	–	91.73	32.99
THINGS-EEG2	Data; [20]	183.78	–	228.22
TUH	Data; [13]	23,653.42	22,554.14	22,158.28
OpenNeuro				
ON_ds001785	Data	–	24.21	21.65
ON_ds001787	Data	–	27.55	26.91
ON_ds001810	Data	–	181.97	3.69
ON_ds001849	Data	–	20.56	20.46
ON_ds001971	Data	–	39.41	39.39
ON_ds002034	Data; [104]	–	35.63	35.51
ON_ds002094	Data	–	19.53	19.28
ON_ds002158	Data; [105]	–	21.01	17.08
ON_ds002218	Data	–	16.48	16.44
ON_ds002336	Data; [106]	–	5.73	5.70
ON_ds002338	Data; [107]†	–	8.26	8.38
ON_ds002578	Data	–	1.45	1.45
ON_ds002680	Data	–	11.71	11.97
ON_ds002691	Data	–	3.90	3.92
ON_ds002718	Data; [108]†	–	8.23	8.24
ON_ds002720	Data; [109]†	–	18.40	18.60
ON_ds002721	Data; [110]†	–	26.01	26.19

Continued on the next page.

Table 5. Pretraining datasets, sources, and recording hours (continued).

Dataset or accession	Source	Quark (h)	Atom (h)	Molecule (h)
ON_ds002722	Data; [109]†	–	22.75	22.86
ON_ds002723	Data; [111]†	–	10.40	10.46
ON_ds002724	Data; [109]†	–	28.49	28.57
ON_ds002725	Data; [112]†	–	22.41	22.34
ON_ds002778	Data; [113]	–	2.44	2.51
ON_ds002791+ds003420	Data 1; Data 2	–	13.41	13.13
ON_ds002814	Data	–	19.01	18.86
ON_ds002833+ds003421	Data 1; Data 2	–	11.63	11.49
ON_ds003004	Data; [114]†	–	28.96	28.99
ON_ds003039	Data; [115]†	–	13.68	13.69
ON_ds003061	Data	–	4.21	4.23
ON_ds003190	Data	–	39.12	39.12
ON_ds003194	Data	–	7.09	7.14
ON_ds003195	Data	–	4.59	4.63
ON_ds003458	Data	–	5.90	5.34
ON_ds003474	Data; [116]†	–	36.40	17.17
ON_ds003478	Data	–	22.91	13.11
ON_ds003490	Data	–	12.64	12.24
ON_ds003506	Data; [117]†	–	20.59	17.11
ON_ds003509	Data; [118]†	–	48.41	45.98
ON_ds003516	Data; [119]†	–	13.54	13.52
ON_ds003517	Data; [120]	–	12.95	12.90
ON_ds003518	Data; [121]†	–	86.40	83.26
ON_ds003519	Data; [122]†	–	12.88	12.87
ON_ds003522	Data; [123]†	–	33.49	30.35
ON_ds003523	Data; [122]†	–	48.80	42.13
ON_ds003555	Data	–	104.72	85.96
ON_ds003570	Data; [124]	–	18.30	14.42
ON_ds003574	Data	–	24.92	24.82
ON_ds003602	Data	–	147.94	148.74
ON_ds003620	Data; [125]†	–	284.66	–
ON_ds003626	Data; [126]†	–	12.91	12.94
ON_ds003638	Data	–	40.50	38.98
ON_ds003655	Data	–	77.80	77.93
ON_ds003670	Data	–	72.66	16.38
ON_ds003690	Data; [127]†	–	28.30	27.82
ON_ds003702	Data; [128]†	–	40.38	<0.01
ON_ds003710	Data; [129]†	–	9.09	9.09
ON_ds003739	Data; [130]†	–	20.38	20.11
ON_ds003751	Data	–	4.59	3.92
ON_ds003753	Data	–	10.06	9.09
ON_ds003766	Data; [131]	–	73.40	40.51
ON_ds003768	Data	–	59.87	58.13
ON_ds003774	Data	–	19.32	7.88
ON_ds003775	Data; [132]†	–	10.13	10.13
ON_ds003800	Data	–	0.74	0.77
ON_ds003801	Data	–	13.68	13.61
ON_ds003810	Data	–	5.10	5.20
ON_ds003822	Data	–	12.83	12.39
ON_ds003846	Data; [133]†	–	22.64	18.96
ON_ds003885	Data; [134]	–	53.89	24.96
ON_ds003887	Data	–	53.54	23.61
ON_ds003969	Data; [135]†	–	66.13	65.05
ON_ds003987	Data	–	51.98	51.04
ON_ds004000	Data	–	11.97	12.31
ON_ds004010	Data; [136]†	–	26.41	26.35
ON_ds004017	Data	–	23.28	10.73
ON_ds004018	Data; [137]†	–	0.79	0.79

Continued on the next page.

Table 5. Pretraining datasets, sources, and recording hours (continued).

Dataset or accession	Source	Quark (h)	Atom (h)	Molecule (h)
ON_ds004019	Data	–	19.01	18.62
ON_ds004022	Data	–	4.66	4.70
ON_ds004024	Data	–	53.90	54.20
ON_ds004040	Data	–	25.52	9.73
ON_ds004043	Data; [138]†	–	36.41	16.38
ON_ds004067	Data	–	58.77	46.24
ON_ds004075	Data	–	12.39	12.54
ON_ds004105	Data; [139]†	–	26.13	25.40
ON_ds004106	Data; [140]†	–	24.31	24.22
ON_ds004118	Data; [141]†	–	65.52	65.28
ON_ds004119	Data; [140]†	–	19.52	19.32
ON_ds004120	Data; [141]†	–	132.93	127.86
ON_ds004121	Data; [141]†	–	31.03	28.15
ON_ds004122	Data; [141]†	–	45.86	37.83
ON_ds004123	Data; [141]†	–	23.09	22.29
ON_ds004147	Data; [142]†	–	9.59	9.57
ON_ds004148	Data	–	118.44	74.98
ON_ds004151	Data; [143]	–	13.27	13.18
ON_ds004152	Data	–	11.41	11.43
ON_ds004196	Data; [144]	–	1.46	1.46
ON_ds004200	Data; [145]†	–	14.08	13.41
ON_ds004252	Data; [146]	–	10.67	8.04
ON_ds004256	Data	–	42.26	26.71
ON_ds004262	Data	–	8.30	8.34
ON_ds004264	Data	–	7.41	7.46
ON_ds004279	Data	–	53.63	52.58
ON_ds004284	Data	–	9.43	8.91
ON_ds004295	Data	–	34.27	32.91
ON_ds004315	Data	–	20.35	21.01
ON_ds004317	Data	–	37.68	37.29
ON_ds004324	Data	–	11.50	10.90
ON_ds004347	Data; [147]	–	6.34	5.18
ON_ds004348	Data; [148]†	–	160.90	82.97
ON_ds004350	Data; [149]†	–	81.86	30.12
ON_ds004356	Data; [150]	–	46.21	38.37
ON_ds004357	Data; [151]†	–	11.28	9.47
ON_ds004362	Data	–	94.44	48.67
ON_ds004367	Data	–	24.74	13.23
ON_ds004381	Data; [152]†	–	19.25	11.62
ON_ds004389	Data	–	29.90	30.20
ON_ds004395	Data; [17]	–	728.70	6,826.75
ON_ds004408	Data	–	20.08	18.71
ON_ds004444	Data; [153]†	–	55.03	41.56
ON_ds004446	Data; [153]†	–	33.22	31.14
ON_ds004447	Data; [153]†	–	23.22	17.72
ON_ds004448	Data; [153]†	–	43.41	42.73
ON_ds004475	Data; [154]†	–	26.85	26.87
ON_ds004477	Data	–	13.55	11.14
ON_ds004502	Data	–	73.79	45.24
ON_ds004504	Data; [155]	–	38.61	19.59
ON_ds004505	Data; [156]†	–	82.86	30.37
ON_ds004514	Data	–	14.59	13.29
ON_ds004515	Data	–	20.53	20.57
ON_ds004517	Data	–	7.68	7.65
ON_ds004519	Data; [157]	–	59.17	39.50
ON_ds004520	Data; [158]	–	49.02	32.73
ON_ds004521	Data; [158]	–	50.31	33.59
ON_ds004532	Data; [121]†	–	49.44	48.79

Continued on the next page.

Table 5. Pretraining datasets, sources, and recording hours (continued).

Dataset or accession	Source	Quark (h)	Atom (h)	Molecule (h)
ON_ds004541	Data	–	5.63	1.33
ON_ds004554	Data; [159]†	–	6.57	–
ON_ds004561	Data	–	11.34	11.35
ON_ds004563	Data	–	178.38	60.01
ON_ds004572	Data; [160]†	–	52.53	53.23
ON_ds004574	Data; [161]†	–	30.80	30.75
ON_ds004577	Data; [162]†	–	22.55	22.84
ON_ds004579	Data; [161]†	–	55.46	54.26
ON_ds004580	Data; [161]†	–	36.25	36.39
ON_ds004582	Data	–	34.13	33.68
ON_ds004584	Data; [161]†	–	6.43	6.61
ON_ds004587	Data	–	25.10	25.26
ON_ds004595	Data	–	9.90	9.93
ON_ds004602	Data	–	52.22	50.19
ON_ds004603	Data; [163]†	–	30.59	23.27
ON_ds004621	Data; [164]†	–	45.19	44.14
ON_ds004625	Data	–	28.27	28.48
ON_ds004626	Data; [165]	–	21.27	21.00
ON_ds004635	Data; [166]†	–	0.31	2.31
ON_ds004657	Data	–	27.05	26.69
ON_ds004660	Data	–	14.80	13.74
ON_ds004661+ds004849+ds004850+ ds004852+ds004853+ds004854+ ds004855	Data 1; Data 2; Data 3; Data 4; Data 5; Data 6; Data 7	–	10.11	9.66
ON_ds004706	Data; [167]†	–	470.22	412.94
ON_ds004718	Data; [168]	–	41.11	6.81
ON_ds004771	Data	–	10.41	–
ON_ds004784	Data; [169]	–	0.92	0.54
ON_ds004785	Data	–	0.87	–
ON_ds004796	Data; [170]†	–	43.74	44.27
ON_ds004802	Data	–	40.74	16.08
ON_ds004816	Data; [171]†	–	21.37	10.16
ON_ds004817	Data; [171]†	–	23.75	10.73
ON_ds004840	Data	–	10.86	10.93
ON_ds004841	Data	–	28.26	27.37
ON_ds004842	Data	–	20.04	18.65
ON_ds004843	Data	–	15.98	16.00
ON_ds004844	Data	–	21.17	20.44
ON_ds004851	Data	–	18.13	18.16
ON_ds004860	Data	–	16.31	15.88
ON_ds004883	Data	–	139.08	124.40
ON_ds004902	Data; [172]	–	17.69	17.70
ON_ds004917	Data; [173]†	–	14.53	14.55
ON_ds004942	Data	–	28.19	27.14
ON_ds004951	Data	–	25.75	22.94
ON_ds004952	Data; [174]	–	122.82	52.58
ON_ds004980	Data	–	36.83	30.56
ON_ds004995	Data; [175]	–	16.16	12.97
ON_ds005021	Data	–	47.41	42.06
ON_ds005034	Data	–	34.77	28.93
ON_ds005079	Data	–	3.80	3.80
ON_ds005087	Data; [176]†	–	26.86	13.05
ON_ds005089	Data	–	62.74	50.47
ON_ds005095	Data	–	16.48	16.09
ON_ds005106	Data; [177]†	–	5.43	3.17
ON_ds005114	Data; [178]†	–	123.32	105.49
ON_ds005121	Data; [179]†	–	40.45	34.09
ON_ds005170	Data	–	100.66	74.70

Continued on the next page.

Table 5. Pretraining datasets, sources, and recording hours (continued).

Dataset or accession	Source	Quark (h)	Atom (h)	Molecule (h)
ON_ds005178	Data	–	1,145.83	86.33
ON_ds005189	Data	–	19.27	17.32
ON_ds005262	Data	–	24.77	17.71
ON_ds005273	Data	–	51.39	49.38
ON_ds005274	Data	–	0.83	1.44
ON_ds005280	Data	–	195.95	98.39
ON_ds005284	Data	–	4.73	2.29
ON_ds005285	Data	–	53.78	26.73
ON_ds005286	Data	–	42.21	18.85
ON_ds005289	Data	–	32.77	16.56
ON_ds005291	Data	–	92.06	45.05
ON_ds005292	Data	–	134.69	60.04
ON_ds005293	Data	–	469.96	177.42
ON_ds005296	Data	–	37.07	36.42
ON_ds005305	Data; [180]	–	13.87	14.22
ON_ds005307	Data	–	1.51	1.57
ON_ds005340	Data; [181]	–	35.26	35.16
ON_ds005342	Data; [182]	–	32.94	33.00
ON_ds005343	Data	–	0.32	2.03
ON_ds005345	Data	–	19.15	19.40
ON_ds005363	Data	–	41.17	40.15
ON_ds005383	Data	–	35.65	8.53
ON_ds005385	Data; [183]†	–	165.92	155.18
ON_ds005397	Data; [184]†	–	27.87	27.85
ON_ds005403	Data	–	13.34	13.36
ON_ds005406	Data	–	15.41	13.28
ON_ds005408	Data	–	56.79	56.57
ON_ds005410	Data; [185]†	–	22.94	22.91
ON_ds005416	Data	–	24.64	0.21
ON_ds005420	Data	–	5.27	5.40
ON_ds005429	Data	–	14.30	14.39
ON_ds005473	Data	–	29.14	14.45
ON_ds005486	Data	–	55.99	51.57
ON_ds005505	Data; [186]†	–	100.24	54.78
ON_ds005506	Data; [186]†	–	107.87	60.01
ON_ds005507	Data; [186]†	–	123.36	66.54
ON_ds005508	Data; [186]†	–	219.08	142.70
ON_ds005509	Data; [186]†	–	182.89	82.60
ON_ds005510	Data; [186]†	–	86.77	41.77
ON_ds005511	Data; [186]†	–	208.25	135.47
ON_ds005512	Data; [186]†	–	132.62	43.73
ON_ds005514	Data; [186]†	–	181.68	24.73
ON_ds005515	Data; [186]†	–	179.61	29.29
ON_ds005516	Data; [186]†	–	245.71	–
ON_ds005520	Data	–	40.31	48.61
ON_ds005530	Data	–	90.40	66.59
ON_ds005540	Data; [37]†	–	166.20	102.46
ON_ds005555	Data; [187]†	–	2,002.10	1,946.74
ON_ds005565	Data	–	11.38	11.41
ON_ds005571	Data; [188]†	–	28.13	27.82
ON_ds005586	Data	–	33.48	24.64
ON_ds005594	Data; [189]†	–	12.28	12.10
ON_ds005620	Data	–	17.85	14.16
ON_ds005628	Data	–	9.18	20.87
ON_ds005642	Data	–	36.90	18.32
ON_ds005648	Data	–	11.54	8.13
ON_ds005662	Data	–	80.35	78.12
ON_ds005672	Data	–	4.58	4.56

Continued on the next page.

Table 5. Pretraining datasets, sources, and recording hours (continued).

Dataset or accession	Source	Quark (h)	Atom (h)	Molecule (h)
ON_ds005697	Data	–	69.64	67.33
ON_ds005779	Data	–	19.30	19.43
ON_ds005795	Data	–	7.90	7.93
ON_ds005811	Data	–	23.25	23.74
ON_ds005815	Data	–	33.53	11.62
ON_ds005841	Data	–	11.10	11.13
ON_ds005857	Data	–	100.61	86.35
ON_ds005863+ds006018	Data 1; Data 2	–	50.25	50.40
ON_ds005866	Data	–	15.85	15.97
ON_ds005868	Data	–	13.00	13.01
ON_ds005872+ds007338	Data 1; Data 2	–	0.09	0.09
ON_ds005876	Data	–	15.95	15.98
ON_ds005907	Data	–	8.37	8.46
ON_ds005932	Data	–	9.88	9.94
ON_ds005946	Data	–	18.29	14.56
ON_ds005960	Data; [190]†	–	66.84	60.23
ON_ds006033	Data	–	2.18	2.19
ON_ds006036	Data; [155]†	–	14.58	7.67
ON_ds006040	Data	–	107.02	38.32
ON_ds006095	Data	–	59.66	60.13
ON_ds006104	Data	–	51.49	18.70
ON_ds006110	Data; [191]†	–	45.73	45.59
ON_ds006142	Data	–	16.98	16.20
ON_ds006159	Data	–	14.28	12.10
ON_ds006171	Data; [192]†	–	40.73	36.74
ON_ds006222	Data	–	52.64	52.09
ON_ds006260	Data	–	22.76	18.64
ON_ds006317	Data	–	62.05	62.02
ON_ds006367	Data	–	68.12	74.82
ON_ds006370	Data	–	95.20	169.88
ON_ds006374	Data	–	40.47	31.61
ON_ds006386	Data	–	96.19	51.27
ON_ds006394	Data	–	17.56	17.38
ON_ds006434	Data	–	262.87	252.36
ON_ds006437	Data	–	16.73	16.80
ON_ds006446	Data	–	17.98	17.96
ON_ds006465	Data	–	28.23	26.30
ON_ds006466	Data; [193]†	–	129.70	129.51
ON_ds006480	Data; [193]†	–	71.52	68.33
ON_ds006525	Data	–	5.67	5.38
ON_ds006547	Data	–	23.71	17.24
ON_ds006554	Data	–	25.86	22.57
ON_ds006563	Data; [194]	–	12.62	12.64
ON_ds006576	Data	–	110.66	108.56
ON_ds006593	Data	–	5.56	5.61
ON_ds006647	Data	–	8.67	8.66
ON_ds006648	Data	–	91.74	73.46
ON_ds006695	Data; [195]†	–	164.22	156.16
ON_ds006735	Data	–	38.78	34.93
ON_ds006761	Data; [196]	–	25.57	12.40
ON_ds006768	Data	–	14.20	13.74
ON_ds006801	Data	–	6.30	6.35
ON_ds006802	Data; [197]†	–	23.38	15.45
ON_ds006803	Data	–	45.64	43.96
ON_ds006817	Data	–	21.63	20.10
ON_ds006839	Data	–	26.40	26.49
ON_ds006840	Data	–	14.60	14.78
ON_ds006848	Data	–	47.33	36.87

Continued on the next page.

Table 5. Pretraining datasets, sources, and recording hours (continued).

Dataset or accession	Source	Quark (h)	Atom (h)	Molecule (h)
ON_ds006850	Data	–	78.12	76.60
ON_ds006861	Data	–	59.19	59.18
ON_ds006866	Data	–	72.68	70.39
ON_ds006897	Data	–	47.85	15.15
ON_ds006921	Data	–	16.87	16.05
ON_ds006923	Data	–	35.29	35.15
ON_ds006945	Data	–	2.08	2.11
ON_ds006963	Data; [198]	–	8.70	22.04
ON_ds006979	Data	–	46.51	23.57
ON_ds007006	Data	–	3.53	3.60
ON_ds007012	Data	–	78.05	76.47
ON_ds007020	Data	–	3.97	4.08
ON_ds007052	Data	–	23.58	23.69
ON_ds007056	Data	–	20.64	20.91
ON_ds007069	Data	–	32.36	32.04
ON_ds007096	Data	–	30.59	30.36
ON_ds007137	Data	–	31.99	32.33
ON_ds007139	Data	–	38.19	37.62
ON_ds007162	Data; [199]†	–	69.87	64.70
ON_ds007169	Data	–	2.82	2.84
ON_ds007172	Data; [200]†	–	30.13	30.52
ON_ds007175	Data	–	45.92	44.68
ON_ds007176+ds007427	Data 1; Data 2; [201, 202]	–	28.29	28.62
ON_ds007180	Data	–	34.68	20.45
ON_ds007181	Data	–	447.89	447.32
ON_ds007216	Data	–	99.32	33.28
ON_ds007221	Data	–	132.70	109.54
ON_ds007222	Data	–	3.68	0.28
ON_ds007262	Data	–	4.53	4.58
ON_ds007314+ds007315	Data 1; Data 2; [203]	–	4.87	4.26
ON_ds007322	Data	–	30.06	27.18
ON_ds007347	Data	–	4.46	3.32
ON_ds007353	Data	–	20.62	20.61
ON_ds007406	Data	–	0.49	0.53
ON_ds007431	Data	–	160.09	119.20
ON_ds007454	Data	–	34.45	32.81
ON_ds007471	Data	–	18.72	18.16
ON_ds007521	Data	–	33.34	30.50
ON_ds007526	Data; [204]†	–	19.31	18.43
ON_ds007537	Data	–	6.60	5.83
ON_ds007554	Data	–	30.02	28.08
ON_ds007558+ds007666+ds007667	Data 1; Data 2; Data 3	–	29.54	29.76
ON_ds007609	Data; [205]	–	7.96	1.96
ON_ds007615	Data; [206]†	–	18.20	18.47
ON_ds007629	Data; [207]†	–	0.98	0.98
ON_ds007647	Data	–	15.46	15.54
ON_ds007648	Data; [208]†	–	16.32	15.50
ON_ds007655	Data	–	20.30	19.59
ON_ds007688	Data	–	13.21	12.69
ON_ds007788	Data	–	32.45	34.04
ON_ds007813	Data	–	1.10	1.09
ON_ds007823	Data; [209]	–	35.90	37.22
ON_ds007827	Data	–	44.24	35.90
Zenodo				
ZEN_10200985	Data; [210]	–	89.53	89.70
ZEN_10229480	Data	–	6.86	6.93
ZEN_10498	Data	–	1.64	1.61

Continued on the next page.

Table 5. Pretraining datasets, sources, and recording hours (continued).

Dataset or accession	Source	Quark (h)	Atom (h)	Molecule (h)
ZEN_10512414	Data; [211]†	–	23.41	23.81
ZEN_10518106	Data; [93]	–	168.11	162.61
ZEN_10696430	Data; [212]	–	32.59	17.48
ZEN_10776597	Data; [213]	–	53.24	21.16
ZEN_10980117	Data; [214]	–	0.03	–
ZEN_11493481	Data; [215]	–	35.37	164.45
ZEN_1170529	Data	–	17.43	4.24
ZEN_13003	Data; [216]	–	28.49	28.46
ZEN_1308615	Data	–	62.73	60.64
ZEN_13341214	Data	–	30.43	30.11
ZEN_13356484	Data	–	16.57	15.47
ZEN_13969130	Data; [217]†	–	18.27	17.68
ZEN_14638129	Data; [218]	–	32.01	35.24
ZEN_14638353	Data; [219]	–	47.39	38.44
ZEN_15454355	Data	–	84.05	75.70
ZEN_15516012	Data; [220]	–	21.04	40.07
ZEN_15665101	Data; [221]	–	21.01	21.28
ZEN_15665660	Data	–	8.31	8.29
ZEN_17178030	Data	–	103.92	102.02
ZEN_20044	Data	–	7.36	7.36
ZEN_2222268	Data; [222]	–	8.89	–
ZEN_2547147	Data	–	111.71	103.40
ZEN_2594667	Data; [223, 224]	–	32.09	32.06
ZEN_29601	Data; [225]	–	12.85	–
ZEN_3581862	Data; [226]	–	63.19	60.52
ZEN_3618205	Data; [227]†	–	82.34	78.59
ZEN_3992359	Data	–	10.81	6.56
ZEN_4002038	Data	–	53.00	50.55
ZEN_4046953	Data	–	34.97	34.26
ZEN_4385970	Data; [228]	–	17.20	17.04
ZEN_4471943	Data	–	89.94	82.32
ZEN_4500849	Data	–	111.84	112.69
ZEN_4501988	Data	–	96.22	85.77
ZEN_4579401	Data; [229]	–	25.10	2.59
ZEN_4579450	Data; [230]	–	74.29	98.37
ZEN_4917218	Data	–	8.25	0.50
ZEN_5002434	Data; [231]	–	7.88	6.75
ZEN_5035511	Data; [232]	–	18.74	17.23
ZEN_5089803	Data; [233]	–	21.55	20.31
ZEN_5121167	Data; [234]	–	11.20	11.40
ZEN_5233656	Data; [235]†	–	27.66	27.60
ZEN_5512578	Data	–	35.17	18.65
ZEN_5513267	Data; [236]	–	40.48	40.52
ZEN_57911	Data; [135]	–	25.39	25.70
ZEN_581188	Data	–	14.35	14.64
ZEN_5907009	Data; [237]	–	9.32	9.43
ZEN_6225534	Data; [238]	–	37.31	38.70
ZEN_6501824	Data; [239]	–	1.33	1.00
ZEN_6642860	Data; [240]	–	55.74	29.18
ZEN_6801531	Data; [241]	–	15.06	8.80
ZEN_6874129	Data	–	56.96	81.60
ZEN_6973571	Data; [242, 243]	–	123.34	71.42
ZEN_6992085	Data; [244]	–	111.80	110.98
ZEN_7149276	Data; [245]	–	9.76	–
ZEN_7196042	Data	–	25.23	24.29
ZEN_7428378	Data; [246]	–	1.54	0.28
ZEN_7496795	Data	–	9.41	9.50
ZEN_7650679	Data; [247]	–	47.51	46.37

Continued on the next page.

Table 5. Pretraining datasets, sources, and recording hours (continued).

Dataset or accession	Source	Quark (h)	Atom (h)	Molecule (h)
ZEN_7664069	Data	–	15.18	15.20
ZEN_7667391	Data	–	24.19	15.54
ZEN_7891964	Data	–	10.00	–
ZEN_8032879	Data; [248]	–	20.24	0.02
ZEN_8086086	Data	–	11.30	11.44
ZEN_8089820	Data; [70]	–	71.88	70.07
ZEN_8255618	Data	–	3.38	3.46
ZEN_8329025	Data; [249]	–	32.03	16.40
ZEN_8369136	Data	–	0.31	31.97
ZEN_8380796	Data	–	43.90	48.88
ZEN_8407716	Data; [250]	–	169.24	149.78

B Pretraining Data Details

B.1 Source composition

Table 6 gives the source breakdown behind Figure 2. Categories summarize the predominant recording setting or task. Quark uses sampling weight 1.0 for TUH, Alljoined, and THINGS-EEG2; 0.1 for HMC; and 0.3 for the self-collected dataset and the four SEED datasets. Their recording hours appear in Table 2.

Table 6. Source composition of Atom and Molecule.

Category	Atom (h)	Share	Molecule (h)	Share
Routine clinical (TUH)	22,554	45.4%	22,158	27.7%
Cognitive / ERP	6,358	12.8%	11,420	14.3%
Unclassified / mixed	6,235	12.6%	4,076	5.1%
Sleep PSG	4,704	9.5%	4,551	5.7%
Auditory / speech	1,954	3.9%	1,459	1.8%
ICU (I-CARE)	1,657	3.3%	28,583	35.8%
Motor imagery	1,333	2.7%	1,429	1.8%
Epilepsy, excluding TUH	1,229	2.5%	2,309	2.9%
Other clinical cohorts	909	1.8%	822	1.0%
Emotion	793	1.6%	726	0.9%
Resting state	790	1.6%	699	0.9%
SSVEP / P300 and related BCI	635	1.3%	588	0.7%
Visual image evoked	477	1.0%	1,089	1.4%

Hours and percentages are rounded independently.

B.2 Windowing, sampling, and padding

For channel count C, Molecule uses $\min(16, \max(4, \lfloor 256/C \rfloor))$ seconds per window, except for the one-second Alljoined and THINGS-EEG2 trials. Above 64 channels, the four-second minimum permits more than 256 channel-by-time tokens. Quark truncates longer samples to 16 one-second patches.

At retained fraction f, the loader keeps every round$(1/f)$-th window within each source/view stream. Selected windows recur across epochs.

The node-local shard cache defaults to 80% of disk capacity; the loader supports persistent workers, four-batch prefetching, and pinned host memory.

The student uses patch-validity weights and excludes fully padded patches as described in Section 3. The tokenizer's reconstruction loss includes zero-padded patches.

B.3 Molecule preparation settings

Table 7 specifies the signal preparation and routing rules used to rebuild Molecule. Quark and Atom retain their earlier preparation pipelines. Atom and Quark use input scale 0.01; Molecule uses 1.0 after recording- or trial-level normalization. Window construction requires finite values, identified channels, and explicit units and padding metadata.

During training, the Atom and Molecule input paths sanitize non-finite values and clip tokenizer inputs to ± 20. For Molecule, the same sanitation is applied before frozen-target extraction and encoder processing.

Table 7. Signal preparation and window routing for Molecule.

Property	Setting
Signal and window construction	
Sampling and filtering	200 Hz; band-pass 0.5–min$(99.5,\ 0.99 f_{\text{Nyquist}})$ Hz; no notch filter.
Bad channels	Remove exactly constant channels and channels with non-finite fraction $\geq 1\%$.
Finite segments	Split the remaining signal into contiguous finite runs of at least 200 samples; no non-finite imputation.
Normalization	Per-channel z-score within each recording or trial, with $\epsilon = 10^{-7}$; clip to ± 15; store in float16; training input scale 1.0.
Window duration	$\min(16, \max(4, \lfloor 256/C \rfloor))$ s for C mapped channels. Alljoined and THINGS-EEG2 use 1 s trials. Windows do not overlap or cross recording/trial boundaries.
Recording tails	Retain tails of at least 1 s; zero-pad and record sample masks and valid-patch fractions.
Quality flags and sampling	
Mild quality flag	Any channel with $\lvert\mu\rvert > 5$ or $\sigma < 0.05$ or $\sigma > 5$; any clipping or padding; any removed bad channel or reader fallback.
Severe quality flag	Any channel with $\lvert\mu\rvert > 10$ or $\sigma < 0.005$ or $\sigma > 10$; clipped fraction $> 5\%$; padded fraction $\geq 50\%$; or all channels bad/flat.
Exclusion priority	Subject exclusions, non-finite/decoding failures, and severe-quality flags precede auxiliary routing and receive zero training weight.
Auxiliary factors	Mild quality flag: 0.5. One to seven channels: 0.25, or 0.125 with a mild flag. Main view: 1.0.
Sampling factors	Multiply source, channel, quality, and provenance factors. Approved processed fallbacks have provenance factor 0.5; raw recordings and complete trial caches use 1.0.

Signal statistics use valid samples in normalized units; μ and σ denote a channel's window mean and standard deviation. A severe flag overrides a mild flag. Sampling factors determine window acceptance probabilities.

B.4 Tokenizer settings

The three versions share the tokenizer architecture. Each uses independently trained weights for frozen target generation during encoder pretraining.

Table 8. Tokenizer training configurations.

Version	GPUs	Global batch	Peak LR	Warmup	Checkpoint
Quark	32	2,048	8×10^{-4}	10	ep39
Atom	8	2,048	2.4×10^{-4}	15	ep16
Molecule	32	6,144	2.4×10^{-4}	15	ep39

Learning rates include batch-size scaling. Warmup is in epochs; planned schedules are 100, 40, and 40 epochs.

C Downstream Datasets and Evaluation Settings

Task inputs are sampled at 200 Hz and retain auxiliary channels where present. TUEV files have 24–32 channels; the table describes the inspected 32-channel input. Extension models use sample/channel z-scoring and clipping to ± 15; external encoders retain their model-specific input scaling. For inputs longer than 16

seconds, such as 30-second sleep windows, we interpolate the extension encoders' temporal embeddings while retaining one-second patches.

Table 9. Downstream datasets, input configurations, and Seed42 window counts.

Dataset and input configuration						Seed42 windows		
Dataset	Category	Classes	Units	Ch.	Window (s)	Train	Validation	Test
AD65 [155]	Neurodeg.	3	88	19	10	2,790	816	697
PD31 [113, 251]	Neurodeg.	2	31	40	1	1,000	773	581
BCIC IV-1 [33, 84]	Motor imag.	2	7	59	4	400	200	200
BCIC IV-2a [33]	Motor imag.	4	9	22	4	1,120	576	576
SHU MI [100]	Motor imag.	2	25	32	4	1,600	977	1,455
BETA SSVEP [252]	SSVEP	40	70	64	1	8,960	1,120	1,120
Monitoring ErrP [253]	ERP	2	6	64	1	320	1,080	1,044
EEGMAT [35]	Cognitive	2	36	21	5	1,371	181	144
LEMON gender [254]	Trait	2	203	59	1	6,480	18,828	19,734
LEMON extraversion [254]	Trait	2	203	59	1	6,480	18,828	19,734
FACED [37]	Emotion	9	123	32	10	8,232	1,008	1,092
SEED [21]	Emotion	3	15	62	10	1,440	1,986	993
SEED-IV [22]	Emotion	4	15	62	4	1,920	5,010	2,505
SEED-V [23]	Emotion	5	16	62	1	2,600	14,718	7,359
SEED-VII [24]	Emotion	7	20	62	10	4,480	2,734	2,734
ISRUC S2 [60]	Sleep	5	8	6	30	1,200	1,975	1,720
ISRUC S3 [60]	Sleep	5	10	6	30	1,600	792	996
SleepEDF [255]	Sleep	5	153[a]	2	30	22,790	40,989	43,334
HMC [16]	Sleep	5	151	8	30	23,691	13,038	13,838
TUAB [256]	Clin. screen	2	2,328	23	10	76,000	42,352	41,962
TUEV [257]	Clin. event	6	370	32	5	16,409	13,206	12,908
TUSL [258]	Clin. event	3	30	23	10	201	28	16
EEGDenoiseNet [259]	Artifact	2	1[b]	1	2	80	1,351	1,352

Neurodeg.: neurodegenerative disease; Motor imag.: motor imagery; Clin. screen: clinical screening; Clin. event: clinical events. Units are subjects, except [a] 153 nights in SleepEDF and [b] one pseudo-subject in EEGDenoiseNet. Training is capped at 40 samples per subject/class unit; validation and test retain all windows. Other seeds use different splits.

Seeds 42, 10, and 5 control splitting and downstream randomness. Splits target 80/10/10 by registered subject identifier. SleepEDF uses night identifiers as split units, allowing person-level overlap; EEGDenoiseNet uses a within-subject sample split. The training cap applies to Full-FT, LP, and joint EEG–LLM evaluation.

C.1 Optimization and readouts

Full-FT uses a 50-epoch schedule to train the encoder and head with AdamW, class-weighted loss, and label smoothing. The search pairs learning rates $\{10^{-4}, 3 \times 10^{-4}, 10^{-3}\}$ with layer-wise decay $\{0.75, 1.0\}$. LP trains a linear head for 30 epochs at constant learning rates $\{10^{-4}, 5 \times 10^{-4}, 10^{-3}\}$ and weight decay $\{0, 0.05, 0.1\}$. Pooled averages patch features; No-pool flattens them in channel/time order. LP candidates and replication fits reuse memory-mapped frozen features within the same task and split; Full-FT fits each candidate independently.

Within each fit, validation scores select the epoch. Seed42 test scores select hyperparameters for replication with seeds 10 and 5. Combined LP results use the higher three-seed test mean of the available readouts, with Pooled used on ties and for SleepEDF.

BENDR uses its stable Full-FT grid except for BCIC IV-2a. BIOT's released weights include supervised TUAB and TUEV training [27]. Its BCIC IV-2a No-pool result uses the fixed 18-channel coordinate mapping. On EEGDenoiseNet, BIOT receives only the native Cz channel. FEMBA Base uses a Cz-only montage: C3–Cz is $-$Cz, Cz–C4 is $+$Cz, and the other 20 bipolar channels are zero. This task adaptation differs from FEMBA's published input configuration [6].

D Results after Excluding Known Overlapping Sources

The 12-task subset contains BCIC IV-1, SHU MI, BETA SSVEP, Monitoring ErrP, LEMON gender, LEMON extraversion, ISRUC S2, SleepEDF, TUAB, TUEV, TUSL, and EEGDenoiseNet. The nine-task subset additionally removes TUAB, TUEV, and TUSL. These exclusions address identified or potential overlap; recording-level separation remains unverified for some retained task–model pairs.

Table 10. Results after excluding known overlapping sources.

Model	FT-12	FT-9	LP-12	LP-9
Quark-B	60.86	58.53	58.07	55.85
Quark-L	60.82	58.65	60.28	58.00
Quark-H	61.00	57.22	59.28	57.44
Quark-G	62.21	59.22	60.75	59.05
Atom-B	60.63	57.69	58.38	55.35
Atom-L	63.75	60.36	59.41	57.26
Atom-H	63.05	59.79	59.32	57.34
Molecule-B	60.86	58.46	58.27	55.47
Molecule-L	62.20	59.53	60.15	57.35
Molecule-H	63.82	61.03	59.62	57.44
REVE Base	61.01	57.88	56.16	53.62
REVE Large	59.99	55.94	57.82	54.94
LUNA Base	58.72	55.81	54.63	52.25
LUNA Large	59.95	56.78	56.68	53.36
LUNA Huge	59.00	55.14	58.13	55.48
CSBrain	62.07	57.88	58.07	54.29
Official LaBraM Base	58.47	54.80	58.13	55.39
CodeBrain	58.15	54.39	58.62	55.20
NeuroLM	58.27	53.23	57.26	52.90
BrainOmni	57.60	53.68	57.47	55.49
NeuroGPT	57.74	55.86	53.57	52.00
EEGMamba	56.46	52.61	54.87	51.52
CBraMod	51.97	50.67	57.29	54.82
BENDR	52.91	52.02	38.91	39.69
BIOT	59.92	54.57	59.39	54.02
FEMBA Base	57.27	55.53	51.28	48.62
EEGNet	52.28	51.88	—	—
EEGConformer	56.08	56.99	—	—
LaBraM Base, random	50.31	48.64	—	—

Extension models use the same checkpoints as Figure 4. LP-12 and LP-9 use selected-readout LP. A dash indicates an unavailable result.

E Linear-Probing Readout Comparison

Table 11 compares Pooled and No-pool on matched tasks at the last evaluated full-data checkpoints. All 26 models cover the same 22 paired tasks; SleepEDF has only Pooled results. The source files provide the available three-seed results, the Atom-H TUAB trajectory mean, and the combined LP rankings.

Table 11. Matched comparison of LP readouts. Each model is compared on the same tasks under both readouts. Values are mean balanced accuracy (%); Δ is No-pool minus Pooled in percentage points.

Model	Paired tasks	Pooled	No-pool	Δ
Quark-B	22	43.68	52.09	+8.40
Quark-L	22	46.56	53.21	+6.65
Quark-H	22	47.05	51.17	+4.12
Quark-G	22	50.19	52.44	+2.25
Atom-B	22	47.53	51.33	+3.79
Atom-L	22	47.30	52.56	+5.26
Atom-H	22	48.16	52.34	+4.18
Molecule-B	22	45.94	50.58	+4.64
Molecule-L	22	45.52	52.58	+7.06
Molecule-H	22	48.66	52.55	+3.90
REVE Large	22	45.52	50.69	+5.17
REVE Base	22	45.31	49.14	+3.84
LUNA Huge	22	48.40	49.11	+0.71
LUNA Large	22	46.34	47.59	+1.25
CSBrain	22	40.89	48.68	+7.79
BIOT	22	48.44	50.83	+2.38
LUNA Base	22	43.79	46.72	+2.92
Official LaBraM Base	22	43.65	51.19	+7.55
CodeBrain	22	45.07	52.29	+7.22
FEMBA Base	22	39.74	45.97	+6.23
NeuroLM	22	47.92	50.46	+2.54
BrainOmni	22	41.04	50.68	+9.65
NeuroGPT	22	43.58	44.41	+0.83
EEGMamba	22	40.28	46.68	+6.40
CBraMod	22	39.91	50.72	+10.81
BENDR	22	32.74	33.92	+1.18

SleepEDF is excluded from every paired comparison. All 26 encoders cover the same 22 tasks under both readouts and form the matched cohort used in the main analysis. Atom-H TUAB No-pool uses the reported trajectory mean.

E.1 FT–LP ranking agreement

LP and Full-FT can favor different encoders. Among the 26 pretrained models, 80 of 325 pairs (24.6%) reverse order when compared by their 23-task means. At the task level, the rate rises to 37.4% (2,797 of 7,475 comparisons).

The extension encoders outperform the 16 baselines in both settings in 131 of 160 comparisons, while 27 favor the extension encoder in only one setting. Tables 12–14 show the overall, per-encoder, and per-task results.

Table 12. FT and LP comparisons of extension encoders against baselines.

Outcome	23-task mean Model pairs	Individual tasks Pair–task comparisons
FT higher, LP higher	131	1835
FT higher, LP lower	5	602
FT lower, LP higher	22	750
FT lower, LP lower	2	486
Tie in either setting	0	7
Total	160	3,680

Table 13. Ordering relative to the 16 baselines for each extension encoder. Scores average 23 tasks.

Encoder	FT higher LP higher	FT higher LP lower	FT lower LP higher	FT lower LP lower
Quark-B	7	0	8	1
Quark-L	9	0	7	0
Quark-H	11	1	4	0
Quark-G	16	0	0	0
Atom-B	12	1	3	0
Atom-L	16	0	0	0
Atom-H	16	0	0	0
Molecule-B	12	3	0	1
Molecule-L	16	0	0	0
Molecule-H	16	0	0	0
Total	131	5	22	2

Table 14. Task-level comparisons of extension encoders against baselines. Each task contributes 160 comparisons.

Task	FT higher LP higher	FT higher LP lower	FT lower LP higher	FT lower LP lower	Ties
AD65	107	33	10	10	0
BCIC IV-1	33	36	63	24	4
BCIC IV-2a	60	7	78	13	2
BETA SSVEP	128	26	2	4	0
EEGDenoiseNet	82	16	13	49	0
EEGMAT	103	34	12	10	1
FACED	50	7	75	28	0
HMC	100	52	3	5	0
ISRUC S2	72	5	73	10	0
ISRUC S3	113	12	19	16	0
LEMON extraversion	74	65	10	11	0
LEMON gender	93	40	10	17	0
Monitoring ErrP	89	21	32	18	0
PD31	45	7	84	24	0
SEED	82	15	28	35	0
SEED-IV	61	20	45	34	0
SEED-V	91	24	31	14	0
SEED-VII	57	20	34	49	0
SHU MI	44	27	68	21	0
SleepEDF	125	26	3	6	0
TUAB	94	51	3	12	0
TUEV	88	51	3	18	0
TUSL	44	7	51	58	0
Total	1835	602	750	486	7

LP ranks encoders by the linear accessibility of their frozen features. Full-FT also depends on how those features adapt to the target task. The reversals show that a lead under LP can disappear after adaptation, so LP rankings alone are insufficient for choosing an encoder for fine-tuning.

Retaining patch structure improves mean LP accuracy by 4.87 points (Section 5.3). Table 15 examines whether this gain also brings encoder rankings closer to Full-FT. Across the same 26 encoders and 22 tasks, excluding SleepEDF, the model-level reversal rate rises from 21.8% with Pooled to 30.2% with No-pool. The task-level rate also rises, from 34.1% to 38.5%.

Table 15. FT–LP ordering reversals on the same 26 encoders and 22 tasks.

LP readout	Mean-score Reversals	Share	Taskwise Reversals	Share	Taskwise ties
Pooled	71	21.8%	2436	34.1%	45
No-pool	98	30.2%	2754	38.5%	15
Per-task best readout	83	25.5%	2699	37.7%	16

Removing average pooling improves LP accuracy without resolving the ranking gap. LP still uses a frozen encoder and linear head, whereas Full-FT can change the representation for the target task. The persistence of the gap with No-pool reinforces the need to evaluate fine-tuning potential after adaptation, alongside frozen-feature performance.

F Results by Checkpoint and Dataset

Cells report three-seed mean balanced accuracy (%). FT is full fine-tuning; LP uses the historical task-selected readout, with Pooled fixed for SleepEDF. Means cover 23 tasks; epochs are zero-indexed. A dash marks an unavailable result. Molecule-B uses a 50-epoch schedule.

Table 16. Quark checkpoints. Balanced accuracy (%).

Dataset	Quark-B ep49		Quark-L ep49		Quark-H ep49		Quark-G ep29	
	FT	LP	FT	LP	FT	LP	FT	LP
AD65	54.05	53.24	60.54	58.72	62.69	50.55	77.16	59.24
PD31	32.35	58.37	34.36	56.46	43.74	49.47	43.46	51.20
BCIC IV-1	52.33	50.17	49.50	52.17	51.00	51.17	53.50	50.50
BCIC IV-2a	34.43	45.37	39.06	43.46	40.39	38.83	41.67	42.36
SHU MI	60.83	58.58	55.23	58.61	52.95	58.33	54.53	61.43
BETA SSVEP	33.81	11.96	41.43	18.99	32.11	18.45	43.66	24.38
Monitoring ErrP	58.23	55.35	59.18	59.01	55.87	60.33	56.33	66.19
EEGMAT	72.69	62.66	63.59	64.82	68.06	65.59	72.99	60.80
LEMON gender	70.78	73.03	72.41	71.09	70.62	67.78	69.80	67.90
LEMON extraversion	51.79	51.87	49.48	51.14	52.46	51.63	52.43	50.99
FACED	23.32	45.61	37.28	43.91	22.34	40.42	34.19	45.01
SEED	55.83	53.12	59.01	55.35	54.23	51.88	60.84	54.71
SEED-IV	27.23	31.84	30.35	30.62	25.60	30.67	31.78	31.57
SEED-V	25.72	25.36	29.42	26.06	28.09	26.72	27.46	28.32
SEED-VII	18.29	18.04	17.76	19.96	19.52	18.77	21.00	21.05
ISRUC S2	25.14	43.40	25.92	47.46	24.82	42.51	27.91	41.96
ISRUC S3	75.61	72.12	61.32	71.55	86.74	67.26	90.30	72.55
SleepEDF	75.88	64.78	75.79	69.07	75.23	70.15	74.89	71.40
HMC	67.00	52.55	68.03	47.36	70.34	52.92	68.71	53.71
TUAB	82.18	77.76	83.82	76.94	82.71	76.93	82.40	78.18
TUEV	66.64	55.80	69.18	59.65	69.38	53.50	71.00	57.27
TUSL	54.76	60.61	49.01	64.75	64.94	63.94	60.12	62.13
EEGDenoiseNet	98.00	93.54	98.93	94.44	99.94	96.64	99.93	96.67
Mean	52.91	52.83	53.50	53.98	54.51	52.37	57.22	54.33

Table 17. Atom-B: full-data checkpoints. Balanced accuracy (%).

Dataset	Atom-B ep9		Atom-B ep19		Atom-B ep29		Atom-B ep39		Atom-B ep49	
	FT	LP	FT	LP	FT	LP	FT	LP	FT	LP
AD65	57.90	53.42	56.77	55.98	67.88	53.72	59.53	53.58	55.52	52.63
PD31	40.29	49.75	44.29	53.15	37.49	52.10	43.26	51.69	53.74	49.39
BCIC IV-1	53.83	52.50	51.83	51.00	49.00	51.67	51.00	49.33	51.83	50.50
BCIC IV-2a	31.83	45.60	32.35	44.56	33.91	44.97	33.28	42.36	36.69	41.90
SHU MI	51.94	60.25	55.52	60.37	54.12	60.39	55.21	59.86	56.24	58.41
BETA SSVEP	37.08	5.45	38.78	6.43	34.94	10.24	35.77	13.18	31.76	11.64
Monitoring ErrP	55.69	53.42	58.31	54.33	56.32	58.87	55.53	57.62	56.05	57.69

Continued on the next page.

Table 17. Atom-B: full-data checkpoints. Balanced accuracy (%) (continued).

Dataset	Atom-B ep9		Atom-B ep19		Atom-B ep29		Atom-B ep39		Atom-B ep49	
	FT	LP	FT	LP	FT	LP	FT	LP	FT	LP
EEGMAT	73.61	64.66	73.93	72.22	74.69	68.36	78.40	68.06	72.87	66.21
LEMON gender	69.78	66.23	73.03	63.83	74.01	64.16	72.94	64.39	72.37	62.15
LEMON extraversion	50.88	51.69	51.96	51.17	52.23	50.96	51.24	50.63	54.00	50.12
FACED	32.89	46.10	33.86	44.43	31.30	43.46	30.90	43.80	25.02	43.70
SEED	57.79	53.04	59.19	54.74	56.31	51.61	61.71	53.18	59.41	56.75
SEED-IV	32.09	32.61	31.40	33.25	32.19	32.43	29.78	32.43	30.57	32.41
SEED-V	28.82	26.39	30.13	27.32	29.77	27.17	29.27	26.91	26.79	26.34
SEED-VII	19.38	20.13	21.30	20.38	21.13	20.42	19.38	20.45	19.35	20.43
ISRUC S2	24.97	44.62	25.61	44.12	25.36	45.02	22.05	41.89	22.87	41.60
ISRUC S3	79.15	72.53	82.65	71.32	81.48	74.80	82.98	73.11	83.36	73.07
SleepEDF	75.21	67.20	74.62	67.62	74.25	67.80	75.03	67.56	74.34	67.10
HMC	69.46	51.08	68.84	51.75	68.23	51.49	68.79	50.07	68.13	49.04
TUAB	81.88	76.22	83.13	76.53	82.86	76.62	82.87	76.71	82.90	76.03
TUEV	67.82	53.08	71.82	57.43	69.83	59.01	66.48	54.47	70.80	55.39
TUSL	53.03	65.29	48.80	67.28	54.08	62.06	48.96	61.05	54.72	70.96
EEGDenoiseNet	98.96	97.93	99.46	97.85	99.56	98.63	99.91	97.91	99.72	98.93
Mean	54.10	52.57	55.11	53.35	54.82	53.30	54.53	52.62	54.74	52.71

Table 18. Atom-L: full-data checkpoints. Balanced accuracy (%).

Dataset	Atom-L ep9		Atom-L ep19		Atom-L ep29		Atom-L ep39		Atom-L ep49	
	FT	LP	FT	LP	FT	LP	FT	LP	FT	LP
AD65	67.53	60.69	59.44	60.85	63.05	58.06	60.01	58.84	59.86	57.16
PD31	48.67	50.98	39.92	51.31	55.48	54.63	42.84	53.48	37.12	53.52
BCIC IV-1	47.50	51.67	50.67	50.00	50.00	50.00	51.83	49.00	53.83	52.33
BCIC IV-2a	41.38	43.75	36.40	42.77	42.65	41.90	34.49	43.17	34.32	41.61
SHU MI	57.31	59.28	56.99	59.26	57.37	60.46	52.11	60.43	52.60	58.97
BETA SSVEP	45.09	15.39	44.02	14.85	45.57	14.64	45.51	13.51	45.48	13.54
Monitoring ErrP	63.67	61.47	65.05	60.29	67.09	61.17	69.02	61.08	57.01	60.53
EEGMAT	77.32	69.60	75.62	64.36	80.40	65.90	72.38	66.21	77.93	64.35
LEMON gender	78.11	64.62	78.71	64.31	78.55	64.12	76.54	64.56	77.99	63.93
LEMON extraversion	50.41	51.56	51.53	50.66	54.21	50.58	54.35	50.81	54.86	50.74
FACED	44.65	44.25	47.03	43.54	47.83	43.29	34.99	43.03	39.98	42.34
SEED	56.59	51.73	64.18	55.98	56.24	55.93	62.59	53.51	62.87	54.69
SEED-IV	30.64	30.81	32.17	31.90	28.83	31.78	31.47	31.93	32.64	31.86
SEED-V	28.18	27.17	28.15	27.17	27.33	28.10	27.66	27.37	28.72	27.14
SEED-VII	23.12	19.85	22.17	20.96	18.93	20.12	22.21	20.33	20.77	20.54
ISRUC S2	22.89	45.83	25.33	47.75	29.52	50.38	22.79	48.11	25.05	47.27
ISRUC S3	85.22	69.25	87.83	67.95	89.08	71.56	83.64	71.46	82.86	72.36
SleepEDF	75.63	68.57	75.73	68.50	75.04	68.93	75.76	69.21	76.65	69.32
HMC	69.64	54.54	71.09	53.84	71.40	53.98	71.43	53.21	70.65	52.67
TUAB	84.07	77.80	84.10	78.37	84.68	78.37	84.52	77.70	84.87	77.46
TUEV	74.07	60.00	77.87	58.33	79.34	58.29	79.01	55.18	82.41	58.07
TUSL	61.99	63.95	56.83	63.27	55.60	63.27	51.22	63.27	54.47	62.04
EEGDenoiseNet	99.91	98.85	99.89	98.26	99.96	98.59	99.89	98.33	99.74	98.67
Mean	57.98	53.98	57.86	53.67	59.05	54.09	56.79	53.64	57.07	53.53

Table 19. Atom-H: full-data checkpoints. Balanced accuracy (%).

Dataset	Atom-H ep9		Atom-H ep19		Atom-H ep29		Atom-H ep39	
	FT	LP	FT	LP	FT	LP	FT	LP
AD65	62.40	58.47	56.45	59.78	63.38	55.25	63.08	55.58
PD31	52.90	57.47	60.47	56.73	44.77	56.07	55.50	54.73
BCIC IV-1	51.83	50.83	50.33	51.33	51.33	54.00	47.50	53.83
BCIC IV-2a	39.58	40.91	38.72	41.09	35.36	42.01	37.27	41.78
SHU MI	53.24	59.23	52.05	58.40	51.10	58.22	51.21	60.09
BETA SSVEP	45.68	17.89	44.64	18.24	45.65	17.59	46.49	17.41
Monitoring ErrP	59.11	61.35	63.53	60.31	64.20	61.89	63.75	60.69

Continued on the next page.

Table 19. Atom-H: full-data checkpoints. Balanced accuracy (%) (continued).

Dataset	Atom-H ep9		Atom-H ep19		Atom-H ep29		Atom-H ep39	
	FT	LP	FT	LP	FT	LP	FT	LP
EEGMAT	71.76	65.43	79.63	66.06	72.69	65.28	77.79	64.51
LEMON gender	70.27	63.86	72.51	64.09	76.28	62.72	74.26	63.43
LEMON extraversion	49.70	51.67	51.12	50.44	52.71	50.53	53.45	50.26
FACED	47.19	43.10	51.53	44.26	31.24	45.03	29.53	44.29
SEED	63.40	53.91	66.33	54.88	63.55	54.12	65.36	53.14
SEED-IV	31.93	31.50	32.37	31.91	30.31	31.28	30.07	30.31
SEED-V	26.52	26.38	27.56	27.26	26.71	27.87	27.71	27.07
SEED-VII	20.20	20.51	18.82	19.93	19.34	20.15	20.46	19.61
ISRUC S2	25.11	45.06	27.77	45.24	25.12	41.76	27.40	43.71
ISRUC S3	84.41	73.27	84.35	72.80	78.83	74.12	84.38	74.09
SleepEDF	74.33	71.14	74.54	70.81	75.03	70.44	74.38	70.19
HMC	71.96	58.04	73.37	56.77	72.33	55.28	71.81	54.60
TUAB	84.14	79.23	84.69	78.59	84.74	78.23	84.63	78.20
TUEV	74.07	61.10	77.04	56.94	77.28	55.70	78.42	56.91
TUSL	74.26	62.59	66.17	63.34	59.37	60.89	55.53	60.65
EEGDenoiseNet	99.80	98.67	99.72	98.54	99.85	97.37	99.63	96.43
Mean	57.99	54.42	58.86	54.25	56.57	53.73	57.37	53.54

Table 20. Molecule-B: 50-epoch run. Balanced accuracy (%).

Dataset	Molecule-B ep9		Molecule-B ep19		Molecule-B ep29		Molecule-B ep39		Molecule-B ep49	
	FT	LP	FT	LP	FT	LP	FT	LP	FT	LP
AD65	51.14	54.89	55.85	52.43	60.97	52.50	60.87	50.75	57.62	49.21
PD31	46.18	54.89	57.81	50.06	56.68	51.30	52.17	51.49	52.55	50.14
BCIC IV-1	51.50	52.67	49.33	51.00	49.67	51.67	55.00	54.00	49.50	53.17
BCIC IV-2a	34.72	42.77	35.76	40.05	33.91	41.61	34.95	40.91	37.67	40.16
SHU MI	57.13	61.86	55.98	61.94	57.04	61.15	55.06	61.06	54.16	59.60
BETA SSVEP	39.73	14.35	39.82	15.12	41.46	14.20	38.36	13.57	39.17	13.33
Monitoring ErrP	64.12	61.61	58.02	61.39	64.36	61.94	62.84	62.50	56.86	62.58
EEGMAT	77.16	68.06	73.15	66.36	79.78	65.90	77.47	60.80	78.24	61.88
LEMON gender	71.97	61.63	72.93	61.80	73.00	60.86	75.34	61.55	74.96	60.64
LEMON extraversion	47.04	51.20	53.93	49.71	51.42	50.67	57.82	51.12	52.59	50.72
FACED	40.69	46.65	42.43	45.93	44.33	44.98	39.18	44.66	36.05	44.67
SEED	59.59	54.36	60.26	53.64	58.75	54.92	60.14	52.79	56.41	53.18
SEED-IV	32.47	30.98	33.19	31.34	31.90	31.06	31.05	31.41	29.76	30.98
SEED-V	28.36	26.71	28.86	26.18	26.12	25.23	26.87	24.80	27.10	24.95
SEED-VII	21.49	20.51	21.62	20.29	21.39	19.95	19.44	19.91	19.15	19.03
ISRUC S2	27.06	44.48	23.89	44.68	27.86	43.78	28.29	41.47	26.04	41.99
ISRUC S3	81.23	69.74	80.99	68.93	79.72	67.15	81.03	68.35	81.26	68.54
SleepEDF	74.61	67.34	74.66	67.46	75.10	66.79	74.31	66.32	74.63	65.95
HMC	66.32	49.47	66.70	50.21	66.82	48.50	66.84	47.80	66.70	47.39
TUAB	82.21	76.38	82.23	75.78	82.98	75.67	82.90	75.53	81.90	75.34
TUEV	66.96	53.94	66.02	50.47	66.99	50.88	69.45	48.82	66.90	54.05
TUSL	55.78	61.88	49.80	65.15	49.01	64.77	56.50	66.61	55.36	70.63
EEGDenoiseNet	99.08	92.27	99.52	92.73	99.33	92.88	98.97	92.55	98.24	91.24
Mean	55.50	52.98	55.77	52.29	56.46	52.10	56.73	51.69	55.34	51.71

Table 21. Molecule-L and Molecule-H checkpoints. Balanced accuracy (%).

Dataset	Molecule-L ep9		Molecule-L ep19		Molecule-L ep29		Molecule-H ep9	
	FT	LP	FT	LP	FT	LP	FT	LP
AD65	62.23	53.48	64.21	52.20	59.90	49.85	61.24	55.45
PD31	45.50	55.65	54.04	53.54	55.55	54.70	58.12	53.88
BCIC IV-1	51.67	50.67	50.50	54.00	51.33	53.50	48.83	53.67
BCIC IV-2a	40.91	42.94	46.82	42.59	34.14	42.59	35.88	43.06
SHU MI	60.11	59.08	54.30	60.93	54.83	60.91	59.28	60.72
BETA SSVEP	45.68	16.67	45.36	15.54	43.72	17.23	49.91	19.88
Monitoring ErrP	72.67	60.05	63.39	61.57	58.03	61.77	63.14	63.03

Continued on the next page.

Table 21. Molecule-L and Molecule-H checkpoints. Balanced accuracy (%) (continued).

Dataset	Molecule-L ep9		Molecule-L ep19		Molecule-L ep29		Molecule-H ep9	
	FT	LP	FT	LP	FT	LP	FT	LP
EEGMAT	74.07	68.21	77.62	68.98	75.62	65.74	74.54	67.29
LEMON gender	74.07	62.59	76.38	62.43	73.89	62.81	74.73	63.48
LEMON extraversion	48.50	49.70	50.28	49.65	51.03	50.82	52.17	49.75
FACED	50.40	46.11	46.59	44.52	43.91	46.19	37.41	46.43
SEED	65.47	51.48	67.85	51.37	57.92	51.25	61.33	53.78
SEED-IV	34.23	32.03	33.88	31.09	31.25	29.99	35.47	30.85
SEED-V	28.17	27.46	28.19	26.00	27.48	25.37	26.66	26.01
SEED-VII	21.02	20.61	21.12	20.55	20.81	19.57	22.74	20.34
ISRUC S2	25.42	46.47	24.30	47.15	28.71	46.30	26.01	43.37
ISRUC S3	80.81	70.60	82.06	72.07	80.52	70.08	83.88	72.42
SleepEDF	75.69	68.68	74.68	68.98	74.77	67.31	75.72	70.14
HMC	70.57	51.46	70.35	51.42	70.55	48.70	71.34	54.42
TUAB	84.26	76.88	83.19	77.08	82.65	76.33	84.17	78.15
TUEV	77.43	57.18	78.11	59.24	75.75	55.17	75.75	55.95
TUSL	61.54	62.96	57.81	64.51	52.18	74.14	56.56	64.35
EEGDenoiseNet	99.83	96.23	99.74	96.41	99.46	95.51	99.52	92.91
Mean	58.71	53.35	58.73	53.56	56.70	53.30	58.02	53.88

Table 22. Atom-B: pretraining data fractions. Balanced accuracy (%).

Dataset	10% data ep49		20% data ep49		50% data ep49		100% data ep49	
	FT	LP	FT	LP	FT	LP	FT	LP
AD65	61.07	53.73	54.39	54.88	65.82	59.06	55.52	52.63
PD31	45.45	47.43	44.35	47.43	44.83	51.32	53.74	49.39
BCIC IV-1	50.50	54.67	50.00	54.33	54.50	51.83	51.83	50.50
BCIC IV-2a	37.21	42.13	33.10	41.78	36.34	42.42	36.69	41.90
SHU MI	53.56	61.17	56.16	59.38	52.88	58.47	56.24	58.41
BETA SSVEP	33.42	10.36	29.70	12.20	37.65	14.52	31.76	11.64
Monitoring ErrP	61.61	59.76	62.28	58.74	56.74	60.80	56.05	57.69
EEGMAT	68.69	68.67	70.69	66.98	76.54	68.22	72.87	66.21
LEMON gender	72.61	66.35	65.60	60.82	75.71	63.33	72.37	62.15
LEMON extraversion	51.34	51.80	52.56	50.80	51.52	50.38	54.00	50.12
FACED	38.07	45.05	35.66	43.47	31.85	43.24	25.02	43.70
SEED	60.39	50.60	60.86	51.71	62.97	54.06	59.41	56.75
SEED-IV	31.40	30.79	29.96	29.68	32.58	30.77	30.57	32.41
SEED-V	27.72	25.38	26.80	25.84	28.78	26.86	26.79	26.34
SEED-VII	20.46	19.31	19.79	20.58	20.45	21.03	19.35	20.43
ISRUC S2	23.05	45.63	22.41	41.36	22.08	40.36	22.87	41.60
ISRUC S3	76.03	70.15	76.51	65.18	83.76	74.57	83.36	73.07
SleepEDF	74.97	65.84	74.36	66.74	75.04	68.71	74.34	67.10
HMC	68.77	49.87	66.77	50.37	66.83	49.82	68.13	49.04
TUAB	82.79	76.16	82.48	77.10	82.17	76.70	82.90	76.03
TUEV	67.67	51.69	66.68	57.53	69.61	55.62	70.80	55.39
TUSL	49.20	68.80	48.43	62.26	60.03	71.00	54.72	70.96
EEGDenoiseNet	98.37	94.28	98.76	97.02	99.33	98.02	99.72	98.93
Mean	54.54	52.59	53.40	52.01	56.00	53.53	54.74	52.71

The 100% reference is repeated from the full-data checkpoint tables.

Table 23. Atom-L: pretraining data fractions. Balanced accuracy (%).

Dataset	10% data ep49		20% data ep49		50% data ep49		100% data ep49	
	FT	LP	FT	LP	FT	LP	FT	LP
AD65	57.02	53.94	51.77	54.57	64.14	52.26	59.86	57.16
PD31	45.61	49.27	45.10	51.89	50.22	50.95	37.12	53.52
BCIC IV-1	48.83	53.00	50.17	52.33	52.00	53.00	53.83	52.33
BCIC IV-2a	40.97	43.06	33.39	41.61	35.53	41.49	34.32	41.61
SHU MI	53.34	58.10	51.20	58.94	52.67	59.44	52.60	58.97

Continued on the next page.

Table 23. Atom-L: pretraining data fractions. Balanced accuracy (%) (continued).

Dataset	10% data ep49		20% data ep49		50% data ep49		100% data ep49	
	FT	LP	FT	LP	FT	LP	FT	LP
BETA SSVEP	41.85	17.86	40.39	16.46	46.58	14.05	45.48	13.54
Monitoring ErrP	56.95	58.98	57.66	58.69	61.34	59.99	57.01	60.53
EEGMAT	66.83	64.66	69.45	69.61	68.99	63.74	77.93	64.35
LEMON gender	76.36	64.50	74.69	63.45	77.42	62.72	77.99	63.93
LEMON extraversion	52.30	50.78	53.77	51.01	53.43	50.80	54.86	50.74
FACED	33.37	43.87	31.17	42.74	36.89	42.87	39.98	42.34
SEED	60.03	53.37	61.16	52.91	63.92	51.37	62.87	54.69
SEED-IV	31.09	30.66	29.96	30.82	30.75	30.85	32.64	31.86
SEED-V	27.12	26.19	28.05	26.42	26.95	25.90	28.72	27.14
SEED-VII	21.47	21.03	20.79	19.91	21.67	20.23	20.77	20.54
ISRUC S2	25.13	46.82	26.89	46.00	28.24	46.00	25.05	47.27
ISRUC S3	79.96	70.12	82.67	71.83	86.41	72.05	82.86	72.36
SleepEDF	75.28	69.46	74.53	69.11	75.97	69.38	76.65	69.32
HMC	69.27	52.58	70.77	52.80	69.55	49.89	70.65	52.67
TUAB	83.62	77.72	83.57	77.74	83.79	77.40	84.87	77.46
TUEV	64.49	58.82	70.17	51.76	71.93	53.43	82.41	58.07
TUSL	53.91	67.56	63.75	61.83	56.68	65.54	54.47	62.04
EEGDenoiseNet	99.69	97.60	99.87	98.54	99.50	97.73	99.74	98.67
Mean	54.98	53.48	55.26	53.09	57.15	52.66	57.07	53.53

The 100% reference is repeated from the full-data checkpoint tables.

Table 24. Molecule-B: pretraining data fractions. Balanced accuracy (%).

Dataset	10% data ep49		20% data ep49		50% data ep49		100% data ep49	
	FT	LP	FT	LP	FT	LP	FT	LP
AD65	53.32	57.48	54.38	54.06	60.13	53.60	57.62	49.21
PD31	52.93	48.74	59.64	57.88	52.08	55.15	52.55	50.14
BCIC IV-1	54.00	52.67	50.67	50.50	52.67	52.33	49.50	53.17
BCIC IV-2a	36.34	44.27	34.72	43.06	41.55	40.39	37.67	40.16
SHU MI	53.37	61.28	58.07	61.31	54.97	60.55	54.16	59.60
BETA SSVEP	38.04	10.27	40.27	12.29	36.82	7.20	39.17	13.33
Monitoring ErrP	63.82	60.78	57.08	62.34	57.82	59.13	56.86	62.58
EEGMAT	75.62	62.81	77.31	70.68	73.46	70.84	78.24	61.88
LEMON gender	75.82	62.79	72.48	66.19	75.30	62.50	74.96	60.64
LEMON extraversion	47.88	51.53	51.51	49.88	52.94	51.85	52.59	50.72
FACED	41.30	46.91	39.04	45.81	42.82	45.88	36.05	44.67
SEED	63.23	52.91	60.81	51.89	63.05	52.71	56.41	53.18
SEED-IV	33.62	31.55	33.62	30.75	32.60	30.30	29.76	30.98
SEED-V	28.77	26.45	30.78	27.50	30.26	26.43	27.10	24.95
SEED-VII	19.00	19.74	19.93	19.33	20.08	20.20	19.15	19.03
ISRUC S2	21.95	45.85	25.23	45.70	25.42	42.77	26.04	41.99
ISRUC S3	79.92	70.65	79.06	68.54	80.49	70.05	81.26	68.54
SleepEDF	74.20	66.00	74.82	67.23	73.76	67.30	74.63	65.95
HMC	68.39	50.61	67.74	49.33	69.03	51.20	66.70	47.39
TUAB	82.23	76.48	82.21	75.64	82.13	75.29	81.90	75.34
TUEV	67.25	59.24	63.03	57.02	61.52	55.68	66.90	54.05
TUSL	49.06	63.75	64.03	66.32	53.24	58.38	55.36	70.63
EEGDenoiseNet	99.00	96.06	99.57	87.90	99.48	98.32	98.24	91.24
Mean	55.61	52.99	56.35	53.09	56.16	52.52	55.34	51.71

The 100% reference is repeated from the full-data checkpoint tables. Molecule-B uses the independent 50-epoch run.

Table 25. External encoders (1/4). Balanced accuracy (%). [3, 4]

Dataset	REVE Base		REVE Large		LUNA Base		LUNA Large	
	FT	LP	FT	LP	FT	LP	FT	LP
AD65	54.71	48.30	58.29	53.68	52.42	44.51	48.72	44.68

Continued on the next page.

Table 25. External encoders (1/4). Balanced accuracy (%). [3, 4] (continued).

Dataset	REVE Base		REVE Large		LUNA Base		LUNA Large	
	FT	LP	FT	LP	FT	LP	FT	LP
PD31	40.54	49.78	50.56	52.25	40.76	40.43	53.66	41.63
BCIC IV-1	54.17	52.00	53.67	53.17	52.00	51.17	51.67	52.67
BCIC IV-2a	41.84	38.02	44.33	44.73	38.25	35.82	37.21	36.57
SHU MI	62.35	59.11	56.50	58.93	58.93	58.27	62.70	55.80
BETA SSVEP	38.39	14.61	35.30	18.33	27.14	4.43	28.24	4.05
Monitoring ErrP	62.48	62.29	56.75	65.35	58.95	57.19	56.81	59.05
EEGMAT	69.92	60.04	74.85	64.67	63.16	59.47	56.92	58.83
LEMON gender	70.75	67.51	73.53	73.30	58.75	53.74	64.59	57.41
LEMON extraversion	50.87	51.44	51.05	52.41	49.54	51.25	49.01	49.76
FACED	51.63	42.37	47.33	44.17	53.24	39.90	57.14	42.70
SEED	60.88	55.63	60.50	50.88	62.12	50.08	57.37	50.54
SEED-IV	33.03	31.69	30.59	31.23	30.10	28.05	33.36	30.61
SEED-V	24.63	24.90	24.63	24.78	28.15	23.84	28.30	23.63
SEED-VII	22.02	19.12	21.85	20.13	21.22	19.54	19.92	20.21
ISRUC S2	24.08	37.61	23.73	39.40	23.81	35.32	24.60	37.39
ISRUC S3	84.48	77.50	83.72	78.26	78.70	60.81	78.88	61.02
SleepEDF	74.13	68.51	72.37	71.66	73.92	63.27	73.96	66.66
HMC	56.50	43.47	58.62	43.24	63.82	44.50	64.77	49.21
TUAB	79.02	76.72	80.20	77.36	79.41	73.40	80.14	74.80
TUEV	61.73	51.91	55.17	51.88	65.16	51.99	68.68	53.75
TUSL	70.47	62.73	81.02	70.15	57.75	59.96	59.54	71.34
EEGDenoiseNet	83.67	69.51	80.58	61.93	99.21	95.58	99.44	97.43
Mean	55.32	50.64	55.44	52.26	53.76	47.94	54.59	49.55

Table 26. External encoders (2/4). Balanced accuracy (%). [1, 9, 10]

Dataset	LUNA Huge		CSBrain		Official LaBraM Base		CodeBrain	
	FT	LP	FT	LP	FT	LP	FT	LP
AD65	51.81	44.40	56.87	51.62	55.53	57.29	56.72	60.25
PD31	54.95	48.42	53.58	41.26	55.64	46.23	51.58	49.95
BCIC IV-1	48.67	51.83	53.17	52.00	50.67	49.33	49.17	49.50
BCIC IV-2a	41.38	35.42	36.17	38.37	37.27	39.70	37.38	42.19
SHU MI	61.12	58.42	56.67	55.59	51.64	58.46	50.94	60.57
BETA SSVEP	18.93	5.95	31.64	19.49	31.73	14.11	12.86	4.64
Monitoring ErrP	58.64	61.37	58.98	56.89	56.39	60.76	59.86	54.63
EEGMAT	66.36	60.81	64.06	56.51	73.92	67.29	67.14	65.59
LEMON gender	59.62	59.19	71.77	65.13	75.79	73.97	64.44	69.83
LEMON extraversion	51.60	50.24	49.49	50.38	48.83	51.79	50.04	50.41
FACED	57.34	44.01	51.27	34.83	21.49	41.59	46.64	46.01
SEED	56.46	53.91	58.41	51.95	61.53	54.29	63.28	59.96
SEED-IV	31.67	31.00	30.28	28.32	33.19	31.35	30.05	31.00
SEED-V	26.99	27.24	24.75	21.38	26.12	26.12	29.31	30.78
SEED-VII	19.58	19.90	20.31	18.44	19.13	20.03	21.70	21.73
ISRUC S2	23.22	43.51	25.91	29.46	21.86	36.31	27.88	39.03
ISRUC S3	81.79	69.98	50.73	52.28	74.80	63.80	64.19	58.90
SleepEDF	74.66	69.57	73.35	64.15	74.83	68.06	74.56	68.92
HMC	66.69	54.66	59.95	44.53	67.24	49.68	58.11	51.04
TUAB	81.07	76.80	80.84	78.23	82.75	76.44	80.01	76.35
TUEV	63.92	61.29	67.83	56.40	59.71	59.62	66.99	58.31
TUSL	66.73	60.11	75.21	73.66	65.99	63.07	61.27	72.01
EEGDenoiseNet	99.76	99.22	99.94	95.50	81.48	85.68	99.80	99.28
Mean	54.91	51.62	54.40	49.41	53.37	51.95	53.21	53.08

Table 27. External encoders (3/4). Balanced accuracy (%). [5, 7, 8, 28]

Dataset	NeuroLM		BrainOmni		NeuroGPT		EEGMamba	
	FT	LP	FT	LP	FT	LP	FT	LP
AD65	57.27	50.44	59.06	53.96	55.73	45.58	52.42	51.29

Continued on the next page.

Table 27. External encoders (3/4). Balanced accuracy (%). [5, 7, 8, 28] (continued).

Dataset	NeuroLM		BrainOmni		NeuroGPT		EEGMamba	
	FT	LP	FT	LP	FT	LP	FT	LP
PD31	58.74	61.57	60.07	48.76	56.44	34.44	49.21	50.92
BCIC IV-1	51.17	50.67	49.33	54.33	52.00	53.67	51.50	49.83
BCIC IV-2a	35.65	40.74	45.43	38.95	36.11	39.18	32.81	37.21
SHU MI	57.13	58.75	57.43	58.61	54.59	51.90	50.55	58.70
BETA SSVEP	12.05	6.34	9.29	17.17	17.23	2.77	3.45	6.55
Monitoring ErrP	53.55	55.40	56.22	58.47	56.30	49.68	52.47	58.19
EEGMAT	68.21	66.67	50.00	66.68	62.81	49.69	50.31	55.41
LEMON gender	67.72	65.59	64.76	63.20	71.58	57.43	64.89	57.79
LEMON extraversion	51.60	52.35	52.45	52.38	50.81	50.43	50.89	50.00
FACED	38.22	43.79	26.02	41.53	43.30	21.62	30.31	39.69
SEED	52.54	50.39	52.57	50.83	51.73	47.89	59.58	51.62
SEED-IV	32.10	32.41	31.36	30.00	26.11	27.03	30.65	28.96
SEED-V	24.91	26.04	25.28	24.31	28.54	23.83	27.55	23.72
SEED-VII	20.68	19.69	21.69	20.95	17.58	18.12	20.21	18.35
ISRUC S2	29.54	37.81	28.31	41.27	30.02	39.86	27.99	28.59
ISRUC S3	63.43	58.83	77.54	67.47	70.97	57.26	65.57	36.70
SleepEDF	70.11	64.53	70.69	63.00	70.25	62.90	71.83	56.72
HMC	53.21	50.65	63.57	53.14	60.69	37.79	49.33	41.61
TUAB	79.78	77.66	80.57	78.42	78.67	71.78	78.68	72.09
TUEV	67.49	60.29	67.41	54.41	63.31	43.44	62.58	53.56
TUSL	72.92	73.05	60.07	57.39	48.11	59.66	62.79	69.08
EEGDenoiseNet	86.22	84.66	94.66	91.01	99.98	99.37	99.89	97.32
Mean	52.36	51.67	52.34	51.58	52.30	45.45	49.80	47.56

Table 28. External encoders (4/4). Balanced accuracy (%). [2, 6, 26, 27]

Dataset	CBraMod		BENDR		BIOT		FEMBA Base	
	FT	LP	FT	LP	FT	LP	FT	LP
AD65	53.71	55.85	46.13	33.33	58.31	47.26	50.21	36.44
PD31	39.91	44.44	66.67	33.33	69.07	61.92	54.02	47.86
BCIC IV-1	55.33	49.50	51.17	52.00	49.83	52.33	50.50	51.00
BCIC IV-2a	30.15	40.16	29.86	27.14	40.45	33.04	41.09	37.09
SHU MI	50.81	61.85	52.81	50.46	53.91	53.53	58.98	60.02
BETA SSVEP	26.61	16.34	3.30	2.59	18.36	5.65	24.17	11.31
Monitoring ErrP	52.55	56.96	54.80	50.72	52.88	52.29	55.65	54.51
EEGMAT	62.09	62.35	64.52	50.00	66.67	62.22	70.70	50.00
LEMON gender	73.58	69.84	57.53	50.00	61.52	59.77	60.46	52.25
LEMON extraversion	49.89	50.79	49.75	50.00	48.77	49.72	51.05	50.13
FACED	34.82	50.07	16.36	16.72	22.56	19.94	27.97	40.34
SEED	57.32	54.02	37.15	34.62	60.15	57.41	52.53	45.58
SEED-IV	25.66	32.49	26.09	25.24	28.01	29.25	29.12	25.89
SEED-V	26.53	23.92	20.04	20.16	24.68	24.91	26.21	28.88
SEED-VII	19.15	19.83	16.01	14.93	20.51	19.26	19.00	19.73
ISRUC S2	25.62	31.87	32.04	23.70	32.95	47.40	25.22	31.63
ISRUC S3	57.74	58.81	57.11	21.69	76.28	71.88	81.76	68.54
SleepEDF	71.68	64.37	73.59	22.67	72.97	66.69	74.17	50.75
HMC	66.87	57.12	63.19	24.86	63.20	55.54	68.99	52.46
TUAB	79.35	79.21	81.32	56.45	90.44	83.34	77.43	71.06
TUEV	57.70	54.93	47.92	16.69	81.41	74.05	67.00	58.68
TUSL	30.56	59.97	37.48	36.63	56.05	69.21	42.97	48.09
EEGDenoiseNet	49.97	91.83	93.18	55.05	99.91	98.74	99.59	75.95
Mean	47.72	51.59	46.87	33.43	54.30	51.97	52.56	46.44

EEGDenoiseNet uses single-channel Cz input for BIOT and a Cz-only, zero-filled montage for FEMBA (Appendix C). BENDR FT uses the stable grid except on BCIC IV-2a.

Table 29. Models trained from random initialization. Balanced accuracy (%). [29, 30]

Dataset	EEGNet		EEGConformer		LaBraM Base	
	FT	LP	FT	LP	FT	LP
AD65	52.70	—	58.60	—	47.71	—
PD31	64.43	—	52.05	—	40.57	—
BCIC IV-1	51.83	—	53.83	—	54.17	—
BCIC IV-2a	36.17	—	50.29	—	33.80	—
SHU MI	57.81	—	59.37	—	51.62	—
BETA SSVEP	5.60	—	12.59	—	5.09	—
Monitoring ErrP	52.48	—	63.42	—	57.45	—
EEGMAT	59.15	—	71.91	—	64.54	—
LEMON gender	70.08	—	71.05	—	57.06	—
LEMON extraversion	50.42	—	50.18	—	50.18	—
FACED	20.24	—	26.49	—	14.79	—
SEED	56.43	—	54.01	—	48.26	—
SEED-IV	29.60	—	31.95	—	26.51	—
SEED-V	31.52	—	23.33	—	20.34	—
SEED-VII	20.72	—	20.44	—	16.65	—
ISRUC S2	25.39	—	32.02	—	24.79	—
ISRUC S3	73.02	—	66.68	—	62.85	—
SleepEDF	70.02	—	71.36	—	67.41	—
HMC	65.77	—	68.70	—	48.42	—
TUAB	79.21	—	79.06	—	75.69	—
TUEV	40.73	—	46.63	—	57.45	—
TUSL	40.49	—	34.42	—	32.84	—
EEGDenoiseNet	83.28	—	99.08	—	69.96	—
Mean	49.44	—	52.06	—	44.70	—

LP was not evaluated for these randomly initialized models.

G Pretraining Trajectories

The training records cover 22 runs and a separate two-epoch continuation of Molecule-L, totaling 941 recorded epochs. Figures 12–14 show shape-preserving cubic curves through the recorded losses, with interpolation in log-loss coordinates on logarithmic axes. Fractions refer to the retained window streams; all reduced-data runs have 50-epoch schedules. Some trajectories extend beyond the checkpoints evaluated downstream.

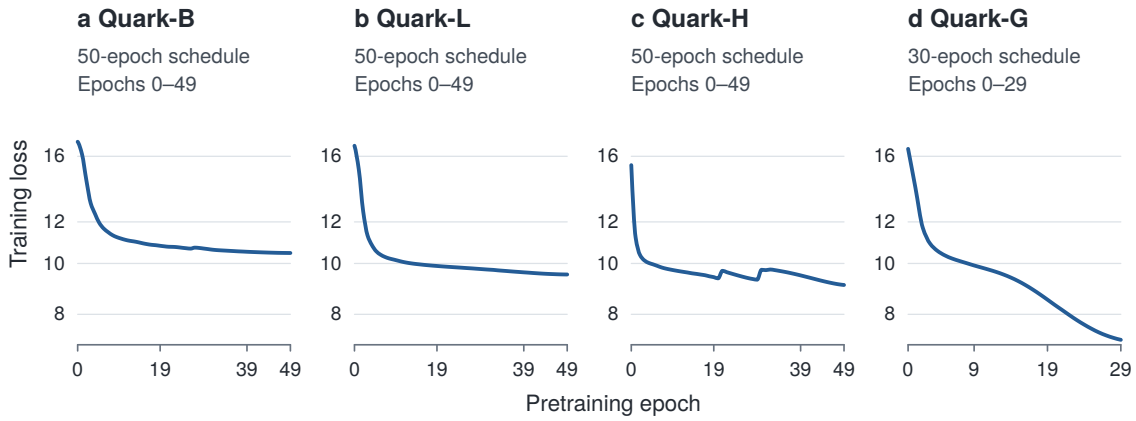


Figure 12. Quark pretraining loss. Base, Large, and Huge use 50-epoch schedules; Giant uses 30 epochs. Each panel shows its full recorded range. Loss axes are logarithmic.

G.1 Atom

Atom's full-data and reduced-data runs use 50-epoch schedules. Figure 13 compares the full-data trajectories with the 10%, 20%, and 50% data fractions at each encoder size.

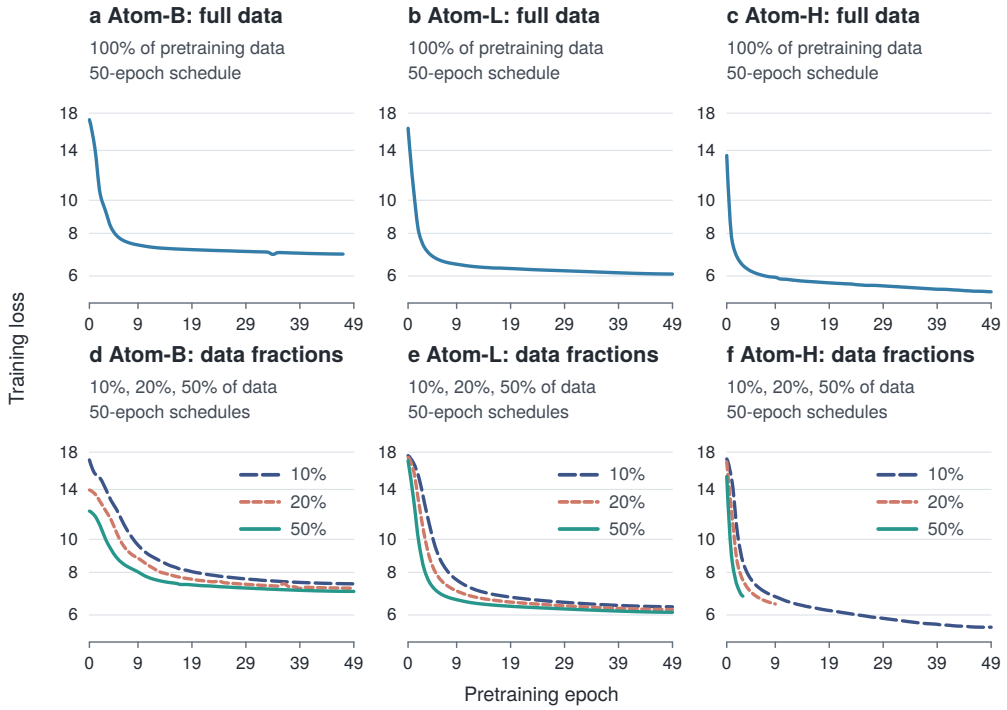


Figure 13. Atom pretraining loss. Full-data runs appear separately from reduced-data runs. Loss axes are logarithmic.

G.2 Molecule

Molecule-B uses a 50-epoch schedule. The Large continuation resumes ep29 with a separate two-epoch rewarm schedule; it has no downstream evaluation. Huge has 20 recorded epochs under a 50-epoch schedule, with its last evaluated checkpoint at ep9. Tokenizers and input preparation differ across collections, limiting direct comparison of their absolute losses.

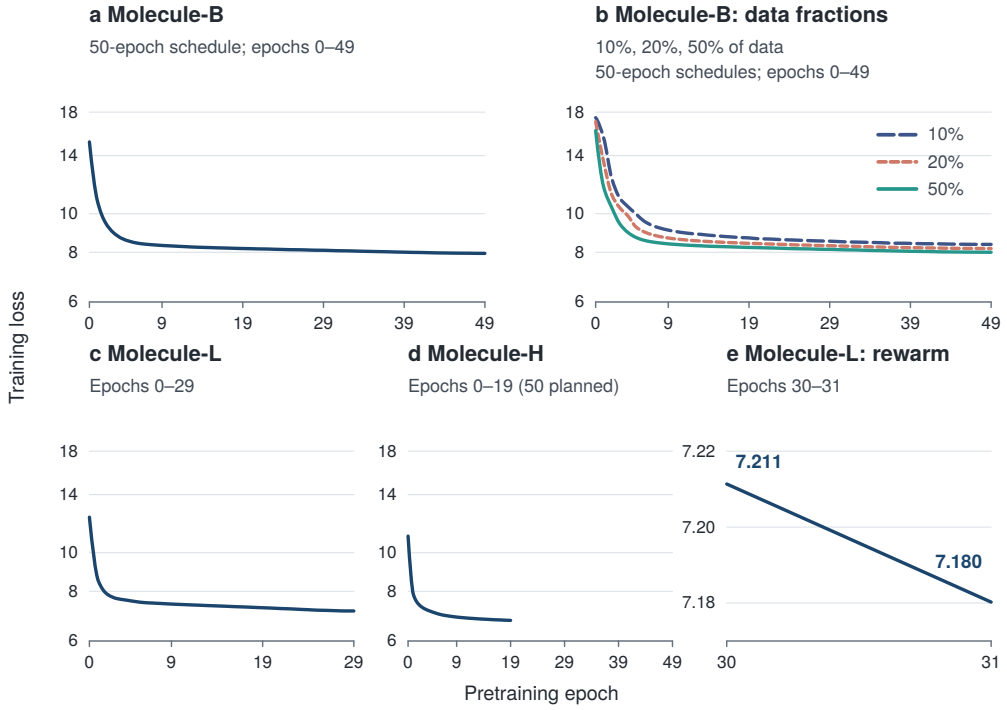


Figure 14. Molecule pretraining loss for Base, reduced-data Base runs, Large, Huge, and the Large continuation. Loss axes are logarithmic except for the continuation.

H Joint EEG–LLM Results

These experiments follow the joint-training protocol in Section 7. Ten EEG encoders are evaluated with six Qwen3 sizes, giving 60 configurations and 180 runs. Tables report test balanced accuracy (%) as mean ± sample SD over seeds 42, 10, and 5. The macro row averages the 23 tasks within each seed before aggregation. †ISRUC S2 retains the original channel-cache defect and remains in the macro mean.

Table 30. Joint EEG–LLM training and evaluation settings.

Setting	Value
EEG encoder	Frozen; the checkpoint is identified in each result-table caption.
Input preparation	Quark/Atom: input scale 0.01, no per-window z-scoring. Molecule: input scale 1, per-window z-scoring.
Cached features	Temporal patch features averaged across channels, standardized using the corresponding task and seed's training-split statistics.
Temporal projector	32 tokens; adaptive average pooling or linear interpolation for shorter sequences; two-layer GELU MLP.
Language model	Qwen3-0.6B, 1.7B, 4B, 8B, 14B, or 32B; original weights frozen.
Trainable modules	Projector and LoRA adapters, trained jointly from step 0.
LoRA	Rank 16, $\alpha = 32$, dropout 0.1; q/k/v/o and gate/up/down projections.
Optimizer	AdamW; $\beta_1 = 0.9$, $\beta_2 = 0.98$.
Learning rate / weight decay	Projector: 10^{-3} / 0; LoRA: 10^{-4} / 0.01.
Schedule	500-step linear warmup, followed by cosine decay to $0.1\times$ the base learning rate.
Global batch / gradient clipping	64 / 1.0.
Training length / early stopping	At most 16,000 steps; validation every 1,000 steps; stop after four evaluations without improvement.
Task sampling	Probability proportional to retained training-set size raised to 0.3; shuffled sampling within each task.
Training supervision	Answer-token loss with inverse training-class-frequency weights for all tasks; prompt template and candidate order vary during training.
Validation selection	One shared checkpoint per seed, maximizing mean task balanced accuracy on up to 1,000 validation examples per task; the earlier checkpoint is retained on ties.
Final evaluation	All test examples; closed-set textual-candidate scoring. The macro score assigns equal weight to all 23 tasks.
Seeds	42, 10, and 5.

The experiments share the data splits and 40-window subject–class training cap used for single-task evaluation (Appendix C). SleepEDF uses night identifiers as split units, allowing person-level overlap; EEGDenoiseNet uses a within-subject sample split. Validation and test use all available windows.

Table 31. EEG–LLM results for Quark-B (epoch 49).

Task	Qwen3-0.6B	Qwen3-1.7B	Qwen3-4B	Qwen3-8B	Qwen3-14B	Qwen3-32B
AD65	49.05 ± 5.89	49.80 ± 7.63	51.20 ± 12.17	47.23 ± 6.03	48.56 ± 5.33	50.21 ± 6.74
PD31	47.41 ± 11.17	39.68 ± 15.09	42.78 ± 13.03	43.71 ± 14.85	37.45 ± 15.96	43.03 ± 15.88
BCIC IV-1	49.33 ± 1.44	48.83 ± 4.04	48.33 ± 1.76	48.00 ± 3.28	50.00 ± 4.50	51.50 ± 2.29
BCIC IV-2a	34.43 ± 0.36	34.49 ± 0.82	34.90 ± 0.52	35.53 ± 0.96	36.05 ± 1.06	35.24 ± 2.39
SHU MI	50.67 ± 1.01	51.27 ± 1.28	51.26 ± 1.66	52.43 ± 0.70	50.58 ± 0.98	50.92 ± 1.06
BETA SSVEP	2.41 ± 0.00	2.32 ± 0.36	2.50 ± 0.47	2.44 ± 0.31	2.35 ± 0.44	2.62 ± 0.46
Monitoring ErrP	52.45 ± 3.31	54.84 ± 2.36	52.08 ± 2.95	54.89 ± 1.95	56.07 ± 3.49	52.93 ± 3.49
EEGMAT	66.07 ± 4.19	65.46 ± 6.58	63.75 ± 11.63	63.90 ± 5.70	61.62 ± 5.59	66.54 ± 9.28
LEMON gender	65.84 ± 3.31	65.99 ± 1.78	65.25 ± 3.16	63.83 ± 2.39	63.23 ± 5.57	63.38 ± 3.93
LEMON extraversion	51.04 ± 1.60	53.23 ± 4.20	51.95 ± 0.43	53.09 ± 1.50	52.02 ± 1.26	51.61 ± 2.44
FACED	14.96 ± 0.83	14.74 ± 1.20	14.89 ± 0.75	14.78 ± 1.16	12.96 ± 1.12	14.16 ± 0.66
SEED	48.18 ± 1.20	48.07 ± 3.57	47.89 ± 2.01	48.93 ± 2.05	47.52 ± 1.91	47.11 ± 3.20
SEED-IV	26.35 ± 1.38	27.29 ± 0.82	27.06 ± 0.34	26.16 ± 1.07	27.37 ± 2.55	26.50 ± 2.16

Continued on the next page.

Table 31. EEG–LLM results for Quark-B (epoch 49) (continued).

Task	Qwen3-0.6B	Qwen3-1.7B	Qwen3-4B	Qwen3-8B	Qwen3-14B	Qwen3-32B
SEED-V	23.12 ± 1.11	23.19 ± 0.54	22.66 ± 1.07	22.68 ± 0.27	22.45 ± 0.52	22.26 ± 0.24
SEED-VII	17.69 ± 0.96	16.54 ± 0.65	17.78 ± 0.75	17.50 ± 0.94	16.14 ± 0.19	16.56 ± 0.38
ISRUC S2†	24.35 ± 5.80	24.87 ± 3.79	24.21 ± 6.80	28.05 ± 4.99	28.93 ± 0.92	23.76 ± 6.29
ISRUC S3	71.10 ± 1.45	67.71 ± 1.99	71.18 ± 5.75	70.71 ± 1.81	73.30 ± 4.41	72.08 ± 1.76
SleepEDF	69.21 ± 2.77	71.52 ± 0.11	70.97 ± 0.90	71.08 ± 0.77	69.60 ± 1.80	69.98 ± 1.06
HMC	54.38 ± 0.22	55.23 ± 2.05	54.54 ± 1.31	54.88 ± 2.37	53.52 ± 2.69	54.03 ± 1.71
TUAB	75.81 ± 1.30	75.28 ± 1.07	75.50 ± 1.13	75.37 ± 0.91	74.14 ± 2.08	75.07 ± 1.79
TUEV	51.10 ± 7.48	50.25 ± 3.11	52.17 ± 7.25	48.77 ± 7.58	50.83 ± 6.96	46.92 ± 4.82
TUSL	54.85 ± 12.55	53.52 ± 14.17	52.81 ± 34.77	60.35 ± 9.38	52.22 ± 24.51	48.98 ± 9.07
EEGDenoiseNet	96.62 ± 0.45	95.80 ± 0.64	95.36 ± 0.22	95.02 ± 0.51	95.65 ± 1.08	95.62 ± 0.15
Macro mean	47.67 ± 1.22	47.39 ± 1.16	47.44 ± 3.06	47.80 ± 1.59	47.07 ± 1.99	47.00 ± 1.31

Table 32. EEG–LLM results for Quark-L (epoch 49).

Task	Qwen3-0.6B	Qwen3-1.7B	Qwen3-4B	Qwen3-8B	Qwen3-14B	Qwen3-32B
AD65	53.42 ± 5.40	53.84 ± 7.42	51.92 ± 1.49	49.66 ± 2.50	53.99 ± 6.61	55.75 ± 9.04
PD31	52.54 ± 15.45	41.52 ± 5.86	43.11 ± 7.80	47.44 ± 8.81	46.80 ± 9.87	42.68 ± 10.46
BCIC IV-1	50.67 ± 1.26	49.83 ± 4.16	48.50 ± 3.28	46.50 ± 1.32	48.33 ± 1.26	50.33 ± 3.06
BCIC IV-2a	35.24 ± 3.22	38.43 ± 2.49	36.11 ± 4.21	35.13 ± 3.53	31.08 ± 2.52	32.70 ± 2.40
SHU MI	50.91 ± 1.81	52.82 ± 1.23	51.61 ± 0.99	52.24 ± 1.87	49.59 ± 0.74	51.18 ± 0.92
BETA SSVEP	2.41 ± 0.31	2.80 ± 0.22	2.38 ± 0.14	2.95 ± 0.82	2.65 ± 0.14	2.47 ± 0.19
Monitoring ErrP	56.24 ± 3.26	55.35 ± 2.06	54.05 ± 4.38	55.04 ± 4.44	50.92 ± 3.14	55.43 ± 5.20
EEGMAT	65.59 ± 7.47	67.91 ± 8.39	69.46 ± 3.01	64.05 ± 3.73	64.98 ± 6.37	71.77 ± 5.78
LEMON gender	66.17 ± 2.78	65.74 ± 3.62	66.27 ± 5.17	63.66 ± 3.01	63.13 ± 6.33	64.49 ± 3.36
LEMON extraversion	50.33 ± 1.36	52.14 ± 1.95	52.70 ± 2.90	51.47 ± 1.68	49.82 ± 2.48	51.31 ± 2.44
FACED	14.60 ± 0.76	17.57 ± 1.56	13.51 ± 0.26	15.86 ± 2.70	13.47 ± 0.36	15.17 ± 0.60
SEED	51.27 ± 0.14	50.24 ± 2.56	48.91 ± 4.40	50.69 ± 0.37	51.22 ± 2.91	49.86 ± 1.69
SEED-IV	26.04 ± 2.14	26.58 ± 0.99	27.87 ± 2.85	27.28 ± 0.71	26.78 ± 1.31	26.81 ± 0.81
SEED-V	24.68 ± 1.44	22.88 ± 1.01	24.31 ± 2.14	22.97 ± 1.61	22.45 ± 0.23	23.10 ± 2.32
SEED-VII	18.55 ± 1.59	18.65 ± 1.74	17.46 ± 1.63	18.57 ± 1.06	17.94 ± 1.99	18.33 ± 1.27
ISRUC S2†	26.92 ± 4.48	23.17 ± 5.45	25.67 ± 5.22	22.47 ± 5.03	25.03 ± 5.32	28.60 ± 1.20
ISRUC S3	73.24 ± 4.97	73.97 ± 1.54	73.28 ± 0.97	74.36 ± 1.11	73.94 ± 3.97	75.45 ± 2.09
SleepEDF	70.50 ± 1.18	71.79 ± 1.09	70.81 ± 3.02	71.62 ± 0.77	70.17 ± 0.41	71.83 ± 1.29
HMC	44.69 ± 1.34	45.97 ± 1.85	44.93 ± 0.67	46.23 ± 1.89	43.70 ± 2.83	46.90 ± 2.61
TUAB	74.86 ± 2.02	75.62 ± 0.94	74.83 ± 1.71	75.30 ± 2.07	73.87 ± 1.64	75.48 ± 1.29
TUEV	47.80 ± 10.04	49.32 ± 9.54	48.05 ± 11.17	49.38 ± 7.15	47.69 ± 6.82	48.82 ± 10.39
TUSL	67.22 ± 3.54	65.26 ± 1.32	69.14 ± 8.03	59.05 ± 11.37	75.94 ± 7.50	65.47 ± 8.77
EEGDenoiseNet	92.48 ± 1.57	92.81 ± 2.52	92.83 ± 1.69	92.55 ± 1.32	85.32 ± 1.20	91.07 ± 2.93
Macro mean	48.54 ± 1.33	48.44 ± 1.38	48.16 ± 1.45	47.59 ± 1.52	47.34 ± 1.35	48.48 ± 2.55

Table 33. EEG–LLM results for Quark-H (epoch 49).

Task	Qwen3-0.6B	Qwen3-1.7B	Qwen3-4B	Qwen3-8B	Qwen3-14B	Qwen3-32B
AD65	51.60 ± 10.18	48.55 ± 9.01	45.63 ± 6.49	45.90 ± 7.73	45.91 ± 1.00	47.84 ± 9.54
PD31	43.42 ± 11.43	49.70 ± 9.15	44.62 ± 11.34	50.82 ± 11.61	47.78 ± 13.38	44.00 ± 10.67
BCIC IV-1	51.67 ± 3.33	48.83 ± 3.88	50.00 ± 1.50	52.50 ± 1.32	48.00 ± 6.54	49.33 ± 0.58
BCIC IV-2a	32.29 ± 1.71	34.03 ± 3.93	34.20 ± 3.94	33.62 ± 1.97	32.52 ± 4.10	30.21 ± 1.59
SHU MI	51.43 ± 0.46	52.50 ± 1.37	50.92 ± 1.11	53.28 ± 0.76	50.27 ± 0.40	51.47 ± 0.70
BETA SSVEP	2.35 ± 0.34	2.68 ± 0.36	2.71 ± 0.36	2.65 ± 0.22	2.86 ± 0.24	2.32 ± 0.24
Monitoring ErrP	54.85 ± 1.59	59.85 ± 2.43	57.43 ± 3.39	56.57 ± 0.86	59.05 ± 3.51	58.21 ± 2.20
EEGMAT	67.91 ± 10.34	63.28 ± 10.27	65.13 ± 7.26	68.69 ± 10.51	62.52 ± 7.32	61.59 ± 10.48
LEMON gender	66.12 ± 1.64	65.74 ± 3.33	67.51 ± 3.03	65.30 ± 3.30	64.91 ± 4.60	66.10 ± 3.67
LEMON extraversion	51.81 ± 1.02	51.87 ± 0.93	52.03 ± 0.24	53.21 ± 1.54	52.24 ± 2.05	49.47 ± 0.57
FACED	12.92 ± 0.44	14.32 ± 1.21	13.83 ± 0.80	15.12 ± 1.83	13.88 ± 1.91	13.23 ± 0.92
SEED	51.51 ± 1.78	49.14 ± 1.74	49.43 ± 1.22	50.83 ± 1.62	48.18 ± 4.52	47.85 ± 2.97
SEED-IV	27.93 ± 1.19	27.02 ± 1.35	26.96 ± 1.47	27.55 ± 0.94	27.64 ± 0.31	27.82 ± 1.07
SEED-V	25.26 ± 0.95	23.53 ± 0.81	25.19 ± 2.50	23.45 ± 1.38	23.65 ± 0.66	23.74 ± 0.88
SEED-VII	18.06 ± 3.10	18.07 ± 0.74	17.88 ± 0.72	18.36 ± 0.75	16.63 ± 1.18	18.43 ± 1.58

Continued on the next page.

Table 33. EEG–LLM results for Quark-H (epoch 49) (continued).

Task	Qwen3-0.6B	Qwen3-1.7B	Qwen3-4B	Qwen3-8B	Qwen3-14B	Qwen3-32B
ISRUC S2†	24.00 ± 3.17	29.50 ± 4.79	24.25 ± 6.08	27.87 ± 5.12	21.97 ± 0.67	24.51 ± 3.39
ISRUC S3	75.31 ± 2.43	73.74 ± 2.77	72.59 ± 3.24	74.52 ± 2.10	70.47 ± 6.75	70.54 ± 1.09
SleepEDF	69.04 ± 1.92	71.33 ± 1.12	69.61 ± 2.58	70.88 ± 2.32	68.04 ± 2.38	68.85 ± 2.61
HMC	52.57 ± 1.57	54.51 ± 2.84	53.33 ± 2.46	54.40 ± 2.55	51.43 ± 4.14	53.01 ± 2.65
TUAB	73.40 ± 1.95	76.04 ± 0.44	75.36 ± 1.48	76.57 ± 1.33	74.90 ± 1.19	74.58 ± 0.44
TUEV	49.63 ± 7.68	47.44 ± 2.54	53.13 ± 6.41	47.58 ± 8.64	46.66 ± 5.58	52.61 ± 7.90
TUSL	64.42 ± 5.26	70.57 ± 3.38	62.42 ± 10.31	62.25 ± 10.06	62.84 ± 7.08	67.72 ± 11.60
EEGDenoiseNet	92.37 ± 0.74	91.22 ± 1.17	90.69 ± 2.85	90.77 ± 1.98	89.34 ± 1.85	87.60 ± 4.44
Macro mean	48.25 ± 0.82	48.84 ± 0.96	48.04 ± 1.28	48.81 ± 1.52	47.03 ± 0.33	47.44 ± 1.56

Table 34. EEG–LLM results for Quark-G (epoch 29).

Task	Qwen3-0.6B	Qwen3-1.7B	Qwen3-4B	Qwen3-8B	Qwen3-14B	Qwen3-32B
AD65	52.81 ± 5.63	51.54 ± 5.15	49.68 ± 5.41	50.71 ± 5.96	52.29 ± 7.07	49.15 ± 6.69
PD31	41.25 ± 7.70	39.61 ± 10.31	38.41 ± 11.48	42.00 ± 5.90	41.28 ± 5.71	44.52 ± 2.40
BCIC IV-1	52.50 ± 1.80	47.83 ± 2.02	45.33 ± 5.51	49.33 ± 1.76	50.67 ± 1.61	48.00 ± 2.18
BCIC IV-2a	31.25 ± 1.08	35.76 ± 1.51	31.83 ± 2.53	34.66 ± 2.09	36.28 ± 0.92	34.03 ± 4.56
SHU MI	54.74 ± 1.03	51.96 ± 3.15	51.94 ± 1.34	52.97 ± 2.03	51.40 ± 2.22	51.42 ± 1.31
BETA SSVEP	2.77 ± 0.54	3.07 ± 0.65	2.89 ± 0.22	3.36 ± 1.00	2.83 ± 0.42	2.74 ± 0.49
Monitoring ErrP	58.13 ± 2.43	55.15 ± 4.37	56.91 ± 2.17	56.84 ± 4.17	57.50 ± 1.40	54.07 ± 3.37
EEGMAT	63.75 ± 10.88	73.32 ± 10.39	70.43 ± 8.40	70.86 ± 9.99	65.62 ± 14.90	67.79 ± 8.97
LEMON gender	67.76 ± 5.26	65.73 ± 2.92	67.93 ± 5.68	65.91 ± 2.74	65.42 ± 2.88	65.14 ± 2.79
LEMON extraversion	50.51 ± 2.83	51.08 ± 1.86	51.25 ± 0.89	51.75 ± 2.21	51.58 ± 0.56	50.86 ± 1.34
FACED	15.84 ± 3.63	18.56 ± 3.00	13.43 ± 2.53	19.71 ± 1.97	16.99 ± 2.98	17.08 ± 2.99
SEED	47.93 ± 3.73	49.67 ± 1.89	44.46 ± 5.71	49.86 ± 1.44	45.91 ± 2.80	45.50 ± 1.86
SEED-IV	27.89 ± 1.14	26.71 ± 0.36	28.83 ± 0.74	28.39 ± 0.36	27.93 ± 0.46	27.38 ± 0.90
SEED-V	25.96 ± 2.18	24.26 ± 1.08	26.00 ± 1.52	24.16 ± 1.21	24.99 ± 0.90	23.13 ± 1.19
SEED-VII	17.71 ± 1.62	19.12 ± 0.53	16.05 ± 2.18	18.49 ± 0.95	17.66 ± 0.61	18.29 ± 0.92
ISRUC S2†	30.05 ± 4.81	26.26 ± 5.44	30.66 ± 2.17	26.56 ± 4.18	31.73 ± 6.48	29.65 ± 3.84
ISRUC S3	68.87 ± 4.64	71.02 ± 3.09	67.82 ± 9.26	70.34 ± 2.74	70.13 ± 1.68	68.78 ± 3.72
SleepEDF	68.63 ± 3.32	70.39 ± 1.46	67.54 ± 3.58	70.44 ± 1.77	69.26 ± 3.58	70.05 ± 1.54
HMC	49.28 ± 2.17	50.32 ± 2.33	45.10 ± 3.26	50.92 ± 1.54	50.16 ± 2.31	49.04 ± 0.62
TUAB	75.68 ± 1.89	75.82 ± 2.00	74.22 ± 3.10	76.59 ± 0.90	75.39 ± 1.80	76.29 ± 1.61
TUEV	54.26 ± 9.30	53.87 ± 12.40	48.82 ± 5.84	56.74 ± 6.24	53.39 ± 11.05	54.77 ± 9.91
TUSL	65.80 ± 16.78	62.87 ± 13.16	67.05 ± 4.94	63.83 ± 9.57	50.31 ± 24.47	61.05 ± 11.57
EEGDenoiseNet	84.18 ± 3.71	81.66 ± 2.14	80.82 ± 6.84	79.80 ± 1.94	74.65 ± 7.79	77.46 ± 6.10
Macro mean	48.15 ± 2.02	48.07 ± 1.86	46.84 ± 1.62	48.44 ± 1.06	47.10 ± 2.41	47.23 ± 1.26

Table 35. EEG–LLM results for Atom-B (epoch 49).

Task	Qwen3-0.6B	Qwen3-1.7B	Qwen3-4B	Qwen3-8B	Qwen3-14B	Qwen3-32B
AD65	49.30 ± 5.27	50.27 ± 5.41	54.80 ± 10.89	54.85 ± 4.42	46.25 ± 6.19	45.74 ± 5.75
PD31	40.56 ± 7.36	43.19 ± 7.66	46.30 ± 12.81	47.04 ± 5.87	41.35 ± 6.01	45.32 ± 6.07
BCIC IV-1	51.00 ± 4.36	49.67 ± 4.04	51.67 ± 7.29	50.00 ± 1.80	51.83 ± 3.88	48.67 ± 3.21
BCIC IV-2a	33.62 ± 5.49	33.80 ± 0.72	34.84 ± 2.37	34.03 ± 2.34	32.64 ± 3.35	33.10 ± 5.28
SHU MI	51.21 ± 0.38	50.42 ± 1.24	51.36 ± 1.31	50.51 ± 1.91	51.54 ± 2.27	49.88 ± 0.89
BETA SSVEP	2.77 ± 0.24	2.65 ± 0.14	2.44 ± 0.31	2.53 ± 0.05	2.68 ± 0.45	2.98 ± 0.19
Monitoring ErrP	54.74 ± 1.98	55.87 ± 2.71	57.01 ± 2.60	54.58 ± 1.70	56.48 ± 0.57	56.89 ± 1.55
EEGMAT	75.19 ± 6.67	68.24 ± 8.37	69.30 ± 7.00	70.11 ± 12.01	67.14 ± 8.65	71.16 ± 2.26
LEMON gender	61.91 ± 2.51	61.94 ± 0.44	61.60 ± 3.02	64.64 ± 4.52	61.58 ± 2.01	62.06 ± 3.67
LEMON extraversion	50.24 ± 1.21	50.27 ± 0.59	50.37 ± 0.21	49.78 ± 0.30	50.41 ± 1.87	49.64 ± 0.16
FACED	14.54 ± 0.53	13.58 ± 0.79	15.66 ± 1.11	14.21 ± 0.49	13.68 ± 1.41	14.29 ± 1.08
SEED	52.94 ± 2.26	52.87 ± 1.71	52.71 ± 3.01	51.70 ± 2.17	50.52 ± 0.46	51.53 ± 4.11
SEED-IV	27.32 ± 1.95	29.42 ± 2.82	26.51 ± 1.33	29.37 ± 0.38	27.53 ± 1.92	27.91 ± 1.00
SEED-V	22.40 ± 0.96	24.19 ± 0.52	23.44 ± 0.65	23.33 ± 0.52	23.86 ± 1.09	22.82 ± 2.32
SEED-VII	18.80 ± 0.79	19.01 ± 0.71	19.54 ± 0.67	17.87 ± 1.28	18.65 ± 1.09	17.90 ± 1.06
ISRUC S2†	27.06 ± 5.08	28.51 ± 2.88	27.16 ± 4.27	27.73 ± 5.50	23.96 ± 4.63	23.21 ± 1.91
ISRUC S3	73.54 ± 1.94	72.43 ± 2.90	75.15 ± 3.90	72.86 ± 2.76	73.60 ± 1.87	72.46 ± 1.31

Continued on the next page.

Table 35. EEG–LLM results for Atom-B (epoch 49) (continued).

Task	Qwen3-0.6B	Qwen3-1.7B	Qwen3-4B	Qwen3-8B	Qwen3-14B	Qwen3-32B
SleepEDF	68.15 ± 1.95	67.49 ± 2.55	69.23 ± 0.21	66.04 ± 4.54	68.30 ± 0.87	67.59 ± 0.77
HMC	55.64 ± 2.72	55.98 ± 3.09	56.12 ± 2.55	54.49 ± 5.95	55.36 ± 1.06	55.50 ± 1.99
TUAB	74.88 ± 2.15	75.10 ± 2.55	75.60 ± 1.04	75.45 ± 1.40	75.00 ± 1.39	75.52 ± 1.43
TUEV	54.81 ± 8.43	52.11 ± 10.32	49.45 ± 5.21	49.52 ± 13.89	54.81 ± 9.18	53.28 ± 12.03
TUSL	50.12 ± 19.01	57.70 ± 14.45	54.75 ± 15.89	65.59 ± 5.54	60.09 ± 9.57	67.24 ± 3.70
EEGDenoiseNet	97.97 ± 0.23	98.04 ± 1.27	98.02 ± 0.22	97.93 ± 0.18	97.17 ± 0.39	97.38 ± 0.40
Macro mean	48.21 ± 1.77	48.38 ± 1.78	48.83 ± 1.80	48.88 ± 1.96	48.02 ± 1.59	48.35 ± 1.40

Table 36. EEG–LLM results for Atom-L (epoch 49).

Task	Qwen3-0.6B	Qwen3-1.7B	Qwen3-4B	Qwen3-8B	Qwen3-14B	Qwen3-32B
AD65	51.49 ± 10.19	52.42 ± 6.16	50.67 ± 2.97	49.93 ± 4.99	50.04 ± 4.36	48.40 ± 1.47
PD31	39.16 ± 3.23	40.19 ± 12.95	42.85 ± 9.57	36.44 ± 5.79	41.83 ± 8.94	41.72 ± 11.59
BCIC IV-1	50.00 ± 1.32	49.50 ± 2.18	48.17 ± 3.01	46.83 ± 1.04	51.33 ± 1.26	50.67 ± 0.76
BCIC IV-2a	35.01 ± 1.76	35.13 ± 2.37	32.64 ± 2.87	33.91 ± 4.02	35.76 ± 2.46	33.28 ± 1.86
SHU MI	51.56 ± 0.92	51.89 ± 0.62	51.43 ± 1.15	51.96 ± 1.47	51.37 ± 0.91	51.08 ± 0.44
BETA SSVEP	2.86 ± 0.78	2.86 ± 0.78	2.47 ± 0.05	2.74 ± 0.10	2.83 ± 0.49	2.38 ± 0.05
Monitoring ErrP	54.32 ± 2.21	56.56 ± 0.80	53.19 ± 2.09	54.62 ± 2.55	54.68 ± 2.93	53.30 ± 3.20
EEGMAT	70.23 ± 6.54	72.55 ± 3.22	66.84 ± 11.51	64.68 ± 3.62	63.44 ± 5.09	69.02 ± 3.73
LEMON gender	61.87 ± 0.84	61.96 ± 2.07	60.24 ± 4.18	61.46 ± 3.51	60.56 ± 2.67	59.96 ± 1.57
LEMON extraversion	50.47 ± 3.13	51.02 ± 2.40	49.39 ± 1.22	50.57 ± 1.84	50.78 ± 2.34	50.89 ± 2.83
FACED	15.04 ± 1.49	13.45 ± 1.32	13.38 ± 0.40	13.07 ± 0.87	12.69 ± 0.55	13.60 ± 1.50
SEED	50.49 ± 2.43	52.50 ± 1.19	50.04 ± 1.71	52.65 ± 2.45	48.15 ± 1.99	51.25 ± 1.77
SEED-IV	26.29 ± 1.01	27.18 ± 1.20	28.48 ± 0.65	27.38 ± 2.39	27.24 ± 1.47	27.49 ± 1.82
SEED-V	24.37 ± 1.35	24.39 ± 1.54	25.84 ± 2.47	22.79 ± 0.50	22.88 ± 2.21	23.39 ± 1.05
SEED-VII	18.03 ± 0.42	18.50 ± 0.59	18.74 ± 1.98	18.77 ± 1.57	18.56 ± 1.73	18.77 ± 1.28
ISRUC S2†	30.23 ± 2.76	32.71 ± 1.75	30.70 ± 4.76	29.12 ± 4.55	30.72 ± 1.67	33.95 ± 2.96
ISRUC S3	74.01 ± 0.34	73.91 ± 1.75	73.46 ± 2.29	75.58 ± 2.95	76.17 ± 1.39	73.39 ± 2.06
SleepEDF	69.46 ± 2.17	69.22 ± 1.22	67.22 ± 2.64	69.64 ± 1.31	70.05 ± 0.84	67.44 ± 3.29
HMC	54.87 ± 1.44	54.01 ± 1.28	51.94 ± 1.46	54.37 ± 1.44	53.53 ± 0.83	53.57 ± 2.16
TUAB	77.16 ± 1.35	76.46 ± 1.31	75.45 ± 3.25	77.39 ± 1.64	76.29 ± 2.96	75.61 ± 2.11
TUEV	58.34 ± 9.75	57.20 ± 11.89	54.38 ± 8.74	55.22 ± 8.52	57.62 ± 11.37	51.81 ± 8.12
TUSL	63.25 ± 16.76	64.66 ± 14.33	63.22 ± 8.45	62.06 ± 9.91	61.32 ± 3.47	71.15 ± 1.24
EEGDenoiseNet	96.95 ± 1.06	97.10 ± 0.71	98.17 ± 0.42	96.95 ± 0.56	96.87 ± 0.56	95.78 ± 1.63
Macro mean	48.93 ± 2.05	49.36 ± 2.16	48.21 ± 1.26	48.18 ± 1.13	48.47 ± 0.82	48.60 ± 1.03

Table 37. EEG–LLM results for Atom-H (epoch 49).

Task	Qwen3-0.6B	Qwen3-1.7B	Qwen3-4B	Qwen3-8B	Qwen3-14B	Qwen3-32B
AD65	54.43 ± 5.61	53.29 ± 6.21	52.31 ± 9.07	57.39 ± 9.42	50.73 ± 6.36	53.02 ± 9.06
PD31	45.29 ± 10.84	46.65 ± 10.20	44.51 ± 4.30	48.46 ± 8.25	49.44 ± 5.88	41.32 ± 12.30
BCIC IV-1	50.83 ± 0.58	51.17 ± 3.06	48.50 ± 1.80	52.50 ± 4.58	49.50 ± 3.91	50.50 ± 3.50
BCIC IV-2a	32.47 ± 3.38	34.43 ± 0.70	35.59 ± 2.30	34.49 ± 1.94	34.78 ± 4.20	28.65 ± 2.56
SHU MI	51.07 ± 0.59	51.59 ± 0.96	50.82 ± 1.68	50.40 ± 0.46	52.21 ± 0.84	51.48 ± 0.94
BETA SSVEP	2.44 ± 0.19	3.69 ± 0.95	2.74 ± 0.34	2.86 ± 0.18	3.45 ± 0.67	3.01 ± 0.14
Monitoring ErrP	54.08 ± 3.30	57.07 ± 3.52	57.40 ± 4.72	56.56 ± 1.38	56.16 ± 3.38	58.08 ± 3.53
EEGMAT	70.68 ± 9.75	71.31 ± 7.63	70.24 ± 5.76	70.53 ± 10.29	69.46 ± 10.86	67.14 ± 3.03
LEMON gender	61.20 ± 1.02	62.33 ± 0.60	61.24 ± 1.40	63.23 ± 0.56	62.29 ± 1.66	61.87 ± 0.21
LEMON extraversion	51.49 ± 3.13	52.03 ± 0.61	49.99 ± 1.35	52.79 ± 1.51	50.77 ± 1.87	51.04 ± 1.25
FACED	14.26 ± 0.50	15.69 ± 2.53	13.13 ± 0.40	15.69 ± 0.83	14.78 ± 1.41	14.75 ± 0.18
SEED	52.66 ± 0.49	53.83 ± 0.82	52.76 ± 0.88	54.03 ± 1.08	53.01 ± 0.98	50.61 ± 1.70
SEED-IV	26.53 ± 0.39	26.98 ± 0.77	28.94 ± 0.81	27.39 ± 1.53	27.64 ± 1.51	27.69 ± 1.01
SEED-V	24.69 ± 1.10	22.86 ± 1.05	24.93 ± 2.38	23.78 ± 1.61	23.36 ± 2.01	22.82 ± 1.22
SEED-VII	19.22 ± 1.27	19.62 ± 2.23	18.94 ± 0.37	18.37 ± 0.89	18.69 ± 1.36	17.84 ± 1.01
ISRUC S2†	26.24 ± 1.86	25.03 ± 1.08	28.89 ± 5.78	23.22 ± 1.71	29.37 ± 5.71	32.77 ± 1.03
ISRUC S3	78.13 ± 2.36	76.30 ± 0.75	75.37 ± 3.41	76.88 ± 0.39	78.36 ± 2.42	74.78 ± 1.54
SleepEDF	69.88 ± 2.09	71.16 ± 1.17	69.99 ± 2.65	71.74 ± 1.56	71.20 ± 1.35	70.64 ± 1.09
HMC	56.99 ± 0.61	56.99 ± 1.06	55.80 ± 2.26	56.54 ± 1.16	57.78 ± 1.04	56.33 ± 1.41

Continued on the next page.

Table 37. EEG–LLM results for Atom-H (epoch 49) (continued).

Task	Qwen3-0.6B	Qwen3-1.7B	Qwen3-4B	Qwen3-8B	Qwen3-14B	Qwen3-32B
TUAB	77.16 ± 2.61	77.82 ± 0.70	76.75 ± 1.81	77.97 ± 1.32	77.64 ± 1.26	77.43 ± 0.95
TUEV	57.32 ± 10.07	54.52 ± 5.92	49.14 ± 4.51	56.69 ± 9.65	53.66 ± 4.82	53.83 ± 3.10
TUSL	69.63 ± 9.81	56.21 ± 16.83	60.82 ± 10.10	54.29 ± 15.21	60.80 ± 10.14	68.51 ± 12.28
EEGDenoiseNet	86.30 ± 3.81	83.91 ± 0.59	84.52 ± 1.97	80.75 ± 2.87	84.85 ± 2.81	83.76 ± 4.94
Macro mean	49.26 ± 1.87	48.89 ± 1.80	48.41 ± 1.69	48.98 ± 1.83	49.13 ± 1.69	48.60 ± 1.27

Table 38. EEG–LLM results for Molecule-B (epoch 49).

Task	Qwen3-0.6B	Qwen3-1.7B	Qwen3-4B	Qwen3-8B	Qwen3-14B	Qwen3-32B
AD65	46.89 ± 4.30	45.28 ± 4.39	43.48 ± 0.90	45.92 ± 1.18	42.74 ± 3.92	46.25 ± 2.28
PD31	48.05 ± 8.00	44.80 ± 7.41	43.44 ± 7.11	44.52 ± 5.00	44.61 ± 6.43	46.16 ± 7.66
BCIC IV-1	51.33 ± 0.58	46.83 ± 3.40	50.67 ± 1.89	48.67 ± 2.36	48.83 ± 2.36	46.33 ± 6.79
BCIC IV-2a	36.52 ± 4.34	33.10 ± 1.26	33.91 ± 2.70	34.03 ± 3.81	33.51 ± 3.30	32.23 ± 0.72
SHU MI	51.70 ± 1.29	53.61 ± 0.97	52.05 ± 1.73	51.47 ± 0.89	50.09 ± 0.26	50.78 ± 1.21
BETA SSVEP	2.41 ± 0.56	2.53 ± 0.34	2.89 ± 0.10	2.74 ± 0.14	2.92 ± 0.22	2.50 ± 0.63
Monitoring ErrP	56.57 ± 2.39	57.89 ± 2.33	54.45 ± 2.46	54.21 ± 4.60	56.13 ± 1.98	53.53 ± 1.76
EEGMAT	65.60 ± 9.36	62.21 ± 8.82	60.81 ± 8.71	63.30 ± 9.95	64.52 ± 12.21	67.46 ± 10.67
LEMON gender	58.79 ± 1.70	56.82 ± 3.94	59.29 ± 2.55	58.82 ± 2.05	55.76 ± 4.31	59.19 ± 3.90
LEMON extraversion	49.88 ± 2.44	50.97 ± 1.32	50.51 ± 1.13	50.06 ± 2.11	52.05 ± 2.93	49.22 ± 1.86
FACED	13.89 ± 1.20	15.30 ± 0.51	14.03 ± 1.73	14.31 ± 2.06	13.62 ± 0.32	13.51 ± 0.51
SEED	50.69 ± 1.53	51.83 ± 0.83	50.45 ± 2.51	50.56 ± 1.90	46.50 ± 2.36	49.51 ± 1.35
SEED-IV	27.13 ± 2.13	26.36 ± 1.32	27.31 ± 0.74	27.82 ± 0.80	27.43 ± 0.79	26.86 ± 1.40
SEED-V	23.48 ± 1.03	22.39 ± 0.98	23.13 ± 0.95	23.44 ± 1.25	22.30 ± 1.54	23.14 ± 1.02
SEED-VII	17.68 ± 1.29	17.91 ± 0.91	19.28 ± 0.68	18.88 ± 0.52	17.43 ± 1.32	18.43 ± 0.82
ISRUC S2†	23.43 ± 3.35	26.80 ± 4.42	25.16 ± 3.46	29.92 ± 3.02	26.55 ± 3.73	28.14 ± 5.25
ISRUC S3	71.72 ± 1.23	70.61 ± 1.36	70.10 ± 1.08	71.29 ± 2.09	69.70 ± 0.84	67.20 ± 2.16
SleepEDF	66.34 ± 1.40	66.95 ± 2.09	66.37 ± 1.49	66.98 ± 1.90	65.97 ± 1.32	64.24 ± 1.53
HMC	48.10 ± 2.20	49.07 ± 2.21	48.25 ± 3.75	48.65 ± 3.46	47.87 ± 2.56	47.45 ± 2.66
TUAB	74.66 ± 3.24	76.05 ± 2.50	75.15 ± 1.15	74.48 ± 2.15	75.14 ± 1.27	75.85 ± 0.97
TUEV	51.43 ± 6.57	53.35 ± 9.21	51.26 ± 9.86	50.50 ± 5.51	49.71 ± 13.98	56.08 ± 8.38
TUSL	67.26 ± 10.67	58.53 ± 6.61	62.35 ± 3.76	70.57 ± 6.93	62.29 ± 8.44	65.44 ± 8.21
EEGDenoiseNet	90.84 ± 1.26	90.07 ± 1.39	89.29 ± 1.80	90.34 ± 0.87	89.11 ± 0.69	91.92 ± 1.15
Macro mean	47.58 ± 1.84	46.92 ± 1.41	46.68 ± 1.33	47.46 ± 1.37	46.29 ± 2.41	47.02 ± 1.59

Table 39. EEG–LLM results for Molecule-L (epoch 29).

Task	Qwen3-0.6B	Qwen3-1.7B	Qwen3-4B	Qwen3-8B	Qwen3-14B	Qwen3-32B
AD65	47.28 ± 4.37	45.84 ± 10.38	47.20 ± 6.55	48.55 ± 3.23	47.89 ± 1.52	51.18 ± 3.29
PD31	47.18 ± 7.62	52.61 ± 5.62	48.94 ± 5.80	48.48 ± 4.93	47.91 ± 9.01	49.00 ± 7.47
BCIC IV-1	47.50 ± 1.80	50.00 ± 2.65	49.33 ± 1.89	45.83 ± 5.13	50.67 ± 1.26	49.67 ± 1.89
BCIC IV-2a	35.47 ± 4.37	36.17 ± 2.54	38.60 ± 1.58	36.00 ± 1.58	34.49 ± 2.34	33.51 ± 1.67
SHU MI	52.05 ± 3.46	52.73 ± 0.59	52.10 ± 2.00	52.90 ± 2.36	52.39 ± 1.63	51.39 ± 1.08
BETA SSVEP	3.27 ± 0.41	2.83 ± 0.27	2.77 ± 0.67	3.12 ± 0.18	3.24 ± 0.44	2.56 ± 0.10
Monitoring ErrP	55.17 ± 2.94	54.99 ± 3.10	56.75 ± 1.94	53.46 ± 3.71	56.63 ± 1.21	58.47 ± 3.99
EEGMAT	64.67 ± 14.35	68.06 ± 5.91	71.31 ± 3.95	64.52 ± 12.43	62.05 ± 4.55	68.23 ± 8.38
LEMON gender	58.97 ± 1.22	57.56 ± 2.48	59.20 ± 3.18	59.51 ± 1.25	57.17 ± 0.62	59.98 ± 3.00
LEMON extraversion	50.64 ± 1.30	50.32 ± 0.79	50.24 ± 0.66	51.59 ± 0.80	50.95 ± 0.64	50.04 ± 2.34
FACED	15.72 ± 1.47	15.19 ± 0.86	14.93 ± 0.44	15.11 ± 1.18	13.61 ± 1.21	14.26 ± 1.63
SEED	51.72 ± 2.22	52.71 ± 1.91	50.86 ± 2.86	49.47 ± 1.17	49.30 ± 1.17	52.26 ± 0.85
SEED-IV	27.50 ± 0.61	26.97 ± 0.32	27.78 ± 1.50	27.01 ± 0.62	26.83 ± 0.85	28.43 ± 1.22
SEED-V	21.86 ± 0.74	22.24 ± 2.27	22.14 ± 0.94	20.96 ± 0.40	21.74 ± 0.20	23.21 ± 1.10
SEED-VII	18.75 ± 0.95	17.98 ± 0.94	18.18 ± 1.05	18.07 ± 1.19	17.54 ± 0.10	18.73 ± 1.09
ISRUC S2†	31.02 ± 4.99	28.17 ± 4.37	34.18 ± 2.89	24.18 ± 0.78	25.66 ± 1.18	28.62 ± 7.42
ISRUC S3	72.86 ± 2.72	72.97 ± 1.88	71.59 ± 0.43	74.36 ± 1.94	73.76 ± 2.17	70.69 ± 1.60
SleepEDF	70.30 ± 0.48	69.80 ± 0.48	69.50 ± 1.52	70.19 ± 0.86	67.82 ± 3.70	68.12 ± 2.54
HMC	54.38 ± 3.27	53.92 ± 3.57	53.65 ± 2.91	54.34 ± 4.14	52.96 ± 1.50	53.83 ± 3.77
TUAB	76.80 ± 1.61	76.49 ± 1.89	75.33 ± 1.95	76.32 ± 1.26	75.00 ± 2.88	75.65 ± 1.94
TUEV	53.91 ± 9.10	53.64 ± 11.65	54.91 ± 4.53	56.91 ± 8.48	47.72 ± 0.65	55.56 ± 8.29

Continued on the next page.

Table 39. EEG–LLM results for Molecule-L (epoch 29) (continued).

Task	Qwen3-0.6B	Qwen3-1.7B	Qwen3-4B	Qwen3-8B	Qwen3-14B	Qwen3-32B
TUSL	62.67 ± 4.21	66.88 ± 12.64	68.11 ± 6.20	64.90 ± 14.50	60.32 ± 6.93	65.80 ± 20.52
EEGDenoiseNet	94.62 ± 0.25	95.51 ± 0.59	93.75 ± 2.33	93.85 ± 1.60	93.05 ± 0.84	93.05 ± 2.79
Macro mean	48.45 ± 1.69	48.85 ± 1.81	49.19 ± 0.82	48.25 ± 1.60	47.33 ± 0.41	48.79 ± 2.06

Table 40. EEG–LLM results for Molecule-H (epoch 9).

Task	Qwen3-0.6B	Qwen3-1.7B	Qwen3-4B	Qwen3-8B	Qwen3-14B	Qwen3-32B
AD65	53.50 ± 4.61	48.32 ± 5.74	49.69 ± 1.72	50.44 ± 8.14	49.30 ± 0.49	50.70 ± 3.55
PD31	43.78 ± 6.05	45.64 ± 2.59	45.79 ± 8.54	45.41 ± 9.53	44.06 ± 12.60	47.94 ± 11.01
BCIC IV-1	52.00 ± 2.78	46.83 ± 2.25	52.83 ± 5.39	49.33 ± 3.51	48.50 ± 1.32	51.83 ± 2.52
BCIC IV-2a	35.94 ± 2.18	36.34 ± 0.20	34.78 ± 6.36	38.19 ± 4.96	34.95 ± 5.78	35.71 ± 1.82
SHU MI	53.79 ± 2.60	51.47 ± 2.28	51.01 ± 0.69	51.30 ± 2.53	49.58 ± 1.27	51.27 ± 2.98
BETA SSVEP	3.84 ± 1.03	3.69 ± 0.74	3.54 ± 0.74	3.24 ± 0.27	3.12 ± 0.59	3.42 ± 0.21
Monitoring ErrP	54.66 ± 3.43	57.80 ± 3.03	57.12 ± 3.45	56.25 ± 3.91	54.18 ± 2.21	58.06 ± 1.28
EEGMAT	70.84 ± 5.20	69.76 ± 6.16	69.45 ± 9.85	71.77 ± 10.23	65.29 ± 14.12	65.44 ± 4.19
LEMON gender	61.78 ± 0.84	60.94 ± 1.66	61.29 ± 1.79	62.04 ± 1.02	58.15 ± 1.88	62.30 ± 0.80
LEMON extraversion	50.40 ± 1.74	50.70 ± 2.33	51.68 ± 1.31	50.52 ± 3.62	50.43 ± 1.58	51.51 ± 1.60
FACED	16.26 ± 2.21	16.66 ± 1.67	14.80 ± 0.07	15.81 ± 1.29	12.80 ± 1.23	14.90 ± 0.44
SEED	52.52 ± 2.89	53.60 ± 2.38	52.52 ± 4.49	53.39 ± 1.46	51.55 ± 2.89	51.58 ± 1.04
SEED-IV	29.34 ± 1.37	28.00 ± 0.95	27.04 ± 0.57	28.73 ± 0.93	28.28 ± 1.07	26.50 ± 0.78
SEED-V	23.97 ± 2.54	22.95 ± 0.71	22.92 ± 0.65	23.14 ± 1.25	24.34 ± 1.93	22.18 ± 0.87
SEED-VII	18.48 ± 1.21	19.07 ± 1.08	18.37 ± 0.18	20.56 ± 1.05	17.93 ± 1.71	17.97 ± 1.45
ISRUC S2[†]	32.46 ± 6.20	33.03 ± 5.32	31.68 ± 4.54	36.28 ± 1.22	28.82 ± 4.32	30.41 ± 5.29
ISRUC S3	75.89 ± 1.63	75.96 ± 0.80	75.78 ± 2.83	75.32 ± 4.01	73.48 ± 4.63	76.74 ± 0.29
SleepEDF	69.95 ± 3.07	70.29 ± 2.09	70.40 ± 1.46	70.98 ± 1.83	68.80 ± 2.38	70.46 ± 1.50
HMC	55.66 ± 1.84	56.90 ± 2.28	57.02 ± 2.39	57.48 ± 3.51	56.32 ± 2.02	57.16 ± 0.67
TUAB	75.90 ± 4.05	77.46 ± 1.59	76.75 ± 1.89	76.99 ± 2.01	75.73 ± 1.55	77.50 ± 1.21
TUEV	56.60 ± 8.77	60.40 ± 13.00	58.65 ± 10.74	61.72 ± 14.45	54.27 ± 15.06	63.53 ± 12.64
TUSL	63.08 ± 4.75	68.13 ± 9.69	61.83 ± 16.11	71.74 ± 6.70	66.11 ± 13.09	63.29 ± 10.49
EEGDenoiseNet	90.89 ± 2.26	89.24 ± 0.48	87.43 ± 2.89	90.96 ± 1.12	87.47 ± 4.89	89.71 ± 1.83
Macro mean	49.63 ± 1.03	49.70 ± 1.68	49.23 ± 2.59	50.50 ± 2.17	47.98 ± 2.89	49.57 ± 1.49

H.1 Training-set size

Table 41 gives the training-window counts shared by the EEG–LLM configurations. Sampling is applied to training only, with at most 40 windows per registered subject–class unit. The split conventions are described in Appendix C; Section 7 gives the joint-training settings.

Table 41. Training windows before and after the per-subject/class cap.

	Seed42		Seed10		Seed5	
Task	Before → after	Retained	Before → after	Retained	Before → after	Retained
AD65	5,500 → 2,790	50.7%	5,572 → 2,790	50.1%	5,609 → 2,790	49.7%
PD31	7,710 → 1,000	13.0%	7,354 → 1,000	13.6%	7,544 → 1,000	13.3%
BCIC IV-1	1,000 → 400	40.0%	1,000 → 400	40.0%	1,000 → 400	40.0%
BCIC IV-2a	4,032 → 1,120	27.8%	4,032 → 1,120	27.8%	4,032 → 1,120	27.8%
SHU MI	9,556 → 1,600	16.7%	9,559 → 1,600	16.7%	9,575 → 1,600	16.7%
BETA SSVEP	8,960 → 8,960	100.0%	8,960 → 8,960	100.0%	8,960 → 8,960	100.0%
Monitoring ErrP	4,313 → 320	7.4%	4,291 → 320	7.5%	4,291 → 320	7.5%
EEGMAT	1,371 → 1,371	100.0%	1,380 → 1,380	100.0%	1,372 → 1,372	100.0%
LEMON gender	152,792 → 6,480	4.2%	152,706 → 6,480	4.2%	152,618 → 6,480	4.2%
LEMON extraversion	152,792 → 6,480	4.2%	152,706 → 6,480	4.2%	152,618 → 6,480	4.2%
FACED	8,232 → 8,232	100.0%	8,232 → 8,232	100.0%	8,232 → 8,232	100.0%
SEED	11,916 → 1,440	12.1%	11,916 → 1,440	12.1%	11,916 → 1,440	12.1%
SEED-IV	30,060 → 1,920	6.4%	30,060 → 1,920	6.4%	30,060 → 1,920	6.4%

Continued on the next page.

Table 41. Training windows before and after the per-subject/class cap (continued).

	Seed42		Seed10		Seed5	
Task	Before → after	Retained	Before → after	Retained	Before → after	Retained
SEED-V	95,667 → 2,600	2.7%	95,667 → 2,600	2.7%	95,667 → 2,600	2.7%
SEED-VII	21,872 → 4,480	20.5%	21,872 → 4,480	20.5%	21,872 → 4,480	20.5%
ISRUC S2	10,446 → 1,200	11.5%	10,446 → 1,200	11.5%	10,581 → 1,200	11.3%
ISRUC S3	7,058 → 1,600	22.7%	6,901 → 1,600	23.2%	6,910 → 1,600	23.2%
SleepEDF	330,638 → 22,790	6.9%	331,030 → 22,863	6.9%	330,374 → 22,861	6.9%
HMC	110,367 → 23,691	21.5%	110,139 → 23,488	21.3%	109,665 → 23,715	21.6%
TUAB	324,769 → 76,000	23.4%	327,890 → 76,280	23.3%	328,390 → 76,320	23.2%
TUEV	87,239 → 16,409	18.8%	94,090 → 16,946	18.0%	93,048 → 16,691	17.9%
TUSL	201 → 201	100.0%	200 → 200	100.0%	218 → 218	100.0%
EEGDenoiseNet	10,809 → 80	0.7%	10,809 → 80	0.7%	10,809 → 80	0.7%
Total	1,397,300 → 191,164	13.7%	1,406,812 → 191,859	13.6%	1,405,361 → 191,879	13.7%

The cap retains at most 40 windows per registered subject–class unit from each training split; validation and test use all available windows. Retention is the after/before ratio. Totals count task entries within each seed, including separate LEMON tasks; seeds reuse overlapping windows.

I Full-FT Duration–Capacity Response

We fit the shared response to the 12 Full-FT observations in Figure 9: four fractions each for Molecule-B, Atom-B, and Atom-L. Each score averages the same 23 tasks and three downstream seeds at the end of a 50-epoch run. Subset hours are estimated as the retained fraction times the collection duration.

In Equation 1, the shared response is $g(u) = 2u/(1+u^2) = [(u+u^{-1})/2]^{-1}$. It approaches $2u$ for small u and $2/u$ for large u, with a maximum at $u = 1$. We use this bounded, log-symmetric shape to describe the measured duration range. Both tails approach the series-specific level A_i.

There are six fitted parameters: three levels A_i, one common amplitude B, and two peak durations. Weighted least squares profiles out the levels and amplitude while searching for the peaks. We set the peak-search bounds to 1,000–1,000,000 hours for Base and 24,814–49,628 hours for Large. The size exponent is then $\eta = \log(H_L/H_B)/\log(46.3/5.8)$.

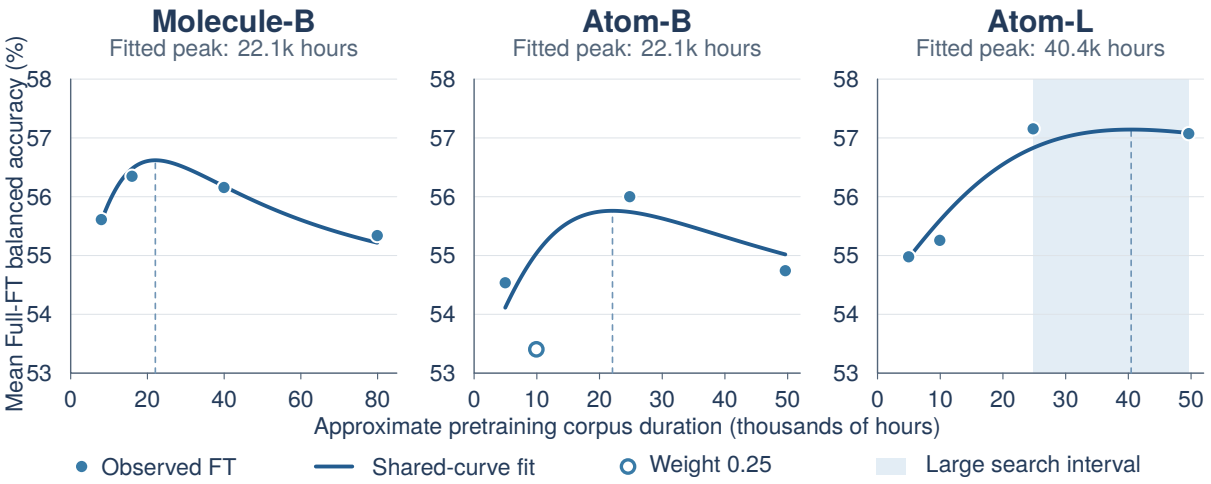


Figure 15. Shared Full-FT response with Atom-B's 20% point assigned weight 0.25. All 12 observations are shown; the hollow point receives reduced weight. Curves share the response shape and amplitude and are drawn only over observed duration ranges. Dashed lines mark fitted peaks; shading marks the imposed Large peak-search interval.

Table 42. Sensitivity to the weight of Atom-B's 20% point. All other points have unit weight. Peaks are in thousands of hours; leave-one-point-out (LOO) RMSE uses all 12 held-out scores, in percentage points.

Weight	Base peak	Large peak	η	LOO RMSE	Constant LOO
1	25.22	40.81	0.232	0.919	1.094
0.25	22.07	40.45	0.292	0.707	1.076
0	20.45	40.09	0.324	0.680	1.080

The displayed fit gives $K = 22{,}065.77$ hours, $\eta = 0.291689$, and $B = 2.8808$ percentage points. The fitted levels A_i are 52.8795, 54.2612, and 53.7392 for Atom-B, Atom-L, and Molecule-B. We use weight 0.25 for this fit and report weights 1 and 0 for sensitivity analysis. The three seeds measure downstream variability for each pretrained checkpoint.

I.1 Response-shape sensitivity and the inverse relation

The peak estimates depend on how the FT response is modeled. For comparison, we fit $S_i = a_i - b_i(u-1)^2$ and $S_i = a_i - b_i[\log u]^2$ with separate levels and amplitudes for each series and the same peak-search bounds (Table 43). These alternatives have eight parameters for 12 observations. The raw-duration quadratic assumes symmetric declines per hour; the log-duration quadratic assumes symmetry per multiplicative change in duration.

Table 43. Alternative response shapes with series-specific amplitudes. Atom-B's 20% point receives the indicated weight. Peak durations are in thousands of hours.

Response shape	Weight	Base peak	Large peak	η
Quadratic in duration	1	31.64	37.62	0.083
	0.25	29.40	37.62	0.119
	0	28.27	37.62	0.138
Quadratic in log duration	1	24.29	49.63†	0.344
	0.25	21.92	49.63†	0.393
	0	16.80	49.63†	0.521

†The fitted Large peak reaches the upper search bound.

An independent raw-duration quadratic places Molecule-B's peak at 37,473 hours, compared with 15,982 hours for its best measured score. Response shape thus has a substantial effect on the inferred peak. The bounded response in Equation 1 gives an interior Large peak for all three point weights.

Sharing the amplitude lowers LOO error within the bounded-response family. At weight 0.25, allowing a separate amplitude for each of the three bounded curves gives a LOO RMSE of 1.259 points, compared with 0.707 for the common-amplitude model. Sharing only between Atom-B and Atom-L gives 0.879.

For each LOO prediction, we remove one observation and refit all parameters. The constant comparison fits one level per series with the same training weights and gives an RMSE of 1.076 points. We use LOO for internal model selection, with the same observations guiding the response shape and amplitude-sharing choices. Endpoint omissions require extrapolation; keeping only held-out 20% and 50% points gives 0.880 points for the shared response and 1.217 for the constant comparison.

The inverse of the fitted peak relation is

$$P_{\text{ref}}(H) = 5.8 \times 10^6 \left(\frac{H}{K}\right)^{1/\eta}, \qquad \frac{d \log P_{\text{ref}}}{d \log H} = \frac{1}{\eta}. \tag{3}$$

The fitted inverse exponent is 3.4283; Equation 2 uses $1/0.292$ after rounding. Thus a relative change in duration produces a larger relative change in the corresponding parameter estimate. Here $P_{\text{ref}}(H)$ identifies the encoder size whose fitted peak occurs at duration H. Selecting a size for maximum FT performance at fixed duration additionally requires estimating how the series levels A_i vary with capacity and collection.

The exponent interpolates the two fitted peak locations for Base and Large. The resulting empirical scaling hypothesis describes FT under the current data collections and 50-epoch schedules, which couple corpus duration to optimization updates. Additional model sizes, intermediate fractions, and matched training exposure would test how this relation generalizes. LP follows a different fraction-dependent response and is analyzed separately in Section 6.3.