BRAINPFN: AMORTIZED BRAIN-BEHAVIOUR PREDICTION VIA PRIOR-DATA FITTED NETWORKS

Anonymous authors
Paper under double-blind review

ABSTRACT

The standard approach to predicting behavioral phenotypes from functional MRI is represented by Kernel Ridge Regression (KRR) applied on functional connectivity (FC) matrices. However, this approach is limited: separate models must be fit and tuned for every score and cohort, they do not transfer across datasets, and accuracy degrades sharply when only a handful of subjects are available, as is typical of clinical or cognitive studies. Despite an increasing interest in developing brain foundation models, none has consistently outperformed KRR on behaviour prediction; most reported gains come from easier targets such as age, sex, or diagnosis rather than continuous behavioural scores. We introduce BrainPFN, a Prior Data Fitted Network for brain behaviour regression. BrainPFN is meta-trained on a domain-specific prior built from ~7k real FC matrices, from which we sample one billion synthetic training samples. At inference, a labelled context is simply passed through the network in a single forward pass, with no per-score or per-cohort refitting required. Across four independent cohorts, BrainPFN beats KRR on small sample prediction, and leads recent foundation models pre-trained at four to nine times larger scale. Our meta-training pool, aggregated from open repositories, is a fraction of the size of the corpora behind competing models, most of which sit behind restrictive data use agreements such as UK Biobank. We release this pool publicly alongside code and checkpoints, showing that a modest, openly available pool can match or beat models trained at Biobank scale.

1 INTRODUCTION

Predicting behavioral and cognitive phenotypes from functional MRI (fMRI) is a central goal of cognitive neuroimaging, with applications ranging from mapping individual differences in cognition to identifying biomarkers of psychiatric and neurological disease. The dominant approach fits a Kernel Ridge Regression (KRR) model on functional connectivity (FC) matrices for each target score and cohort (Ooi et al., 2022). This baseline is simple and hard to beat (He et al., 2020), but comes with a structural cost: a new model must be fit and tuned from scratch for every phenotype and dataset, and its accuracy degrades sharply when only a handful of labeled subjects are available (Ooi et al., 2025). This is precisely the regime of most clinical and cognitive studies (Szucs & Ioannidis, 2020), where cohorts number in the tens or low hundreds rather than thousands, as shown in Fig. 1.

A growing line of work has sought to replace this per-dataset fitting with brain foundation models (Dong et al., 2024; Thomas et al., 2022; Gijsen et al., 2026; Caro et al., 2024; Yang et al., 2024; Dong et al., 2026), pre-trained once on large fMRI corpora and then adapted, in principle cheaply, to new tasks. Despite this effort, no such model consistently outperforms a simple KRR baseline in continuous behavioral prediction (Marraffini et al., 2026): most reported gains are concentrated on coarser targets such as age, sex, or categorical diagnosis, and evaluations on genuine behavioral scores in the small-sample regime remain scarce. The result is a puzzling gap between the promise of large-scale pre-training and its demonstrated benefit for the settings where a transferable prior should matter most: small, heterogeneous, hard-to-pool cohorts.

We breach the fMRI-cognitive prediction gap with *Prior-data Fitted Networks* (PFNs, Müller et al., 2022), a model class trained to perform Bayesian inference in a single forward pass. PFNs are meta-trained on synthetic datasets drawn from a task-specific prior, rather than on a fixed pre-training corpus. This design replaces per-dataset optimization with amortized inference: at test time, a

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

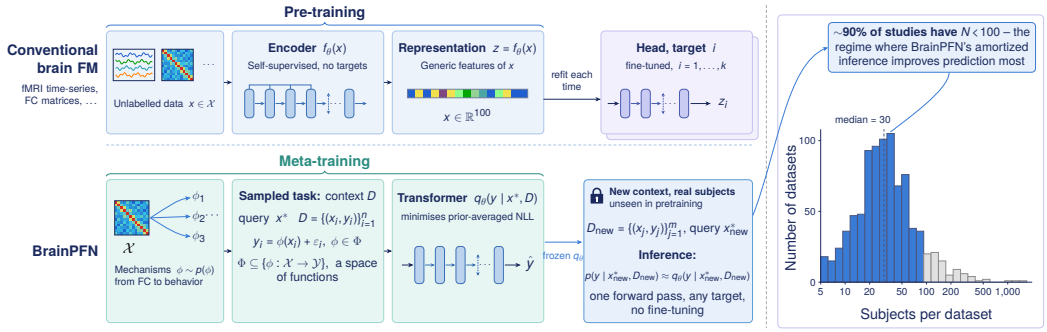


Figure 1: **BrainPFN overview:** BrainPFN replaces the per-target, per-cohort refitting of conventional brain foundation models with a single frozen forward pass. Conventional brain FMs learn a generic representation via self-supervision, then require fitting a separate prediction head for each target. BrainPFN instead meta-trains a transformer q_θ entirely on synthetic tasks generated from a prior $\phi \in \Phi$ over brain-behavior to directly approximate the posterior predictive. At deployment, the same frozen q_θ is given a new context from real, previously unseen subjects and predicts any target in one forward pass, with no retraining. The right panel reports the distribution of cohort size across ~1000 neuroimaging datasets on OpenNeuro with fMRI data for at least 5 subjects: ~90% have fewer than 100 subjects (highlighted in blue), precisely the small-cohort regime where BrainPFN's amortized inference improves prediction most.

labeled context of any size is simply passed through the network alongside query samples, with no gradient steps, refitting, or hyperparameter search required. We instantiate this idea for brain-behavior prediction by constructing a domain-specific prior directly from a modest pool of ~7,000 real FC matrices, and meta-training on one billion synthetic regression problems sampled from it. Notably, our meta-training step is fully self-supervised and does not require any real behavioral label.

Our contribution is three-fold:

- We introduce *BrainPFN*, a Prior-data Fitted Network for brain-behavior prediction that amortizes inference across scores and cohorts in a single forward pass. This approach excels in the small-sample regime where existing approaches struggle most.
- We show that BrainPFN consistently outperforms KRR and existing brain foundation models – several of them pre-trained on corpora four to nine times larger – in low-data settings, while remaining competitive with the strongest baselines as more labeled data becomes available.
- We publicly release the meta-training pool underlying BrainPFN, a fraction of the size of, and free of the restrictive data-use agreements attached to, corpora such as UK Biobank, together with code and checkpoints, demonstrating that a small, openly available pool can match or exceed models trained at UK BioBank scale.

2 Related Work

Classical brain-behaviour prediction and the small-sample regime. Connectome-based predictive modelling predicts individual scores from functional connectivity (FC) with per-score regularized regression (Finn et al., 2015; Rosenberg et al., 2016; Shen et al., 2017), and systematic benchmarks find simple linear models remarkably competitive on such data (Dadi et al., 2019; He et al., 2020; Schulz et al., 2020). Recent studies on fMRI foundation models show evidence that larger datasets can improve downstream prediction (Wang et al., 2026), however most studies do not typically focus on prediction on small cohorts. This matters practically because most clinical and cognitive neuroimaging studies never reach the large-N regime: highly-cited fMRI studies have a median sample size around a dozen participants (Szucs & Ioannidis, 2020). With standard approaches, this results in underpowered designs that inflate and destabilize reported effects (Button et al., 2013), whereas brain-wide association studies require samples in the thousands to reproduce

reliably (Marek et al., 2022): a scale most single-site and, especially, clinical acquisitions never reach, where limited sample size is repeatedly identified as the main bottleneck to single-subject prediction (Arbabshirani et al., 2017). A predictive model whose accuracy is concentrated at small N would therefore target the regime where most neuroimaging and clinical cohorts actually sit.

Foundation models for brain imaging. Foundation models pre-trained on large fMRI corpora include BrainLM (Caro et al., 2024), BrainMass (Yang et al., 2024) (masked modelling on FC matrices across 30 datasets, ~70k samples), Brain-JEPA (Dong et al., 2024) and Brain-Semantoks (Gijsen et al., 2026) on ROI time series. These are typically deployed as frozen encoders with a per-target linear probe; Sec. 5 compares BrainPFN against three of them under that protocol. All four pre-train, in whole or in part, on UK Biobank (Bycroft et al., 2018): Brain-JEPA on ~32k and Brain-Semantoks on ~39k UKB scans, BrainLM on a further ~61k, and BrainMass's 30-dataset pool includes it alongside HCP and Cam-CAN. BrainPFN's meta-training pool aggregates ~7.7k FC matrices from public OpenNeuro releases (Markiewicz et al., 2021) with no access agreement, 4–9× smaller than the corpora above, and we release it publicly alongside the model and code.

Prior-data Fitted Networks and In-Context Learning. PFNs (Müller et al., 2022) have been popularized in the tabular domain, with TabPFN (Hollmann et al., 2022; 2025) and TabICL (QU et al., 2025), with strong results in the small-data regime. Beyond tabular data, PFNs have been applied to time-series forecasting (Dooley et al., 2023), Bayesian optimization (Müller et al., 2023), learning-curve extrapolation (Adriaensen et al., 2023), causal effect estimation (Robertson et al., 2026; Balazadeh Meresht et al., 2026) and genomics (Ubbens et al., 2025), and in-context learners have also reached neuroimaging (Czolbe & Dalca, 2023; Yu et al., 2025). In general, PFNs rely on fully synthetic priors: both inputs and targets come from randomly sampled generative processes such as structural causal models. Our prior is instead grounded in real data. We do not sample FC matrices; we draw them from a meta-training pool of real connectomes.

3 Method: BrainPFN

3.1 Amortized inference for Brain-Behavior Prediction

We represent each subject by their resting-state FC matrix parcellated with the Schaefer-400 atlas (Schaefer et al., 2018). FC matrices are symmetric positive-definite (SPD) matrices, thus we map them to log-Cholesky coordinates (Lin, 2019), which respects the curved geometry of the SPD manifold. The log-Cholesky representation is then reduced by PCA to $d = 100$ components, yielding feature vectors $x \in \mathbb{R}^d$. The target $y \in \mathbb{R}$ is a continuous behavioral score. Given a labelled training set $D_{\text{train}} = (X_{\text{train}}, y_{\text{train}})$, where $X_{train} \in \mathbb{R}^{n \times d}$ is composed of n FC feature vectors, and a new test subject x, the goal is to predict y without refitting any model parameters, i.e. purely in-context. The training set D_{train} is also called *context*.

Suppose the data in D_{train} are generated by some unknown mechanism $\phi\colon \mathcal{X} \to \mathbb{R}$, drawn from a prior $\phi \in \Phi \sim p(\phi)$ that encodes everything we know about how FC relates to behaviour. Each mechanism induces a conditional $p(y \mid x, \phi)$. Given D_{train}, we seek the predictor $q(y \mid x)$ that minimizes the Bayes risk under log loss:

$$\mathcal{R}(q) = \mathbb{E}_{\phi \sim p(\phi \mid D_{\text{train}})}\, \mathbb{E}_{y \sim p(y \mid x, \phi)} \big[-\log q(y \mid x) \big] = -\int p(y \mid x, D_{\text{train}}) \log q(y \mid x)\, dy\,. \tag{1}$$

This is cross-entropy, so by Gibbs' inequality it is minimised by the posterior predictive:

$$p(y \mid x, D_{\text{train}}) = \int_{\Phi} p(y \mid x, \phi)\, p(\phi \mid D_{\text{train}})\, d\phi\,. \tag{2}$$

A Bayes-optimal predictor therefore integrates over *every* possible mechanism weighted by how well it explains the context. This integral is generally intractable, but it can be approximated with a Prior-Fitted Network (PFN, Müller et al., 2022), a Transformer trained to approximate:

$$q^* := \arg\min_q\ \mathbb{E}_{\phi \sim p(\phi)}\, \mathbb{E}_{D_{\text{train}},\, x,\, y \sim p(\cdot \mid \phi)} \big[-\log q(y \mid x, D_{\text{train}}) \big] \underset{\text{PFN}}{=} \int_{\Phi} p(y \mid x, \phi)\, p(\phi \mid D_{\text{train}})\, d\phi\,, \tag{3}$$

Approximating Eq. 3 is called *amortized inference*. In practice, at each training step a mechanism ϕ is sampled, and a synthetic dataset $D = \{(x_i, y_i)\}_{i=1}^{n}$ is drawn from it. Specifically, once we sampled ϕ, we produce $y_i = \phi(x_i)$ where the feature vectors x_i come from a set of real FC matrices, i.e. the *meta-training pool*. D is then split into a context D_{train} and a held-out test point $(x_{\text{test}}, y_{\text{test}})$. Taking this into account the network is trained by minimizing $L_\phi(\theta)$:

$$L_\phi(\theta) = \mathbb{E}_{D\sim p(D|\phi)} \left[-\log q_\theta(y_{\text{test}} \mid x_{\text{test}}, D_{\text{train}}) \right] . \tag{4}$$

Because a different ϕ is sampled at each step, a sufficiently expressive network trained in this way approximates q_θ, a parametric minimizer for q^* of Eq. 3, as $L(\theta) = \mathbb{E}_{\phi\sim p(\phi)} L_\phi(\theta)$ without ever having to compute the posterior predictive explicitly (Müller et al., 2022).

The key choice in this framework is therefore how to build the prior $p(\phi)$, which in our case must capture the structure of FC-behaviour associations.

3.2 Brain-Behavior Association Prior

Regularised kernel methods, and kernel ridge regression (KRR) in particular, remain among the strongest and most reproducible predictors of behavior from FC data, matching or exceeding deep learning architectures on this exact task (He et al., 2020; Shen et al., 2017; Dadi et al., 2019; Ooi et al., 2022). In KRR, predictions take the form $\hat{y} = K(x, X_{\text{train}})\,\alpha$, where K is a kernel matrix computed from FC vectors, $\alpha = (K(X_{\text{train}}, X_{\text{train}}) + \lambda I)^{-1} y_{\text{train}}$ are the dual coefficients, and λ is a regularisation parameter. BrainPFN is accordingly built to approximate the posterior predictive distribution over the space of KRR functions mapping FC to behavior, with prior $p(\phi)$ defined over mechanisms $\phi = (K, \alpha, \sigma_\epsilon^2)$ parametrising the kernel choice, the dual coefficients α, and the noise $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$. For each synthetic task, we randomly select a set of FC vectors X_{train} from our meta-training pool and generate synthetic continuous targets as:

$$y = K(X_{\text{train}}, X_{\text{train}})\,\alpha + \epsilon, \qquad (K, \alpha, \sigma_\epsilon^2) = \phi \sim p(\phi), \tag{5}$$

where the kernel K is sampled from the linear, RBF, and Laplacian families, together with the kernel coefficients. A held-out test set X_{test} is drawn from the same pool, the targets $y_{\text{test}} = K(X_{\text{test}}, X_{\text{train}})\,\alpha + \epsilon$ are generated from the same mechanism ϕ, and the network q_θ is trained to predict y_{test} from $(x_{\text{test}}, D_{\text{train}})$ by minimising the loss in Eq. 4. We draw X_{train} and X_{test} from real FC matrices (the meta-training pool) rather than simulated ones, as the statistical structure of functional connectivity is not trivial to capture with generative models. It is worth noting that for a single kernel with fixed hyperparameters the posterior predictive of Eq. 3 is available in closed form (App. C.2). In our prior, however, the kernel bandwidth is itself random so the exact posterior predictive is a mixture of GP posteriors weighted by their marginal likelihoods, with no closed form. The PFN learns to perform this averaging in-context.

Sampling dual coefficients. In the dual formulation of KRR, predictions take the form $f(x) = K(x, X_{\text{train}})\,\alpha$, and the regularisation term in the KRR objective penalises the squared reproducing kernel Hilbert space (RKHS) norm of f: $\|f\|_{\mathcal{H}}^2 = \alpha^\top K_{\text{train}}\,\alpha = \sum_j \hat{f}_j^2/\lambda_j$ where λ_j and v_j are the eigenvalues and eigenvectors of K_{train}, and $\hat{f}_j = \lambda_j\, v_j^\top \alpha$ is the prediction component along v_j. The RKHS norm thus penalises directions with small eigenvalues: the same prediction amplitude $\hat{f}_j$ incurs a cost $1/\lambda_j$ times larger in low-variance directions. To sample dual coefficients that respect this geometry we draw them from:

$$\alpha \sim \mathcal{N}\big(0,\, \sigma_\alpha^2 K_{\text{train}}^{-1}\big)\,, \tag{6}$$

which concentrates the sampled functions along the high-eigenvalue (smooth) directions of the kernel. This ties their effective complexity to the spectrum of the FC manifold rather than producing arbitrarily rough targets. We show this outperforms an isotropic prior in Sec. 5.4.

Task difficulty and calibration. Brain-behavior associations vary widely in strength across cognitive domains and datasets. We model behavior as $y = s + \epsilon$ (where s is the KRR prediction), meaning that some part is explained by the actual FC data and some part is aleatoric or epistemic noise. By varying the amount of noise ϵ, we can control the strength of the correlation between FC and behavior $r = \mathrm{Corr}(s, y)$, which represents the task difficulty. With $\epsilon \sim \mathcal{N}(0, \sigma^2)$ independent of s and $\mathrm{Var}(s) = 1$, a short calculation gives (full derivation in Sec. A.1):

$$r = \mathrm{Corr}(s, y) = \frac{\mathrm{Cov}(s,\, s+\epsilon)}{\sqrt{\mathrm{Var}(s)\,\mathrm{Var}(y)}} = \frac{1}{\sqrt{1+\sigma^2}}, \tag{7}$$

so we can solve for $\sigma_\epsilon^2 = (1 - r^2)/r^2$, and the observed target is:

$$y = s + \epsilon, \qquad \epsilon \sim \mathcal{N}\left(0, \frac{1 - r^2}{r^2}\right). \tag{8}$$

To avoid circularity, the raw signal $s = K(X, X_{\text{train}})\,\alpha$ must be standardized to unit variance using subjects not involved in fitting it. Reusing the context or test block to estimate this variance would bias it differently across kernel families, so we draw a third, disjoint block of subjects for this purpose only. We therefore draw a third block of subjects from the pool, disjoint from both the context and the test points, and use it only to estimate the scale of s. As a result, r acts as a single, kernel-agnostic difficulty knob across the whole prior.

4 Experimental Setup

4.1 Datasets

Meta-training pool BrainPFN is meta-trained entirely on ~7.7k real resting-state FC matrices from ~4.9k unique subjects across 300 public OpenNeuro studies (Markiewicz et al., 2021) (full list in Sec. E). No dataset in the pool requires a data-use agreement. Every subject is processed with an identical fmriprep-based pipeline (full preprocessing parameters in App. B.1) and parcellated with the Schaefer-400 atlas (Schaefer et al., 2018). We collect no behavioral scores for this pool: BrainPFN is meta-trained purely on synthetic regression tasks built from real FC features.

Downstream datasets We evaluate cross-dataset transfer of BrainPFN on four independent cohorts spanning different age ranges, acquisition sites, and scanning protocols, processed with the same FC pipeline as the meta-training pool: HCP Young Adult (HCP-YA, Van Essen et al., 2013), HCP-Aging (HCP-A, Bookheimer et al., 2019), Cam-CAN (Shafto et al., 2014; Taylor et al., 2017), and AOMIC-ID1000 (Snoek et al., 2021), none of which overlaps with the meta-training pool (Tab. 1). Target scores span cognition, personality, emotion, physical, and demographic (age, sex) measures, drawn from each dataset's standard behavioural battery.

Table 1: **Evaluation datasets** used to benchmark BrainPFN generalization to external cohorts. # scores is the number of target behavioral measures evaluated per cohort.

Dataset	Train / Test	Age range	Acquisition	# scores
HCP-YA	1029	22–37 (healthy young adult)	resting-state	62
→ *Task*	1008		task-based	2
HCP-Aging	1198	36–90+ (healthy older adult)	resting-state	35
Cam-CAN	597	18–88 (lifespan)	resting-state	3
AOMIC-ID1000	881	19–26 (general population)	movie-watching	14

4.2 Implementation Details

BrainPFN architecture and meta-training. BrainPFN uses the asymmetric-attention Transformer of TabPFN (Hollmann et al., 2022) with a single linear regression head in place of the original softmax classifier. We report two presets: *BrainPFN-0.8M* (4 layers, d_{model}=128) and *BrainPFN-4.7M* (6 layers, d_{model}=256); full architectural and optimization hyperparameters are given in App. B.1. The model is trained for 10k steps on ~1B synthetic context/test (x, y) pairs sampled online from the meta-training pool, using a Laplace kernel and a fixed task difficulty $r = 0.5$ by default (App. B.1); Sec. 5.4 ablates both choices.

Evaluation protocol. For each dataset we perform 5-fold cross-validation, repeatedly sampling 20 contexts of size N from each split's train fold; BrainPFN (only its PCA basis refit per context) and the Ridge/KRR baseline are evaluated on the same held-out test fold via the Pearson correlation r between predictions and ground truth, at context sizes $N \in \{5, 10, 20, \dots\}$ up to the train-fold size. We summarize each learning curve with the log-area under the learning curve, **log-AULC**, computed via the trapezoidal rule, so that a higher value indicates a more data-efficient model across the whole

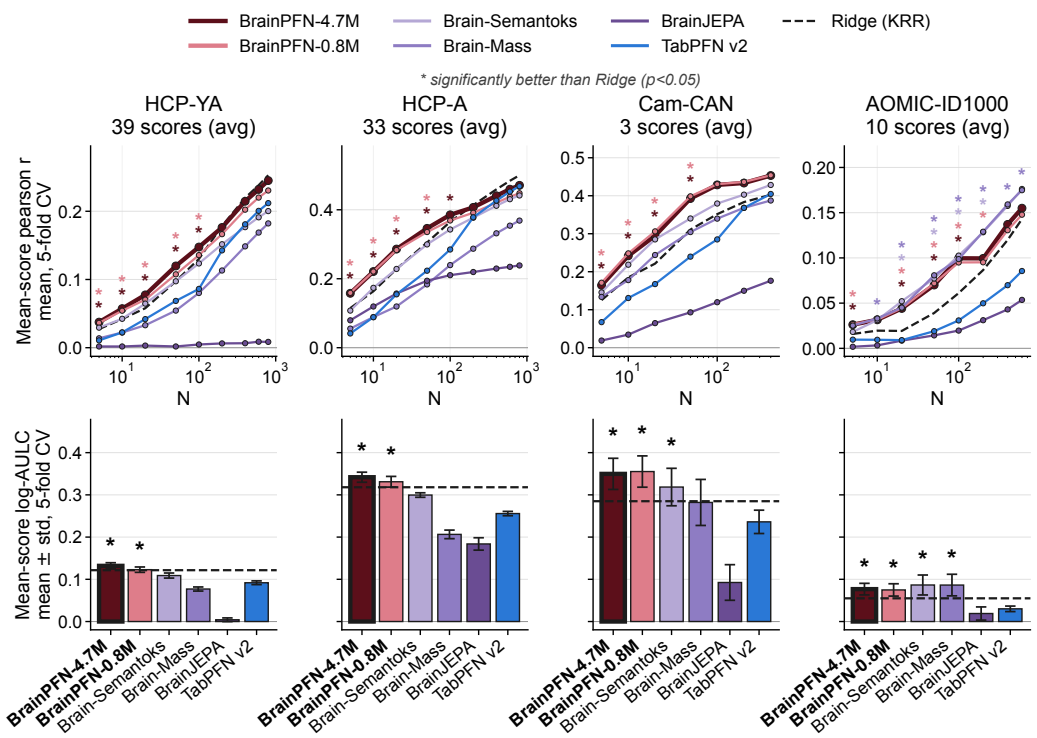


Figure 2: **Grand average across all behavioral tasks**. On a total of 85 tasks, BrainPFN beats Ridge and leads other foundation models on brain-behavior association. Top: mean Pearson r vs. context size N; bottom: area under the learning curve (log-AULC). BrainPFN is the only model consistently above Ridge (dashed; * $p < 0.05$) on all four cohorts.

N range rather than at one arbitrarily chosen size. We restrict analysis to *real-signal* targets, those decodable with at least $r \geq 0.1$ by any of the models, to avoid diluting results with near-random-level targets. Full per-score results, significance-testing, and evaluation-grid details are given in App. D.2 and App. B.1.

5 Results

5.1 BrainPFN leads the small N regime

Fig. 2 summarizes brain-behaviour prediction results across all four downstream cohorts on all 85 real-signal targets, plotting the learning curve of mean Pearson r against context size N (top) and its log-AULC summary (bottom) against Ridge and four existing foundation models: Brain-JEPA, Brain-Mass, Brain-Semantoks, and the generic-prior TabPFN v2. We restrict this comparison to *real-signal* targets, those decodable with at least $r \geq 0.1$ by any of the compared models, so as not to dilute the average with near-random-level targets; results on the full, unfiltered set of targets are reported in App. D. As a baseline we use linear-kernel KRR on log-Cholesky features, PCA with 800 components, and $\alpha = 1$. This configuration was found to be the best during preliminary sweeps.

In the low-data regime this work targets, BrainPFN is the only model that beats Ridge consistently: both BrainPFN-4.7M and BrainPFN-0.8M are significantly above the Ridge baseline at $N \leq 20$ across all four cohorts (Fig. 2, top, asterisks), and this advantage carries through to the log-AULC summary, where BrainPFN is the only model to significantly outperform Ridge on HCP-YA, HCP-A, and Cam-CAN. None of the other foundation models matches this consistency. Brain-Semantoks is the strongest alternative, tracking Ridge closely and occasionally exceeding it, while Brain-Mass is competitive on Cam-CAN and AOMIC-ID1000 but noticeably weaker elsewhere; notably, Cam-CAN was part of Brain-Mass's pretraining corpus. Cognition is the behavioral domain where brain foundation models have so far struggled most to improve on KRR (He et al., 2020), so we examine

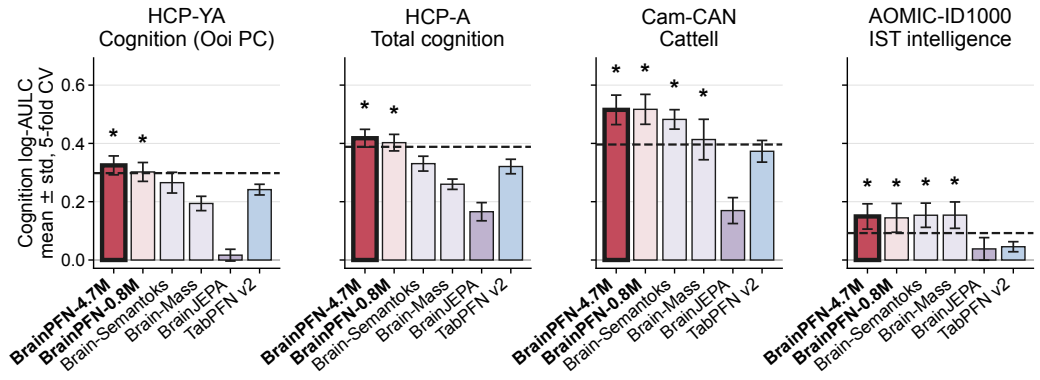


Figure 3: **BrainPFN beats ridge and leads other foundation models on cognition prediction.** Log-AULC (mean ± std, 5-fold CV; dashed: Ridge; *: $p < 0.05$ vs. Ridge). BrainPFN is the only model significantly above Ridge on all four cohorts; other foundation models mostly fall below it.

it separately in Fig. 3, using one general-cognition score per cohort: a cognition component in the style of Ooi et al. (2022) on HCP-YA (more details in App. C.1), the NIH Toolbox total cognition composite on HCP-A, the Cattell fluid-intelligence score on Cam-CAN, and the IST intelligence score on AOMIC-ID1000. Both BrainPFN presets are the only models whose log-AULC is significantly above Ridge on all four cohorts. On HCP-YA and HCP-A, every competing foundation model falls below the Ridge baseline, with Brain-JEPA close to chance on HCP-YA.

The curves and log-AULC summaries in Fig. 2 average over 85 real-signal targets in total. Hence, we check whether BrainPFN's average reflects a broad advantage or is carried by a few favourable scores. Fig. 4 reports, per dataset and context size N, how many targets BrainPFN predicts more accurately than Ridge: BrainPFN wins on nearly every target on all four datasets up to N=100, the win rate then drops between N=100 and N=200 on HCP-YA and HCP-A specifically, and it never drops on Cam-CAN or AOMIC-ID1000. Cam-CAN is the one exception to the pattern above: BrainPFN's lead over Ridge does not shrink to zero as N grows, and it is the one cohort where BrainPFN also stays ahead of the strongest competing foundation model at full context.

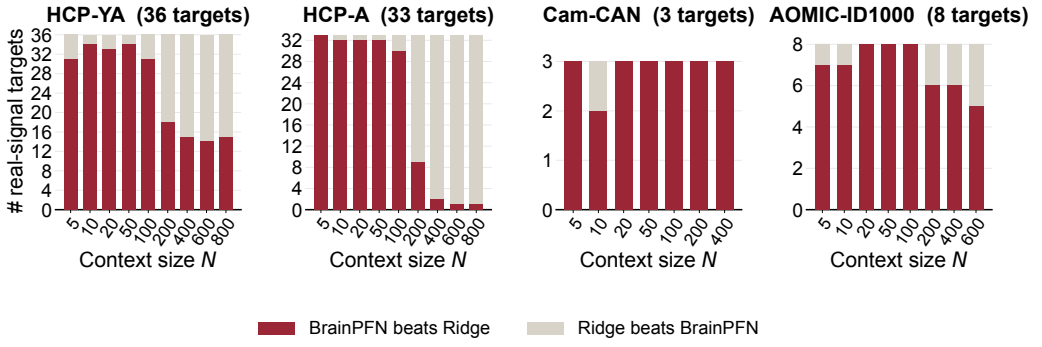


Figure 4: **Win-rate of BrainPFN versus Ridge** at each context size N. Bar height is the number of targets tested. BrainPFN's small-N advantage is not carried by a few favorable targets: it wins on nearly every target on small contexts.

BrainPFN's advantage is largest in the small-N regime it was designed for. As N grows, Ridge gradually closes the gap and overtakes BrainPFN between $N \approx 200$, several times the median OpenNeuro cohort size (Fig. 1), most visibly on HCP-A. On Cam-CAN and AOMIC-ID1000, no target crosses over within the available sample size. The same holds for every other foundation model we tested, none of which keeps a consistent advantage over Ridge at large N. BrainPFN remains competitive in this regime, while being the only model that consistently beats Ridge where data are scarce.

5.2 Existing Brain Foundation Models fall short

Impact of the domain-specific prior. BrainPFN and TabPFN are the same kind of model: an in-context learner built on the same Transformer backbone. They differ in what each one is trained to predict: BrainPFN's synthetic KRR-over-real-FC prior versus TabPFN's generic, domain-agnostic prior. This leaves the prior itself as the source of BrainPFN's advantage. This evidence is further supported by the RKHS sampling ablation in Sec. 5.4, which removes the RKHS structure from the prior over dual coefficients and produces the largest degradation of any ablation tested. Furthermore, simply training on domain-specific data is not by itself sufficient for good few-shot behavioral prediction: Brain-JEPA is the weakest model in every cohort despite being pretrained on fMRI data from more than 40k subjects, under-performing even the generic-prior TabPFN, which has seen no brain data at all. Although to a lesser extent, similar observations can also be made for the other foundation models. We read this as supporting context for the prior argument above: self-supervised fMRI pretraining and BrainPFN's explicit synthetic-function prior are different mechanisms, and Brain-JEPA's weak showing only demonstrates that the former is not sufficient on its own.

Model capacity, saturation, and transferability. We find that model size is not predictive of performance: even though BrainPFN-4.7M gives a small, consistent edge over BrainPFN-0.8M, both models have far fewer parameters than the other foundation models. A compact model trained on a limited real-data footprint, augmented into ~1B synthetic training samples, is enough to learn a prior that outperforms models with substantially larger parameter counts (BrainJEPA has ~86M parameters, Brain-Semantoks ~62M, and BrainMass ~10M). We speculate that this gap could also reflect these larger models' pretraining corpora not scaling with their parameter count. Furthermore, behavioural scores are defined differently across cohorts – different instruments and scales, and sometimes different underlying constructs even when the name is similar (Sec. 4.1), so a per-score model like Ridge must be refit from scratch for every phenotype and every cohort, with nothing carried over from one dataset to the next. BrainPFN, in contrast, is used as-is, without retraining or refitting, on all four downstream cohorts and every target reported in this section.

5.3 Generalization beyond resting-state FC

Beyond the results on naturalistic (movie-watching) fMRI from AOMIC-ID1000 (Figs. 2 and 3), where BrainPFN beats Ridge despite being pre-trained only on resting-state data, here we investigate whether this advantage also carries over to task-based fMRI. Using the HCP-YA dataset, we test whether BrainPFN's small-context advantage extends beyond resting-state functional connectivity by predicting in-scanner working-memory (WM) task performance from FC estimated on the task fMRI run itself, compared against FC estimated from the resting-state run of the same subjects. Tab. 2 reports 2-back and 0-back accuracy predicted from both FC sources. The small-context pattern from Sec. 5.1 reproduces on both FC sources: BrainPFN beats Ridge on small contexts, and Ridge overtakes at larger N. Task FC is also simply a better predictor of task performance than resting FC, for either model, at every context size tested (Table 2). This is consistent with neuroscientific evidence in favor of task-based fMRI for behavioral prediction (Zhao et al., 2023).

5.4 Ablations

We perform an ablation study on HCP-Aging in terms of log-AULC to assess the contribution of key design choices in BrainPFN (Fig. 5); App. C.3 extends the same comparison on the other three downstream cohorts. **Kernel:** sampling randomly or restricting the kernel family to linear or RBF kernels instead of the Laplace kernel likewise underperforms the best configuration. **Task-difficulty:** randomly sampling r seems to slightly under-perform a fixed r, but not always significantly. **RKHS and log-Cholesky:** consistent with Sec. 5.2, removing the RKHS prior on the α coefficients ("no RKHS") produces the largest degradation of all ablations, and working directly in SPD space instead of using the log-Cholesky parameterization also reduces performance, though to a lesser extent. These are the two most impactful design choices tested. **Model size:** reducing model size from 4.7M to 0.8M parameters results in a modest decrease in performance. Hence, our preferred model is 4.7M because the computational overhead is minimal.

Table 2: **BrainPFN's small-context advantage carries over to task fMRI.** The table reports working-memory accuracy (2-back, 0-back) predicted from resting-state or task FC on the same HCP-YA subjects, as mean Pearson r ± std over 5 test folds. BrainPFN beats Ridge up to N = 100 on resting-state FC and up to N = 200 on task FC.

	`WM_Task_2bk_Acc` (2-back)				`WM_Task_0bk_Acc` (0-back)			
	Resting-state FC		Task FC		Resting-state FC		Task FC	
N	BrainPFN	Ridge	BrainPFN	Ridge	BrainPFN	Ridge	BrainPFN	Ridge
5	**0.047±0.009**	0.029±0.015	**0.104±0.039**	0.081±0.042	**0.037±0.034**	0.023±0.033	**0.113±0.064**	0.083±0.052
10	**0.091±0.017**	0.055±0.018	**0.196±0.043**	0.140±0.027	**0.050±0.028**	0.025±0.022	**0.171±0.029**	0.103±0.033
20	**0.123±0.031**	0.091±0.024	**0.230±0.019**	0.162±0.014	**0.084±0.028**	0.061±0.034	**0.229±0.037**	0.148±0.043
50	**0.167±0.038**	0.131±0.026	**0.312±0.077**	0.241±0.066	**0.118±0.041**	0.096±0.038	**0.294±0.061**	0.223±0.059
100	**0.222±0.052**	0.200±0.047	**0.347±0.083**	0.298±0.073	**0.146±0.016**	0.129±0.034	**0.325±0.063**	0.268±0.066
200	0.265±0.057	**0.269±0.047**	**0.357±0.075**	0.352±0.076	0.182±0.050	**0.188±0.056**	**0.347±0.086**	0.324±0.083
400	0.314±0.049	**0.331±0.038**	0.389±0.088	**0.409±0.090**	0.212±0.062	**0.223±0.074**	0.372±0.082	**0.375±0.078**
600	0.338±0.054	**0.362±0.037**	0.415±0.093	**0.441±0.093**	0.225±0.062	**0.240±0.078**	0.389±0.080	**0.399±0.068**
800	0.344±0.055	**0.375±0.035**	0.429±0.089	**0.461±0.093**	0.235±0.060	**0.256±0.087**	0.405±0.077	**0.421±0.065**

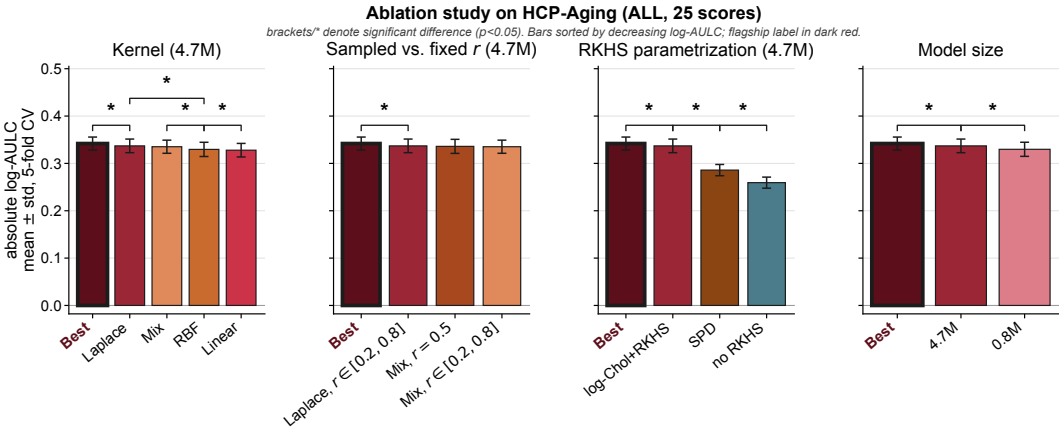


Figure 5: **Ablation study on HCP-Aging.** BrainPFN-4.7M (Laplace kernel family, fixed task difficulty) is the best configuration in every score group and significantly outperforms all variants. Removing the RKHS-sampling or working directly in SPD space causes the largest drops. *Best* denotes BrainPFN-4.7M (Laplace, r=0.5).

6 Conclusions

We presented BrainPFN, a Prior-data Fitted Network that recasts brain-behavior prediction as amortized inference: Meta-trained once on synthetic KRR tasks built on ∼7.7k real, unlabeled FC matrices from OpenNeuro, it predicts any continuous target from a labeled context in a single forward pass, with no refitting or hyperparameter search. Across four independent cohorts and 85 targets, BrainPFN is the only model out of all the brain foundation models we compared that consistently beats KRR in the small-sample regime. That includes general cognition in every cohort. The advantage extends to movie-watching and task FC, even if BrainPFN was meta-trained using only resting-state data. BrainPFN closes for the first time, at least on small cohorts, the gap between the promise of pre-trained brain foundation models and simple linear baselines. Still, on large cohorts Ridge maintains the lead. Future work will focus on designing richer prior to extend BrainPFN's lead to larger N, other parcellations, and clinical cohorts. Notably, our meta-training pool is a fraction of the size of UK BioBank data, typically used to train brain foundation models, and free of the access restrictions attached to it. By releasing this pool alongside code and checkpoints, we hope to lower the barrier to building and evaluating foundation models for neuroimaging, and to encourage a shift in how the field benchmarks such models: toward the low-data, high-heterogeneity settings where they are needed most.

AI USE STATEMENT

In this work, we used generative AI tools for: propose or refine hypothesis, design or provide feedback on research methodology or experiments, implement methods, clean and reformat dataset, interpret results. We have not used generative AI tools for: generating synthetic dataset, help develop theoretical models or conceptual frameworks, formulate mathematical claims, provide critical ingredients for proving mathematical claims, assisting in the writing of proofs, support qualitative and thematic data analysis. Additionally, we used generative AI tools for: creating or modifying scientific figures of images, suggesting experimental parameters, create or edit software code, creation of artifacts, drafting parts of a research paper, summarize or analyze existing literature, discover research topics or identify gaps, brainstorming, sourcing/searching for information, edit a research paper to improve readability, identify relevant literature, suggest a structure for a research paper. We have reviewed all AI-assisted work. We have not used AI to generate references, they were hand copied from the original sources. AI-generated code was verified and tested for correctness. We take responsibility for the final content of this work, including text, claims or artifacts produced with the aid of generative AI.

ETHICS STATEMENT

This work uses recordings of human brain activity (resting-state and task fMRI) to predict behavioral and cognitive phenotypes. All data are either fully public or obtained under standard data use agreements (DUAs) from their original repositories; no new human data were collected, and we defer to the consent and IRB protocols of the original collectors. Our meta-training pool draws entirely from unrestricted OpenNeuro releases, and our downstream cohorts (HCP-YA, HCP-Aging, Cam-CAN, AOMIC-ID1000) are accessed under their standard DUAs.

Our aim is to advance cognitive and clinical neuroscience by making accurate brain-behavior prediction feasible in the small-sample regime typical of real research and clinical cohorts, ultimately improving our understanding of the brain and its diseases and supporting better diagnosis and treatment response.

We recognize a potential dual use of this technology: discriminating against or profiling individuals based on predicted cognitive ability or other traits, e.g., in employment, insurance, or education. We do not condone this use. BrainPFN is intended strictly for research and clinical purposes, and we release our pool, code, and checkpoints on that basis, encouraging downstream users to restrict use to appropriately consented research and clinical contexts.

REPRODUCIBILITY STATEMENT

BrainPFN's meta-training is fully reproducible: the meta-training pool is built entirely from public OpenNeuro data with no access restrictions, and we release the full source code, preprocessing pipeline, and model checkpoints, all of which will be made public upon acceptance. Evaluation on the four downstream cohorts (HCP-YA, HCP-Aging, Cam-CAN, AOMIC-ID1000) is likewise reproducible, provided the user agrees to the standard data use agreement (DUA) required by each dataset's distributor. Full details of the preprocessing pipeline, architecture, and training hyperparameters needed to reproduce our results are given in App. B.1, with dataset and evaluation-grid details in App. D.2.

REFERENCES

Steven Adriaensen, Herilalaina Rakotoarison, Samuel Müller, and Frank Hutter. Efficient bayesian learning curve extrapolation using prior-data fitted networks. *Advances in Neural Information Processing Systems*, 36:19858–19886, 2023.

Mohammad R Arbabshirani, Sergey Plis, Jing Sui, and Vince D Calhoun. Single subject prediction of brain disorders in neuroimaging: Promises and pitfalls. *Neuroimage*, 145:137–165, 2017.

Vahid Balazadeh Meresht, Hamidreza Kamkari, Valentin Thomas, Junwei Ma, Bingru Li, Jesse Cresswell, and Rahul Krishnan. Causalpfn: Amortized causal effect estimation via in-context learning. *Advances in Neural Information Processing Systems*, 38:154945–154984, 2026.

Susan Y Bookheimer, David H Salat, Melissa Terpstra, Beau M Ances, Deanna M Barch, Randy L Buckner, Gregory C Burgess, Sandra W Curtiss, Mirella Diaz-Santos, Jennifer Stine Elam, et al. The lifespan human connectome project in aging: an overview. *Neuroimage*, 185:335–348, 2019.

Katherine S Button, John PA Ioannidis, Claire Mokrysz, Brian A Nosek, Jonathan Flint, Emma SJ Robinson, and Marcus R Munafò. Power failure: why small sample size undermines the reliability of neuroscience. *Nature reviews neuroscience*, 14(5):365–376, 2013.

Clare Bycroft, Colin Freeman, Desislava Petkova, Gavin Band, Lloyd T Elliott, Kevin Sharp, Allan Motyer, Damjan Vukcevic, Olivier Delaneau, Jared O'Connell, et al. The uk biobank resource with deep phenotyping and genomic data. *Nature*, 562(7726):203–209, 2018.

Josue Ortega Caro, Antonio Henrique de Oliveira Fonseca, Syed A Rizvi, Matteo Rosati, Christopher Averill, James L Cross, Prateek Mittal, Emanuele Zappala, Rahul Madhav Dhodapkar, Chadi Abdallah, and David van Dijk. BrainLM: A foundation model for brain activity recordings. In *The Twelfth International Conference on Learning Representations*, 2024. URL https://openreview.net/forum?id=RwI7ZEfR27.

Steffen Czolbe and Adrian V. Dalca. Neuralizer: General neuroimage analysis without re-training. In *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 6217–6230, 2023. doi: 10.1109/CVPR52729.2023.00602.

Kamalaker Dadi, Mehdi Rahim, Alexandre Abraham, Darya Chyzhyk, Michael Milham, Bertrand Thirion, Gaël Varoquaux, Alzheimer's Disease Neuroimaging Initiative, et al. Benchmarking functional connectome-based predictive models for resting-state fmri. *NeuroImage*, 192:115–134, 2019.

Zijian Dong, Ruilin Li, Yilei Wu, Thuan T Nguyen, Joanna S Chong, Fang Ji, Nathanael R Tong, Christopher L Chen, and Juan H Zhou. Brain-jepa: Brain dynamics foundation model with gradient positioning and spatiotemporal masking. *Advances in Neural Information Processing Systems*, 37:86048–86073, 2024.

Zijian Dong, Ruilin Li, Joanna Chong, Niousha Dehestani, Yinghui Teng, Yi Lin, Zhizhou Li, Yichi Zhang, Yapei Xie, Leon Ooi, et al. Brain harmony: a multimodal foundation model unifying morphology and function into 1d tokens. *Advances in Neural Information Processing Systems*, 38:122100–122127, 2026.

Samuel Dooley, Gurnoor Singh Khurana, Chirag Mohapatra, Siddartha V Naidu, and Colin White. Forecastpfn: Synthetically-trained zero-shot forecasting. *Advances in Neural Information Processing Systems*, 36:2403–2426, 2023.

Emily S Finn, Xilin Shen, Dustin Scheinost, Monica D Rosenberg, Jessica Huang, Marvin M Chun, Xenophon Papademetris, and R Todd Constable. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nature neuroscience*, 18(11):1664–1671, 2015.

Sam Gijsen, Marc-Andre Schulz, and Kerstin Ritter. Brain-semantoks: Learning semantic tokens of brain dynamics with a self-distilled foundation model. In *International Conference on Learning Representations*, volume 2026, pp. 13186–13208, 2026.

Tong He, Ru Kong, Avram J Holmes, Minh Nguyen, Mert R Sabuncu, Simon B Eickhoff, Danilo Bzdok, Jiashi Feng, and BT Thomas Yeo. Deep neural networks and kernel regression achieve comparable accuracies for functional connectivity prediction of behavior and demographics. *NeuroImage*, 206:116276, 2020.

Noah Hollmann, Samuel Müller, Katharina Eggensperger, and Frank Hutter. TabPFN: A transformer that solves small tabular classification problems in a second. In *NeurIPS 2022 First Table Representation Workshop*, 2022. URL https://openreview.net/forum?id=eu9fVjVasr4.

Noah Hollmann, Samuel Müller, Lennart Purucker, Arjun Krishnakumar, Max Körfer, Shi Bin Hoo, Robin Tibor Schirrmeister, and Frank Hutter. Accurate predictions on small data with a tabular foundation model. *Nature*, 637(8045):319–326, 2025.

Zhenhua Lin. Riemannian geometry of symmetric positive definite matrices via cholesky decomposition. *SIAM Journal on Matrix Analysis and Applications*, 40(4):1353–1370, 2019.

Scott Marek, Brenden Tervo-Clemmens, Finnegan J Calabro, David F Montez, Benjamin P Kay, Alexander S Hatoum, Meghan Rose Donohue, William Foran, Ryland L Miller, Timothy J Hendrickson, et al. Reproducible brain-wide association studies require thousands of individuals. *Nature*, 603(7902):654–660, 2022.

Christopher J Markiewicz, Krzysztof J Gorgolewski, Franklin Feingold, Ross Blair, Yaroslav O Halchenko, Eric Miller, Nell Hardcastle, Joe Wexler, Oscar Esteban, Mathias Goncavles, et al. The openneuro resource for sharing of neuroscience data. *Elife*, 10:e71774, 2021.

Giovanni Marraffini, Gabriel Mahuas, Trinidad Borrell, Victoria Shevchenko, and Demian Wassermann. The variance brain foundation models forgot: Third-order statistics predict cognition where billion-parameter models fail. *arXiv preprint arXiv:2606.04010*, 2026.

Samuel Müller, Noah Hollmann, Sebastian Pineda Arango, Josif Grabocka, and Frank Hutter. Transformers can do bayesian inference. In *International Conference on Learning Representations*, 2022. URL https://openreview.net/forum?id=KSugKcbNf9.

Samuel Müller, Matthias Feurer, Noah Hollmann, and Frank Hutter. Pfns4bo: In-context learning for bayesian optimization. In *International Conference on Machine Learning*, pp. 25444–25470. PMLR, 2023.

Leon Qi Rong Ooi, Jianzhong Chen, Shaoshi Zhang, Ru Kong, Angela Tam, Jingwei Li, Elvisha Dhamala, Juan Helen Zhou, Avram J Holmes, and BT Thomas Yeo. Comparison of individualized behavioral predictions across anatomical, diffusion and functional connectivity mri. *NeuroImage*, 263:119636, 2022.

Leon Qi Rong Ooi, Csaba Orban, Shaoshi Zhang, Thomas E Nichols, Trevor Wei Kiat Tan, Ru Kong, Scott Marek, Nico UF Dosenbach, Timothy O Laumann, Evan M Gordon, et al. Longer scans boost prediction and cut costs in brain-wide association studies. *Nature*, 644(8077):731–740, 2025.

Jingang QU, David Holzmüller, Gaël Varoquaux, and Marine Le Morvan. TabICL: A tabular foundation model for in-context learning on large data. In *Forty-second International Conference on Machine Learning*, 2025. URL https://openreview.net/forum?id=0VvD1PmNzM.

Jake Robertson, Arik Reuter, Siyuan Guo, Noah Hollmann, Frank Hutter, and Bernhard Schölkopf. Do-pfn: In-context learning for causal effect estimation. *Advances in Neural Information Processing Systems*, 38:174811–174848, 2026.

Monica D Rosenberg, Emily S Finn, Dustin Scheinost, Xenophon Papademetris, Xilin Shen, R Todd Constable, and Marvin M Chun. A neuromarker of sustained attention from whole-brain functional connectivity. *Nature neuroscience*, 19(1):165–171, 2016.

Alexander Schaefer, Ru Kong, Evan M Gordon, Timothy O Laumann, Xi-Nian Zuo, Avram J Holmes, Simon B Eickhoff, and BT Thomas Yeo. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity mri. *Cerebral cortex*, 28(9):3095–3114, 2018.

Marc-Andre Schulz, BT Thomas Yeo, Joshua T Vogelstein, Janaina Mourao-Miranada, Jakob N Kather, Konrad Kording, Blake Richards, and Danilo Bzdok. Different scaling of linear models and deep learning in ukbiobank brain images versus machine-learning datasets. *Nature communications*, 11(1):4238, 2020.

Meredith A Shafto, Lorraine K Tyler, Marie Dixon, Jason R Taylor, James B Rowe, Rhodri Cusack, Andrew J Calder, William D Marslen-Wilson, John Duncan, Tim Dalgleish, et al. The cambridge centre for ageing and neuroscience (cam-can) study protocol: a cross-sectional, lifespan, multidisciplinary examination of healthy cognitive ageing. *BMC neurology*, 14(1):204, 2014.

Xilin Shen, Emily S Finn, Dustin Scheinost, Monica D Rosenberg, Marvin M Chun, Xenophon Papademetris, and R Todd Constable. Using connectome-based predictive modeling to predict individual behavior from brain connectivity. *nature protocols*, 12(3):506–518, 2017.

Lukas Snoek, Maite M van der Miesen, Tinka Beemsterboer, Andries Van Der Leij, Annemarie Eigenhuis, and H Steven Scholte. The amsterdam open mri collection, a set of multimodal mri datasets for individual difference analyses. *Scientific data*, 8(1):85, 2021.

Denes Szucs and John PA Ioannidis. Sample size evolution in neuroimaging research: An evaluation of highly-cited studies (1990–2012) and of latest practices (2017–2018) in high-impact journals. *NeuroImage*, 221:117164, 2020.

Jason R Taylor, Nitin Williams, Rhodri Cusack, Tibor Auer, Meredith A Shafto, Marie Dixon, Lorraine K Tyler, Richard N Henson, et al. The cambridge centre for ageing and neuroscience (cam-can) data repository: Structural and functional mri, meg, and cognitive data from a cross-sectional adult lifespan sample. *neuroimage*, 144:262–269, 2017.

Armin Thomas, Christopher Ré, and Russell Poldrack. Self-supervised learning of brain dynamics from broad neuroimaging data. *Advances in neural information processing systems*, 35:21255–21269, 2022.

Jordan Ubbens, Ian Stavness, and Andrew G. Sharpe. Gpfn: Prior-data fitted networks for genomic prediction. *IEEE Transactions on Computational Biology and Bioinformatics*, 22(6):2642–2649, 2025. doi: 10.1109/TCBBIO.2025.3596744.

David C Van Essen, Stephen M Smith, Deanna M Barch, Timothy EJ Behrens, Essa Yacoub, Kamil Ugurbil, Wu-Minn HCP Consortium, et al. The wu-minn human connectome project: an overview. *Neuroimage*, 80:62–79, 2013.

Mo Wang, Wenhao Ye, Junfeng Xia, Junxiang Zhang, Xuanye Pan, Minghao Xu, Haotian Deng, Hongkai Wen, and Quanying Liu. Omni-fMRI: A universal atlas-free fMRI foundation model. In *Forty-third International Conference on Machine Learning*, 2026. URL https://openreview.net/forum?id=Rc8th2rXXe.

Yanwu Yang, Chenfei Ye, Guinan Su, Ziyao Zhang, Zhikai Chang, Hairui Chen, Piu Chan, Yue Yu, and Ting Ma. Brainmass: Advancing brain network analysis for diagnosis with large-scale self-supervised learning. *IEEE transactions on medical imaging*, 43(11):4004–4016, 2024.

Muquan Yu, Mu Nan, Hossein Adeli, Jacob S. Prince, John A. Pyles, Leila Wehbe, Margaret Marie Henderson, Michael J. Tarr, and Andrew Luo. Meta-learning an in-context transformer model of human higher visual cortex. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*, 2025. URL https://openreview.net/forum?id=B3iPTZh7Za.

Weiqi Zhao, Carolina Makowski, Donald J Hagler, Hugh P Garavan, Wesley K Thompson, Deanna J Greene, Terry L Jernigan, and Anders M Dale. Task fmri paradigms may capture more behaviorally relevant information than resting-state functional connectivity. *NeuroImage*, 270:119946, 2023.

A Derivations

A.1 Task difficulty

We model the observed behavioral target as $y = s + \epsilon$, where $s = K(X, X_{\text{train}})\,\alpha$ is the KRR signal (standardised to unit variance, $\mathrm{Var}(s) = 1$) and $\epsilon \sim \mathcal{N}(0, \sigma^2)$ is additive noise independent of s. We derive the noise variance σ^2 required to achieve a target correlation $r = \mathrm{Corr}(s, y)$.

Covariance. Since $\epsilon \perp s$:

$$\begin{aligned}\mathrm{Cov}(s, y) &= \mathrm{Cov}(s,\ s + \epsilon)\\ &= \mathbb{E}[s(s+\epsilon)] - \mathbb{E}[s]\,\mathbb{E}[s+\epsilon]\\ &= \big(\mathbb{E}[s^2] - \mathbb{E}[s]^2\big) + \big(\mathbb{E}[s\epsilon] - \mathbb{E}[s]\,\mathbb{E}[\epsilon]\big)\\ &= \mathrm{Var}(s) + \mathrm{Cov}(s, \epsilon)\\ &= 1 + 0 = 1. \end{aligned} \tag{9}$$

Variance of the target. Since s and ϵ are independent:

$$\mathrm{Var}(y) = \mathrm{Var}(s + \epsilon) = \mathrm{Var}(s) + \mathrm{Var}(\epsilon) = 1 + \sigma^2. \tag{10}$$

Correlation.

$$r = \mathrm{Corr}(s, y) = \frac{\mathrm{Cov}(s, y)}{\sqrt{\mathrm{Var}(s)\,\mathrm{Var}(y)}} = \frac{1}{\sqrt{1 \cdot (1 + \sigma^2)}} = \frac{1}{\sqrt{1 + \sigma^2}}. \tag{11}$$

Solving for σ^2. Setting the above equal to a target effect size r and solving:

$$\frac{1}{\sqrt{1 + \sigma^2}} = r \implies 1 + \sigma^2 = \frac{1}{r^2} \implies \sigma^2 = \frac{1 - r^2}{r^2}, \tag{12}$$

which gives Eq. (7).

B Additional Details

B.1 Additional experimental setup details

Meta-training pool preprocessing. Every subject in the meta-training pool (Sec. 4.1) is processed with the same pipeline based on fmriprep (25.2.3): resampling of volumes to the MNI152 template at 2mm isotropic resolution, followed by scrubbing of high-motion volumes (framewise displacement $>$ 0.5mm), 24-parameter confound regression including motion and tissue intensity parameters, and a 0.01–0.1Hz bandpass filter, before parcellation with the Schaefer-400 atlas (Schaefer et al., 2018).

Architecture and optimization hyperparameters. Training tokens concatenate x_i and the scalar label y_i before a linear projection; test tokens project x_j alone. Features and labels are normalised within each forward pass using statistics computed from the training tokens and de-normalised at output. *BrainPFN-0.8M* uses 4 layers, d_{model}=128, 4 heads, d_{ff}=256, and *BrainPFN-4.7M* uses 6 layers, d_{model}=256, 4 heads, d_{ff}=512; both parameter counts exclude the input projection layers, whose size scales with the (data-dependent) feature dimensionality d. In preliminary sweeps, further increasing depth and width gave no accuracy improvement over BrainPFN-4.7M, so we do not report larger presets. The model is trained for 10k steps with AdamW (learning rate 3×10^{-4}, weight decay 10^{-2}, no warmup), batch size 512, and dropout 0.1. Each synthetic task in a batch draws 200 context subjects, a disjoint calibration block of 200 subjects (Eq. 7), and 100 held-out subjects from the meta-training pool, i.e. 500 subjects per synthetic task; over the full run this amounts to $10^4 \times 512 \times 300 \approx 1.5 \times 10^9$ individual context/test (x, y) pairs seen by the network – the ~1B figure reported in the abstract. Training uses a single V100 or A100 GPU.

Significance testing and evaluation grid. The other foundation models are evaluated at each context by linear probing on the same context, rather than by refitting Ridge/KRR. Significance against the Ridge baseline is computed on the per-fold averages: we first average across the 20 within-fold context draws, then compute significance across the 5 resulting fold-level values. The context-size grid $N \in \{5, 10, 20, \dots\}$ extends up to each cohort's train-fold size (Tab. 1).

B.2 Data preprocessing for the other foundation models

BrainJEPA. Input is the raw resting-state BOLD timeseries (not a connectivity matrix), parcellated into 450 regions (400 cortical + 50 subcortical) and reordered to match BrainJEPA's own pretraining atlas convention. Each region's signal is robust-scaled – $(x - \text{median})/\text{IQR}$ – using reference statistics computed once from the training-fold subjects and then frozen and reused identically across every dataset (never refit per dataset). The timeseries is cut into non-overlapping 160-timepoint windows (fixed by the pretrained model's positional embeddings); every window from every run of a subject is passed through the frozen pretrained encoder and the resulting window-level embeddings are averaged into a single 768-dimensional subject embedding.

BrainMass. Input is a functional-connectivity (correlation) matrix computed from a 100-region parcellation, matching the resolution BrainMass was pretrained on (coarser than this project's own 400-region atlas). The frozen pretrained transformer encoder processes this matrix through its attention layers and a dimensionality-reduction step, yielding an 800-dimensional subject embedding (taken before the self-supervised pretraining projection head).

Brain-Semantoks. Input is again the raw BOLD timeseries, resampled to a fixed 2-second TR, over the same 450-region parcellation used for BrainJEPA plus a cerebellar atlas that this project's data does not cover – that channel is zero-filled and explicitly masked out inside the model so it contributes no signal. The timeseries is split into non-overlapping 100-timepoint windows, encoded by the frozen pretrained model, and window/run embeddings are averaged per subject into a 1536-dimensional embedding (a concatenation of pooled local tokens and a global summary token).

TabPFN. Unlike the three embedding models above, TabPFN is a general-purpose tabular foundation model with no neuroimaging pretraining. Input is the same raw functional-connectivity features used for BrainPFN. The PCA basis is refit separately for every few-shot training-set size evaluated, rather than fixed once – matching the same per-context convention used for the ridge and BrainPFN baselines, so TabPFN always sees a feature representation scaled appropriately to how much data it is being given.

C Additional Results

C.1 Ooi et al. (2022)-style behavioural components (HCP-YA and HCP-Aging)

In addition to the individual behavioural scores evaluated throughout the paper, we follow the dimensionality-reduction approach of Ooi et al. (2022) to construct a small set of interpretable, orthogonal behavioural components for HCP-YA and HCP-Aging, and evaluate every model on them as additional regression targets. This appendix gives the method, the resulting components, and the corresponding model comparison.

Method. For each cohort we start from its full pool of behavioural measures – 58 scores for HCP-YA, 29 for HCP-Aging, spanning cognition, motor, and emotion/well-being domains. Missing entries are imputed with an EM-style iterative imputer (scikit-learn's `IterativeImputer`, matching the EM imputation used by Ooi et al. (2022)), the imputed matrix is z-scored per column, and we fit PCA retaining a fixed number of components – 3 for HCP-YA, 6 for HCP-Aging, chosen to mirror the component counts reported for these cohorts in the original Ooi et al. analysis. The raw PCA loadings are then rotated with varimax to sharpen each component onto a smaller subset of scores, which is what makes the components individually interpretable (Figures 8 and 9 show, for each component, the behavioural scores with the largest rotated loadings). The rotated per-subject component scores are saved alongside the original behavioural scores and enter the benchmark as ordinary regression targets, one per component, evaluated with the exact same 5-fold, context-size-sweep protocol as every other score in this paper.

For HCP-YA, the three components separate into a *cognition* axis (dominated by working memory, arithmetic, reading, vocabulary, and fluid intelligence), an *emotion recognition* axis (dominated by task-based emotion-recognition accuracy), and a *dissatisfaction* axis (dominated by perceived stress, sadness, loneliness, and neuroticism, loading negatively on positive affect). For HCP-Aging, the six

components separate NIH Toolbox cognition into *fluid* and *crystallized* sub-factors, and add *motor* (grip strength, sex, walk endurance), *wellbeing* (psychological well-being and social satisfaction, loading negatively on negative affect and poor sleep), *memory* (RAVLT verbal learning/recall), and *processing speed* (Trail Making A/B) components, merged into the per-subject score tables the same way as HCP-YA's three components above.

Model comparison. Figures 6 and 7 show pearson r vs. context size N for all 7 models on every OOI component, and Tables 3 and 4 give the corresponding log-AULC summary (5-fold CV pooled, same protocol and significance test – one-sided paired t-test over the 5 fold-level log-AULC values – as the rest of the paper). BrainPFN is the most consistent winner on both cohorts: on HCP-Aging it beats Ridge on Fluid Cognition, Memory, and Processing Speed, and is competitive on Crystallized Cognition and Motor; on HCP-YA it beats Ridge on Cognition and Emotion Recognition. Wellbeing (HCP-Aging) and Dissatisfaction (HCP-YA) are the two components no model beats Ridge on – consistent with the rest of the paper's finding that affect/self-report targets are the hardest domain for FC-based prediction regardless of which model does the decoding. BrainJEPA is consistently the weakest model on both cohorts' OOI components, in line with its performance on the individual behavioural scores in Appendix D.1.

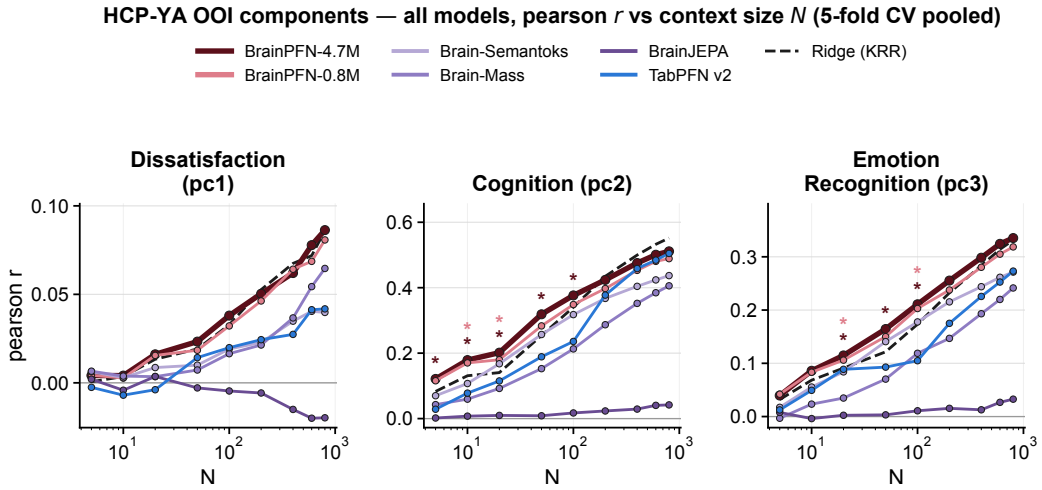


Figure 6: HCP-YA OOI components: pearson r vs. context size N for all 7 models (5-fold CV pooled). Dashed black = canonical Ridge reference. "*" marks N where a model is significantly ahead of its own Ridge baseline (one-sided paired t-test on the 5 fold-level means, Holm-corrected across that panel's N-grid).

Table 3: HCP-YA OOI components: log-AULC (Pearson r vs. context size N, 5-fold CV pooled). **Bold** = model beats the canonical Ridge row (one-sided paired t-test over 5 folds, $p < 0.05$).

Model	Dissatisfaction	Cognition	Emotion Recognition
Ridge	0.032	0.298	0.162
BrainPFN-4.7M	0.034	**0.325**	**0.186**
BrainPFN-0.8M	0.031	0.302	0.175
Brain-Semantoks	0.017	0.265	0.150
Brain-Mass	0.017	0.194	0.099
BrainJEPA	-0.005	0.017	0.009
TabPFN v2	0.014	0.241	0.124

Component loadings. Figures 8 and 9 show the top-7 rotated loadings (by $|\text{loading}|$) per component, reproduced from the saved PCA+varimax component loadings computed above.

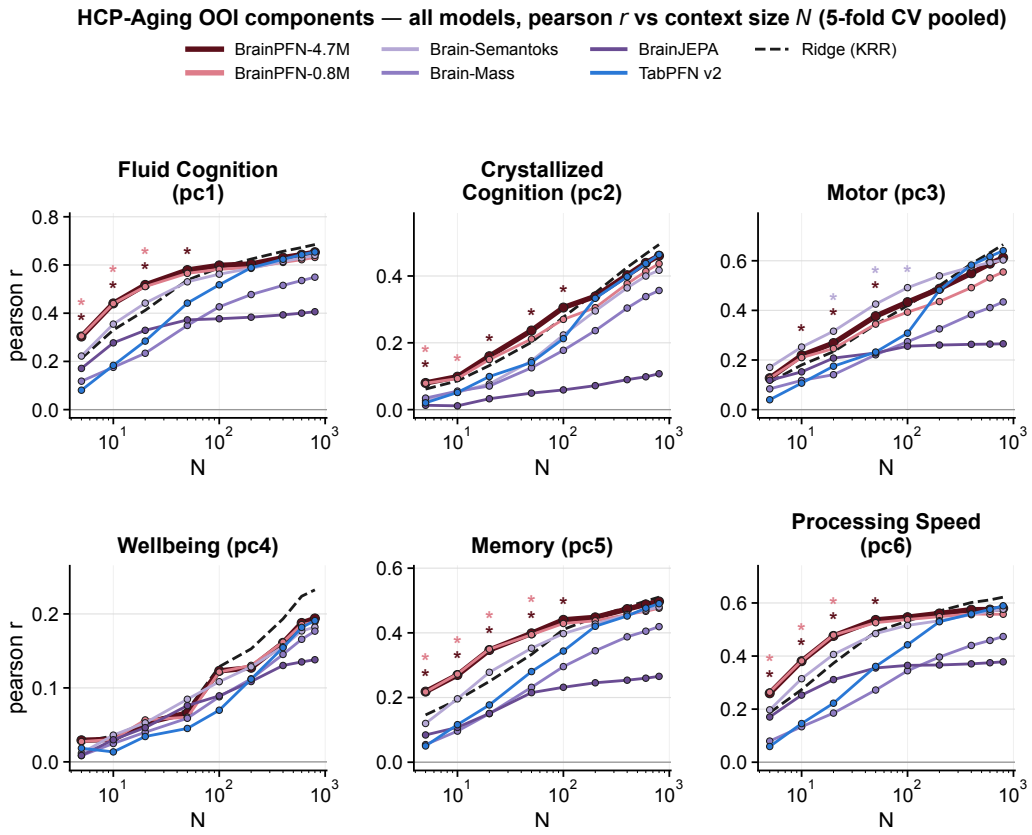


Figure 7: HCP-Aging OOI components: pearson r vs. context size N for all 7 models (5-fold CV pooled), same conventions as Figure 6.

Table 4: HCP-Aging OOI components: log-AULC (Pearson r vs. context size N, 5-fold CV pooled). **Bold** = model beats the canonical Ridge row (one-sided paired t-test over 5 folds, $p < 0.05$).

Model	Fluid Cognition	Crystallized Cognition	Motor	Wellbeing	Memory	Processing Speed
Ridge	0.512	0.246	0.375	0.107	0.347	0.462
BrainPFN-4.7M	**0.550**	**0.256**	0.384	0.095	**0.390**	**0.499**
BrainPFN-0.8M	**0.538**	0.235	0.349	0.095	**0.383**	**0.492**
Brain-Semantoks	0.502	0.194	**0.424**	0.094	0.343	0.456
Brain-Mass	0.356	0.164	0.243	0.079	0.246	0.290
BrainJEPA	0.346	0.053	0.223	0.078	0.196	0.328
TabPFN v2	0.427	0.207	0.313	0.075	0.292	0.366

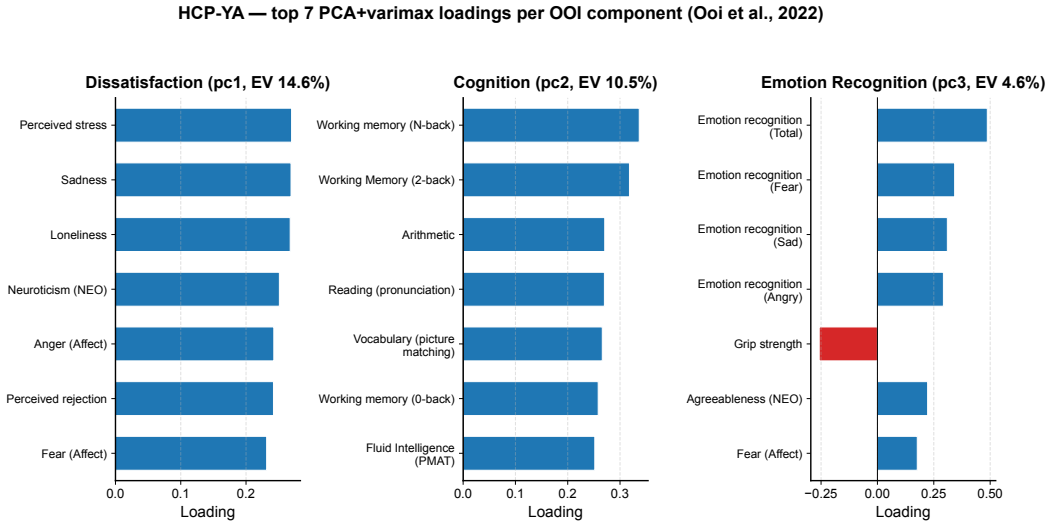


Figure 8: HCP-YA: top-7 PCA+varimax loadings per OOI component.

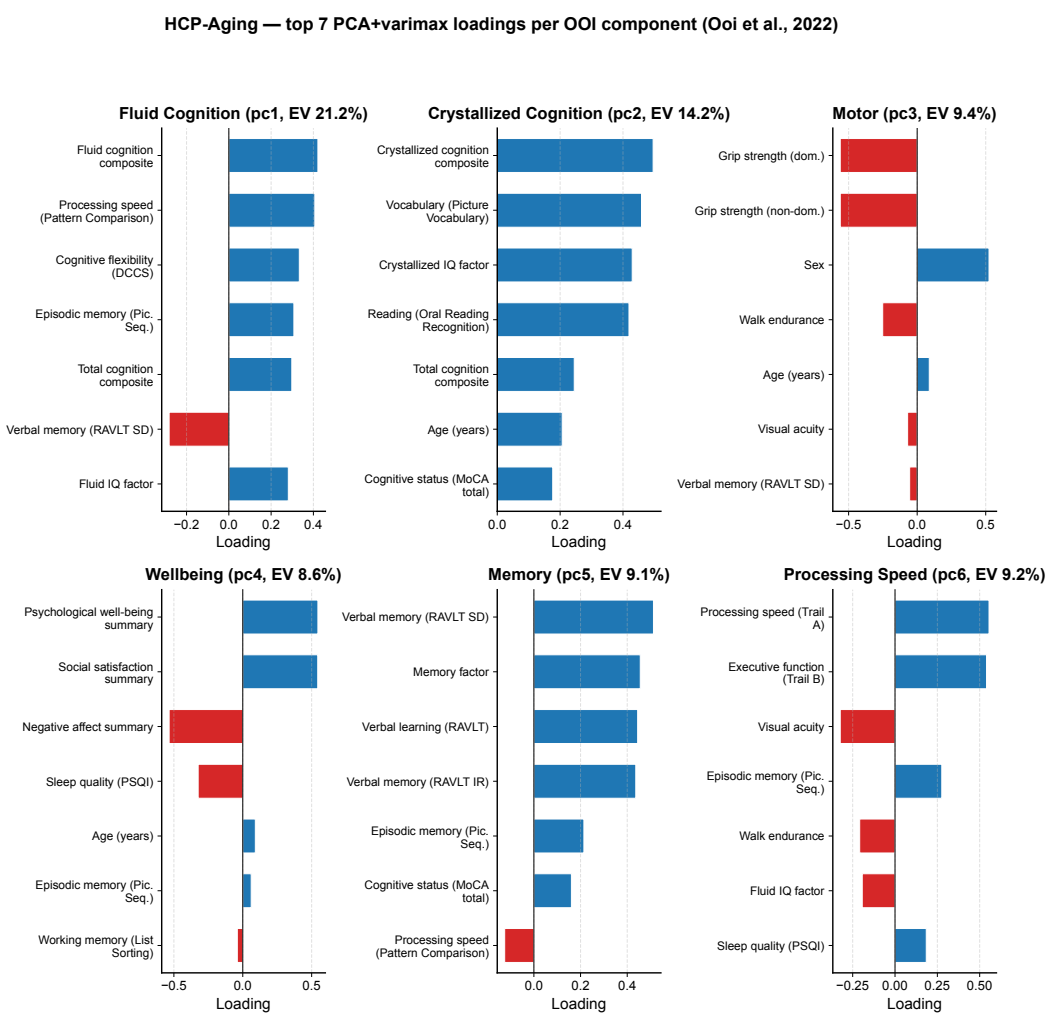


Figure 9: HCP-Aging: top-7 PCA+varimax loadings per OOI component.

C.2 EXACT GAUSSIAN-PROCESS BASELINE

BrainPFN is meta-trained on synthetic tasks whose targets are kernel functions of FC plus calibrated noise (Sec. 3.2). For a fixed kernel and fixed hyperparameters, the posterior predictive of such a task is that of a Gaussian process (GP) and is available in closed form; with random hyperparameters it is a mixture of GP posteriors weighted by their marginal likelihoods. A GP with the same kernel, fitted on the same features, is therefore the natural analytic counterpart of BrainPFN: it tests whether the network does more than reproduce this Bayesian computation.

GP baselines. Each target is standardised on the context set and modelled as $y = f(x) + \epsilon$, with $f \sim \mathcal{GP}(0, r^2 k_\gamma)$ and $\epsilon \sim \mathcal{N}(0, 1 - r^2)$, where r is the signal–target correlation (the task-difficulty parameter of the prior) and $k_\gamma(x, x') = \exp(-\gamma \|x - x'\|_1)$ is the prior's Laplace kernel, with bandwidth $\gamma = s/m$ and m the median pairwise ℓ_1 distance in the context. The inputs are exactly the features BrainPFN sees (log-Cholesky FC, with the PCA basis refit on each context). For fixed (s, r) the posterior mean is $\mu_* = r^2 K_* (r^2 K + (1 - r^2) I)^{-1} y$. The two baselines differ only in how they treat (s, r), both using the exact marginal likelihood $p(y \mid s, r)$ on a grid:

- **GP-BMA** (Bayesian model averaging) places a uniform hyperprior on $s \in [1, 3]$ and $r \in [0.2, 0.8]$ – the bandwidth range of the prior and the task-difficulty range of its sampled-r variant, which contains BrainPFN-4.7M's fixed r=0.5 – and averages the GP predictions over an 11×13 grid, weighted by their posterior probability. It integrates the hyperparameters out, which is the computation BrainPFN is trained to amortise, and is the strongest analytic competitor.
- **GP-ML** (type-II maximum likelihood) keeps the single (s, r) that maximises the marginal likelihood over a wider grid ($s \in [0.1, 10]$, $r \in [0.05, 0.95]$), the standard way GPs are fitted in practice.

We also report the ablation variant of BrainPFN-4.7M trained with $r \sim U[0.2, 0.8]$ instead of the fixed r=0.5 ("sampled r", Sec. 5.4). All models share the same folds, 20 context draws per N, and test splits; log-AULC is computed over $N \in \{5, \ldots, 800\}$, and significance is a two-sided paired t-test over the 5 fold-level values, reported as $p < 0.05$ or n.s. Real-signal targets are selected as in the rest of the paper, over the models compared here.

Results. BrainPFN-4.7M significantly beats both GPs over the whole learning curve in all three settings, by $+0.010$ to $+0.020$ log-AULC over GP-BMA (Tab. 5); the sampled-r variant is ahead of GP-BMA everywhere but significantly so only on HCP-A. The gain comes from the few-shot regime (Tab. 6): BrainPFN-4.7M is significantly ahead of GP-BMA at every N from 5 to 100 in all three settings. From N=200 the two are within 0.01 of each other in every setting (GP-BMA is slightly but significantly ahead on HCP-A for $N \geq 400$), and Ridge is the best model there, so the GP does not close BrainPFN's large-N gap to Ridge either. GP-ML has the lowest log-AULC of all models: with few subjects the marginal likelihood is nearly flat and its maximiser lands on degenerate corners of the grid, whereas averaging stays near the centre of the hyperprior.

Table 5: log-AULC of all models ($N \in \{5, \ldots, 800\}$, mean over 5 folds; bold: best model) and two-sided paired t-test across folds against GP-BMA and GP-ML: ↑/↓ marks the row model as significantly better/worse ($p < 0.05$), n.s. otherwise.

	HCP-YA, real-signal (36)			HCP-YA, cognition composite			HCP-A, real-signal (34)		
Model	log-AULC	vs. GP-BMA	vs. GP-ML	log-AULC	vs. GP-BMA	vs. GP-ML	log-AULC	vs. GP-BMA	vs. GP-ML
BrainPFN-4.7M	**0.137**	<0.05↑	<0.05↑	**0.325**	<0.05↑	<0.05↑	**0.333**	<0.05↑	<0.05↑
BrainPFN-4.7M, sampled r	0.129	n.s.	<0.05↑	0.314	n.s.	<0.05↑	0.329	<0.05↑	<0.05↑
GP-BMA	0.127	–	<0.05↑	0.305	–	<0.05↑	0.319	–	<0.05↑
GP-ML	0.110	–	–	0.272	–	–	0.290	–	–
Ridge	0.126	n.s.	<0.05↑	0.298	n.s.	<0.05↑	0.310	<0.05↓	<0.05↑

Table 6: Mean Pearson r at each context size N (mean over targets and 5 folds; bold: best model at that N). The last row of each block is the two-sided paired t-test of BrainPFN-4.7M vs. GP-BMA across folds ($p < 0.05$ or n.s., in either direction).

Model	N=5	10	20	50	100	200	400	600	800
HCP-YA, real-signal (36 targets)									
BrainPFN-4.7M	**0.039**	**0.059**	**0.081**	**0.125**	**0.155**	0.182	0.221	0.238	0.250
BrainPFN-4.7M, sampled r	0.038	0.056	0.074	0.117	0.146	0.172	0.211	0.231	0.243
GP-BMA	0.034	0.049	0.067	0.110	0.142	0.180	0.216	0.232	0.244
GP-ML	0.028	0.039	0.050	0.084	0.112	0.154	0.207	0.227	0.241
Ridge	0.029	0.043	0.060	0.102	0.136	**0.184**	**0.226**	**0.245**	**0.260**
p (4.7M vs. GP-BMA)	<0.05	<0.05	<0.05	<0.05	<0.05	n.s.	<0.05	n.s.	n.s.
HCP-YA, cognition composite									
BrainPFN-4.7M	**0.121**	**0.179**	**0.201**	**0.319**	**0.376**	0.424	0.475	0.501	0.511
BrainPFN-4.7M, sampled r	0.120	0.173	0.188	0.299	0.364	0.409	0.467	0.497	0.509
GP-BMA	0.096	0.136	0.168	0.281	0.355	0.427	0.483	0.504	0.515
GP-ML	0.076	0.106	0.131	0.225	0.276	0.422	0.479	0.499	0.509
Ridge	0.083	0.131	0.141	0.254	0.340	**0.435**	**0.500**	**0.533**	**0.552**
p (4.7M vs. GP-BMA)	<0.05	<0.05	<0.05	<0.05	<0.05	n.s.	n.s.	n.s.	n.s.
HCP-A, real-signal (34 targets)									
BrainPFN-4.7M	0.154	**0.215**	**0.278**	**0.337**	**0.375**	0.396	0.430	0.449	0.460
BrainPFN-4.7M, sampled r	**0.155**	0.214	0.277	0.332	0.369	0.391	0.422	0.440	0.450
GP-BMA	0.136	0.187	0.247	0.314	0.364	0.399	0.435	0.455	0.468
GP-ML	0.115	0.154	0.202	0.263	0.313	0.393	0.433	0.453	0.465
Ridge	0.112	0.159	0.222	0.297	0.355	**0.403**	**0.447**	**0.472**	**0.488**
p (4.7M vs. GP-BMA)	<0.05	<0.05	<0.05	<0.05	<0.05	n.s.	<0.05	<0.05	<0.05

C.3 Complete Ablation Study Results

Section 5.4 reports the ablation study (kernel choice, model size, and fixed vs. sampled r, all against the flagship BrainPFN-4.7M) on HCP-Aging only. Here we give the same comparison, first summarized across all four datasets, then broken down by score domain for each dataset individually (HCP-Aging's own by-domain breakdown is Figure 5).

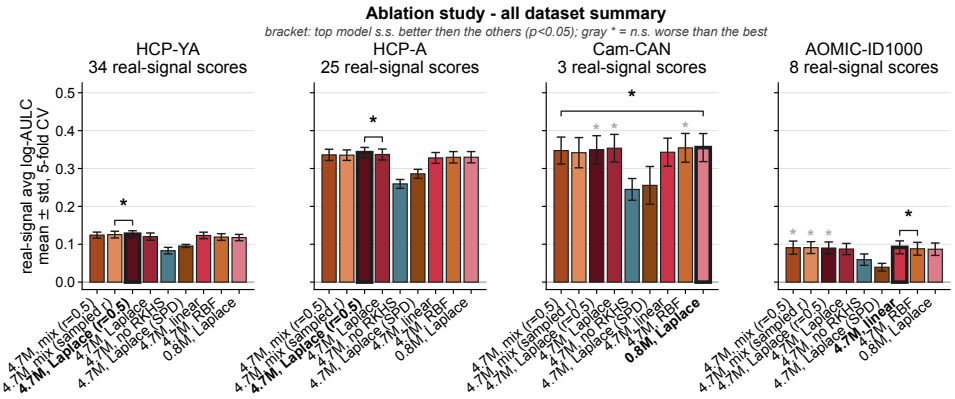


Figure 10: Ablations vs. the flagship (BrainPFN-4.7M, Laplace kernel, r=0.5), summarized per dataset. Unlike the by-domain panels below, each dataset's bar here averages only over that dataset's real-signal scores (the union, across the five plotted checkpoints, of each model's own scores with detectable signal vs. Ridge), so these numbers do not match the "ALL" panel of the corresponding by-domain figure, which averages unfiltered over every baseline-covered score.

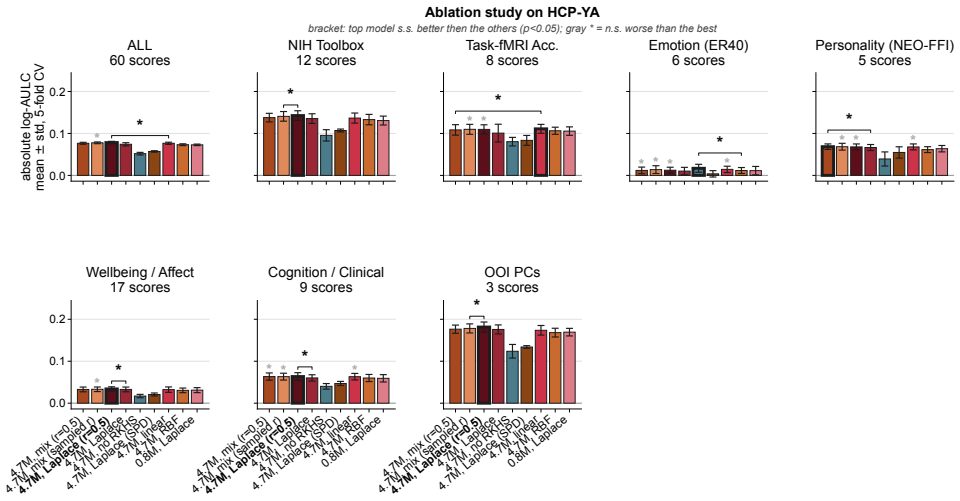


Figure 11: Ablation study on HCP-YA, by score domain.

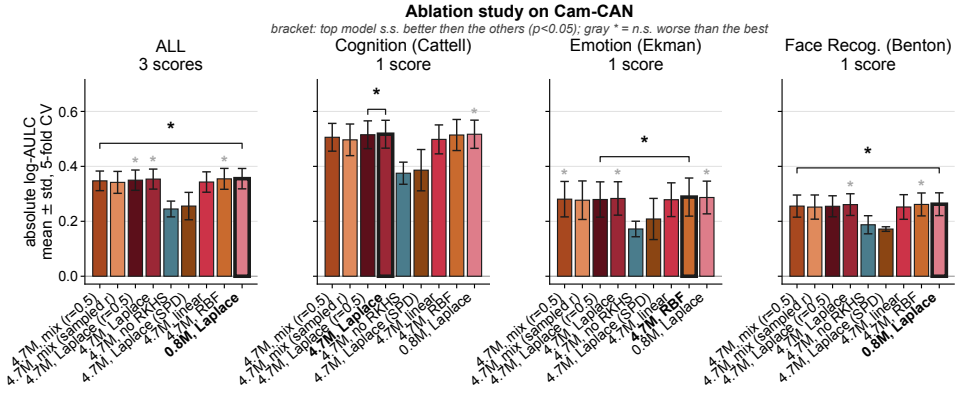


Figure 12: Ablation study on Cam-CAN, by score domain.

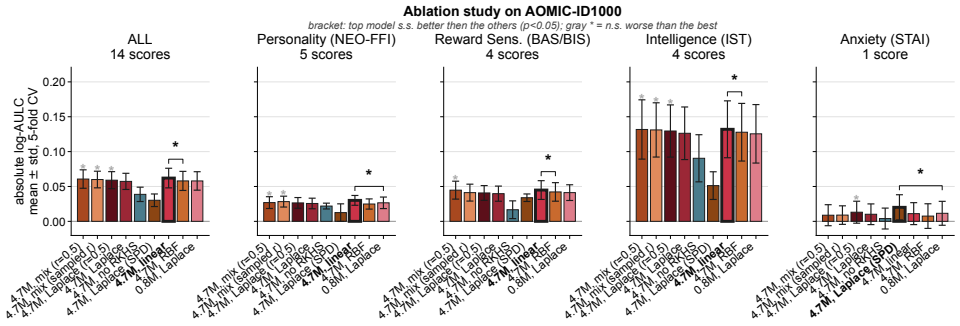


Figure 13: Ablation study on AOMIC-ID1000, by score domain.

C.4 Crossover between BrainPFN and Ridge across context sizes

We define each target's *crossover* N as the smallest context size beyond which Ridge exceeds BrainPFN for every larger N tested: the practical sample size below which BrainPFN is the better choice. Most HCP-YA and HCP-A targets cross over near N=150–200; no target on Cam-CAN or AOMIC-ID1000 crosses over within their respective pool size. Table 7 reports this crossover for representative targets spanning cognition, physical, and personality measures; a working-memory accuracy score on HCP crosses earliest, at N=50, while a matrix-reasoning score never crosses within the tested range. Figure 14 shows the same crossover directly for every real-signal target rather than this representative handful.

Table 7: Context-size crossover for representative targets: the smallest N at which Ridge's Pearson r overtakes BrainPFN's and remains higher for every larger N tested. Most HCP targets cross between N=50 and N=200; no target on Cam-CAN or AOMIC-ID1000 crosses within the tested range.

Dataset	Target	Crossover N
HCP	`WM_Task_Acc` (working-memory accuracy)	50
HCP	`pc2` (cognition composite)	200
HCP	`PicVocab_AgeAdj`	200
HCP	`Strength_Unadj` (grip strength)	200
HCP	`PMAT24_A_CR` (matrix reasoning)	never ($N \leq 800$)
Cam-CAN	`Cattell` (fluid intelligence)	never ($N \leq 400$)
Cam-CAN	`BentonFaces` (face recognition)	never ($N \leq 400$)
AOMIC-ID1000	`IST_intelligence_total`	never ($N \leq 600$)

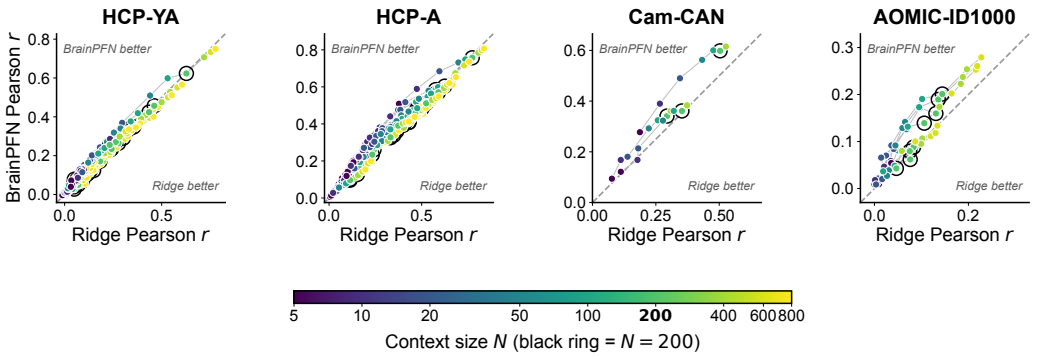


Figure 14: Ridge Pearson r versus BrainPFN Pearson r for every real-signal target, one point per (target, context size N), coloured by N (light = small, dark = large); points for the same target at consecutive N are joined by a thin grey line. Points above the dashed diagonal are context sizes at which BrainPFN wins; N=200 points are marked with a black ring. On HCP and HCP-Aging, trajectories cross the diagonal right around N=200; on Cam-CAN and AOMIC-ID1000, they never do within the tested range.

C.5 Movie-watching vs. resting-state functional connectivity on HCP

Figure 15 compares FC computed from the HCP 7T movie-watching scans against FC from each subject's own resting-state scan (subject-matched to the movie cohort), for BrainPFN-4.7M and Ridge alike. Of the four HCP 7T movie-watching runs (`MOVIE1--4_7T`), we use `MOVIE1_7T_AP`, which covers the most subjects (176) among the four. Movie-watching FC gives a consistently higher signal than resting-state FC for both models at every context size, and BrainPFN's margin over Ridge is preserved across both FC inputs – the gain from movie-derived FC and the gain from BrainPFN over Ridge are additive rather than substituting for one another.

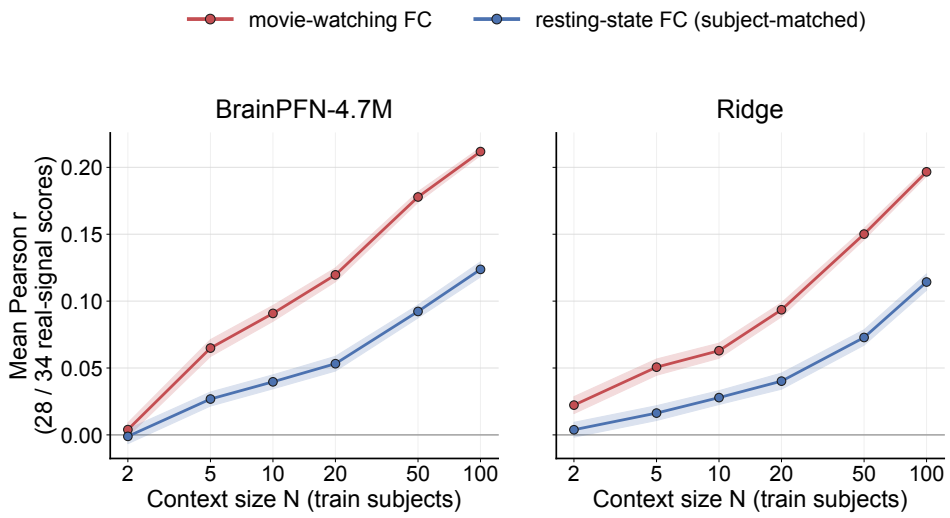


Figure 15: HCP: mean Pearson r vs. context size N for movie-watching FC and subject-matched resting-state FC, restricted to each FC input's own real-signal scores (28 for movie-watching, 34 for resting-state), 5-fold CV, 20 repeats per fold.

D Full List of Results

D.1 Per-score, all-model comparison, all 5-fold benchmarks

The per-score breakdown above (Sec. D.2) pairs each model with its own Ridge baseline, one model per figure. This section instead gives every score its own panel with all 7 models plotted together (same style as Figure 2's curve panels: one line per model, colored and sized as in the legend, BrainPFN variants shown bold, plus the shared canonical-Ridge dashed reference curve, "*" marking N where a model is significantly ahead of its own Ridge – one-sided paired t-test on the 5 fold-level means, Holm-corrected across that panel's own N-grid), split per dataset and paginated the same way the per-model perscore figures are (panels ordered by peak Pearson r across every model; `COG` tags the dataset's designated cognition target, `SIG` tags a score with real signal for BrainPFN-4.7M vs. Ridge). HCP-YA (60 scores) is split across 7 pages, HCP-A across 3, and AOMIC-ID1000 across 2; Cam-CAN (3 scores) fits on one page.

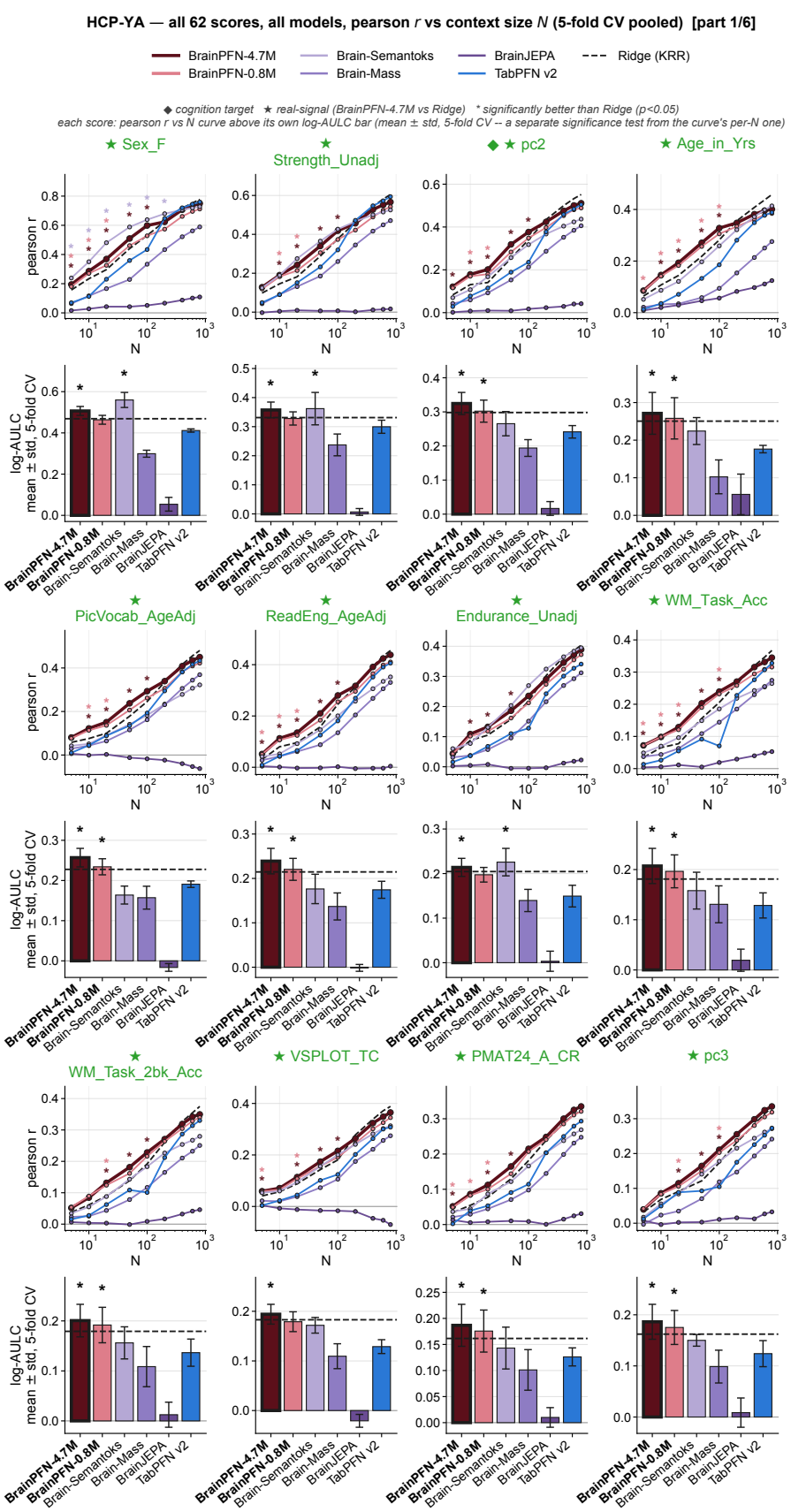


Figure 16: All 7 models vs. shared Ridge reference, all scores, HCP-YA, 5-fold CV (1/6).

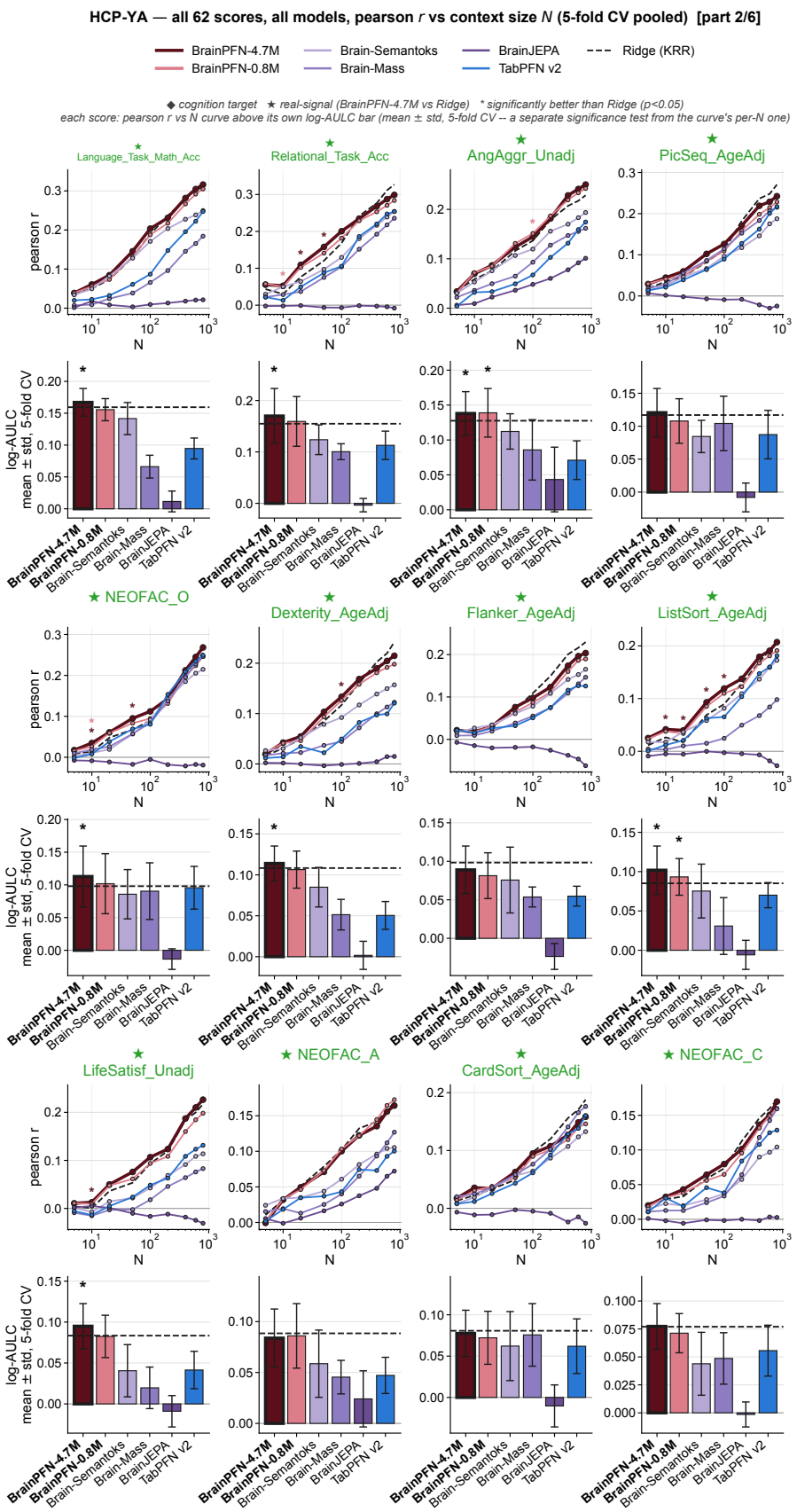


Figure 17: All 7 models vs. shared Ridge reference, all scores, HCP-YA, 5-fold CV (2/6).

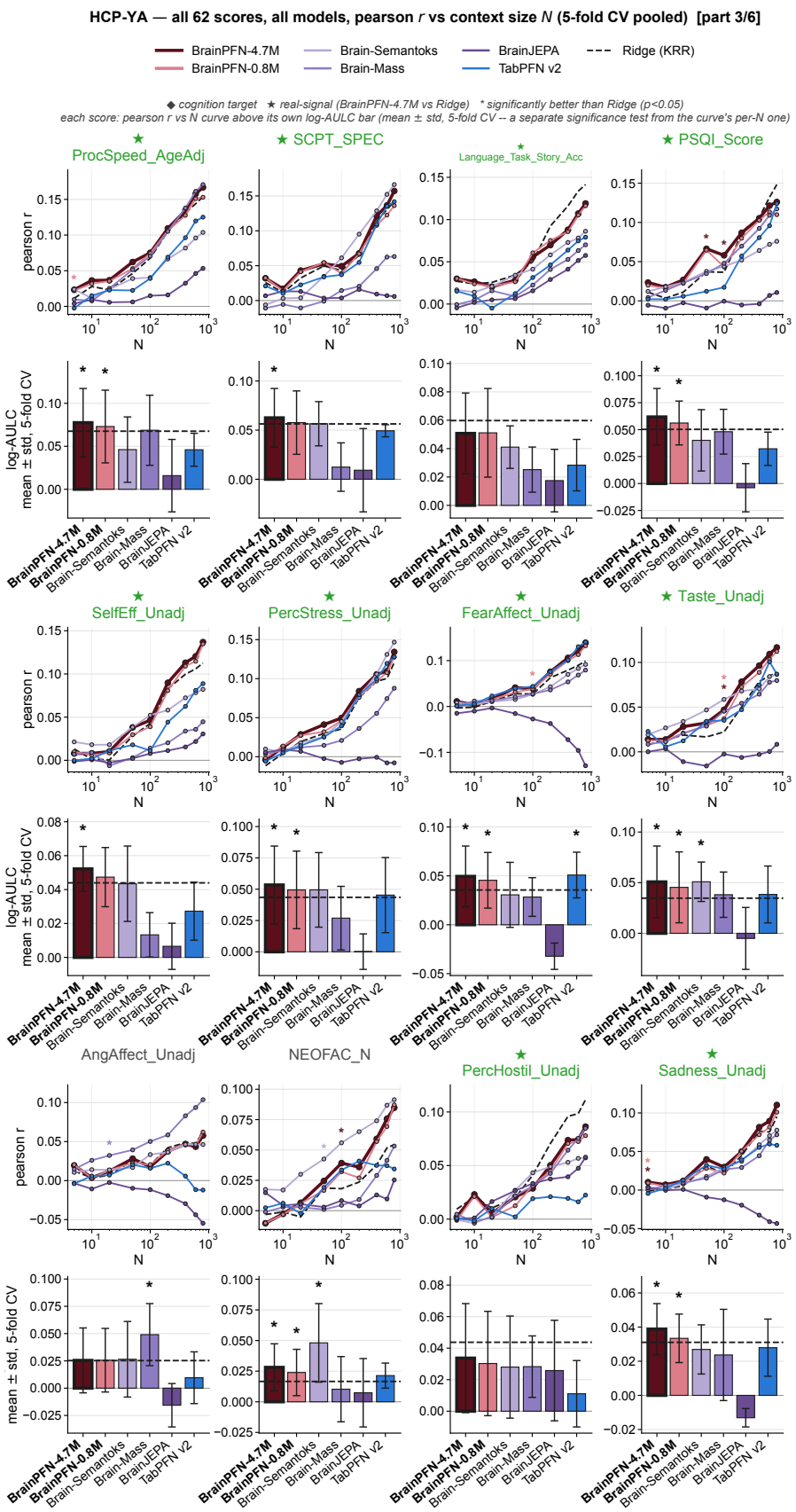


Figure 18: All 7 models vs. shared Ridge reference, all scores, HCP-YA, 5-fold CV (3/6).

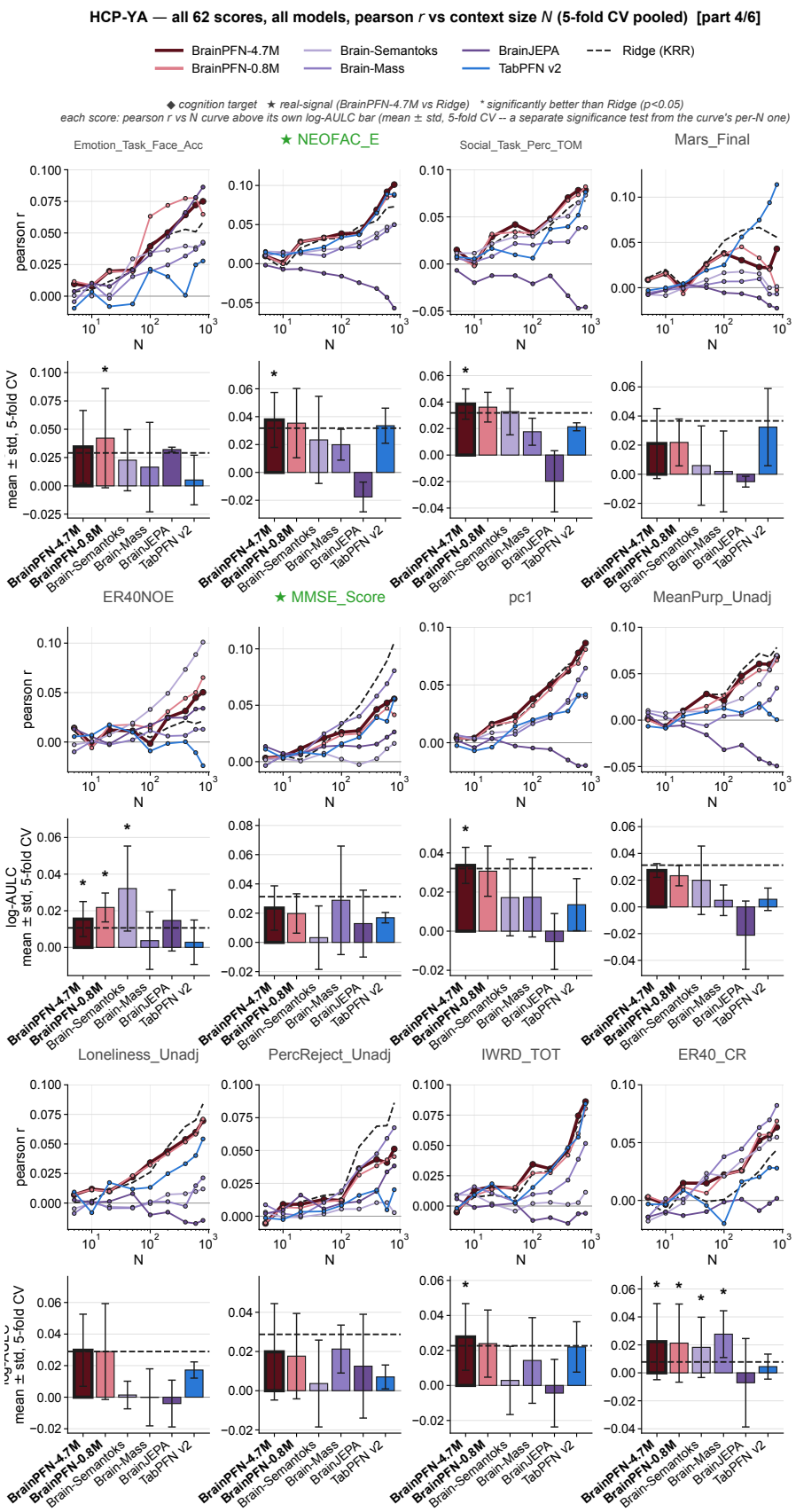


Figure 19: All 7 models vs. shared Ridge reference, all scores, HCP-YA, 5-fold CV (4/6).

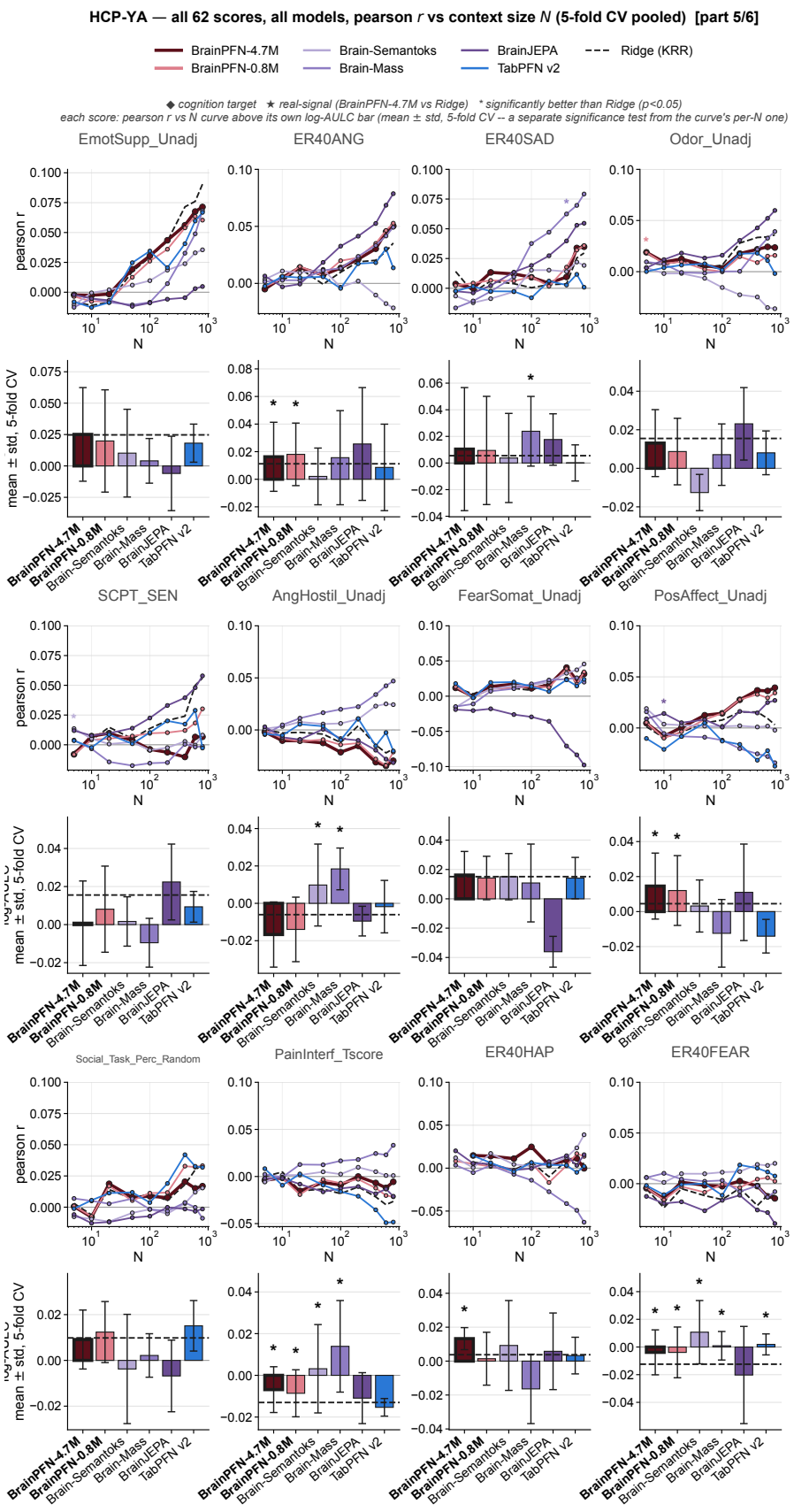


Figure 20: All 7 models vs. shared Ridge reference, all scores, HCP-YA, 5-fold CV (5/6).

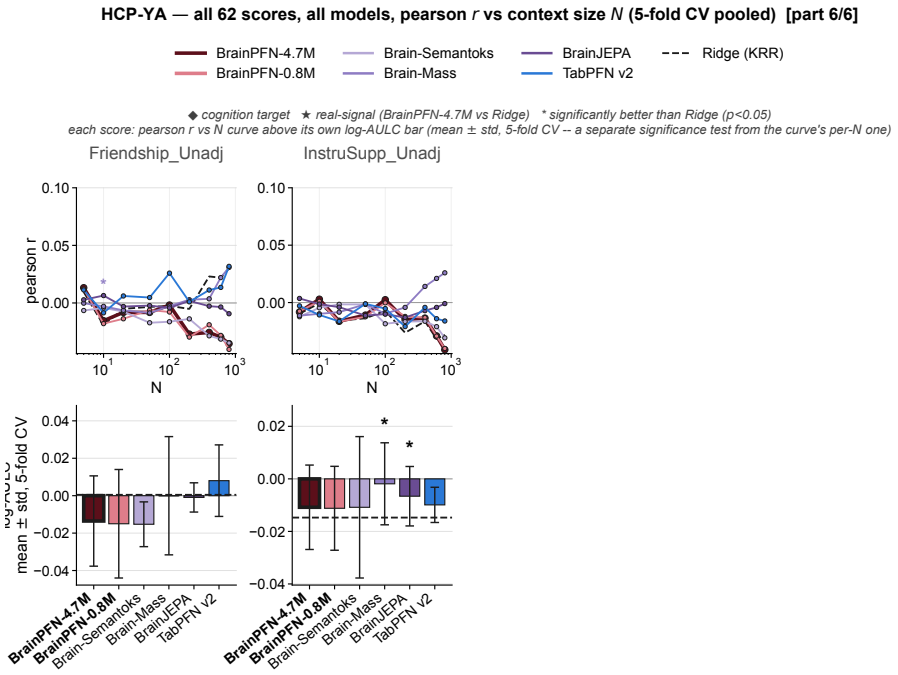


Figure 21: All 7 models vs. shared Ridge reference, all scores, HCP-YA, 5-fold CV (6/6).

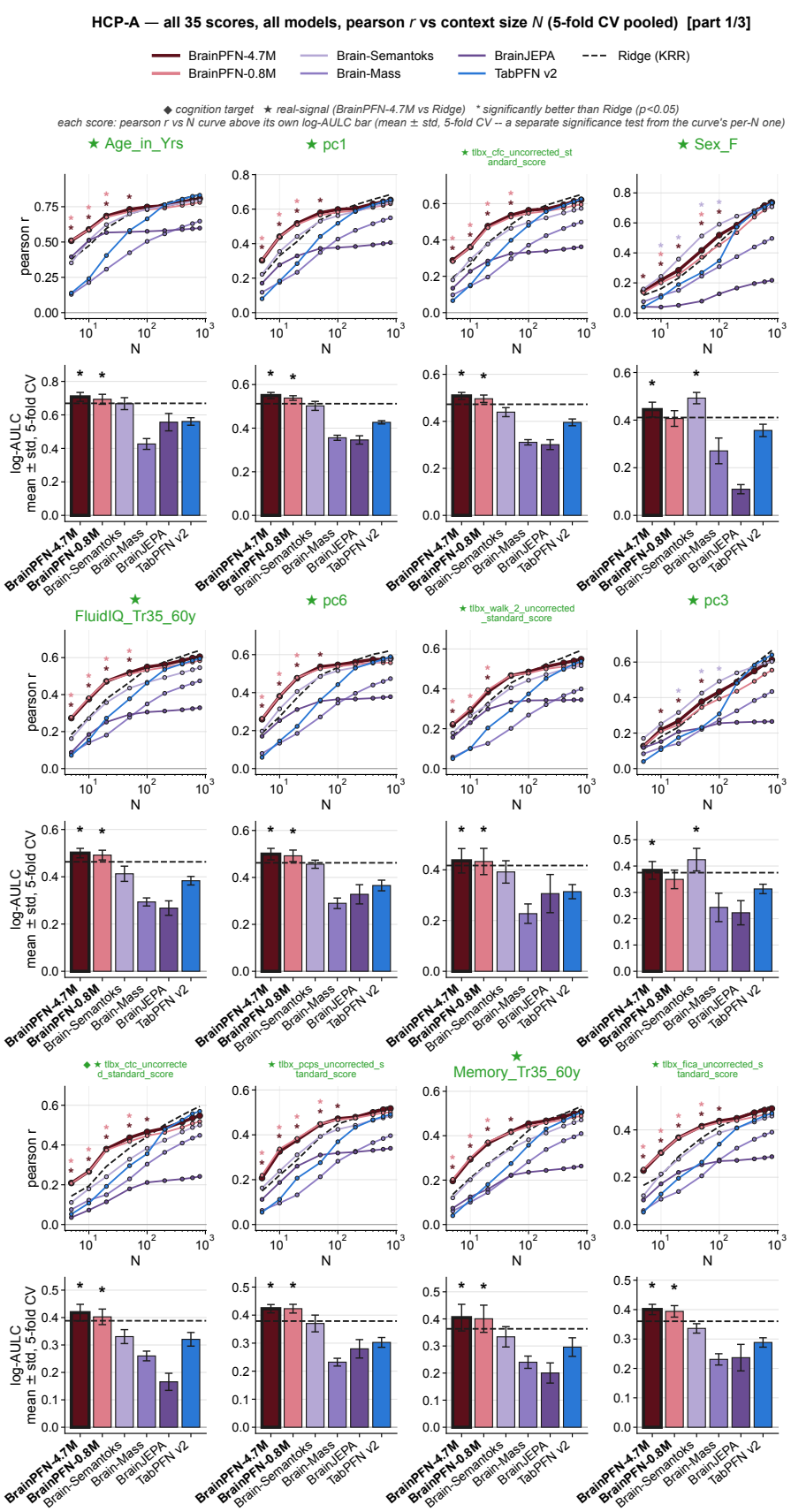


Figure 22: All 7 models vs. shared Ridge reference, all scores, HCP-A, 5-fold CV (1/3).

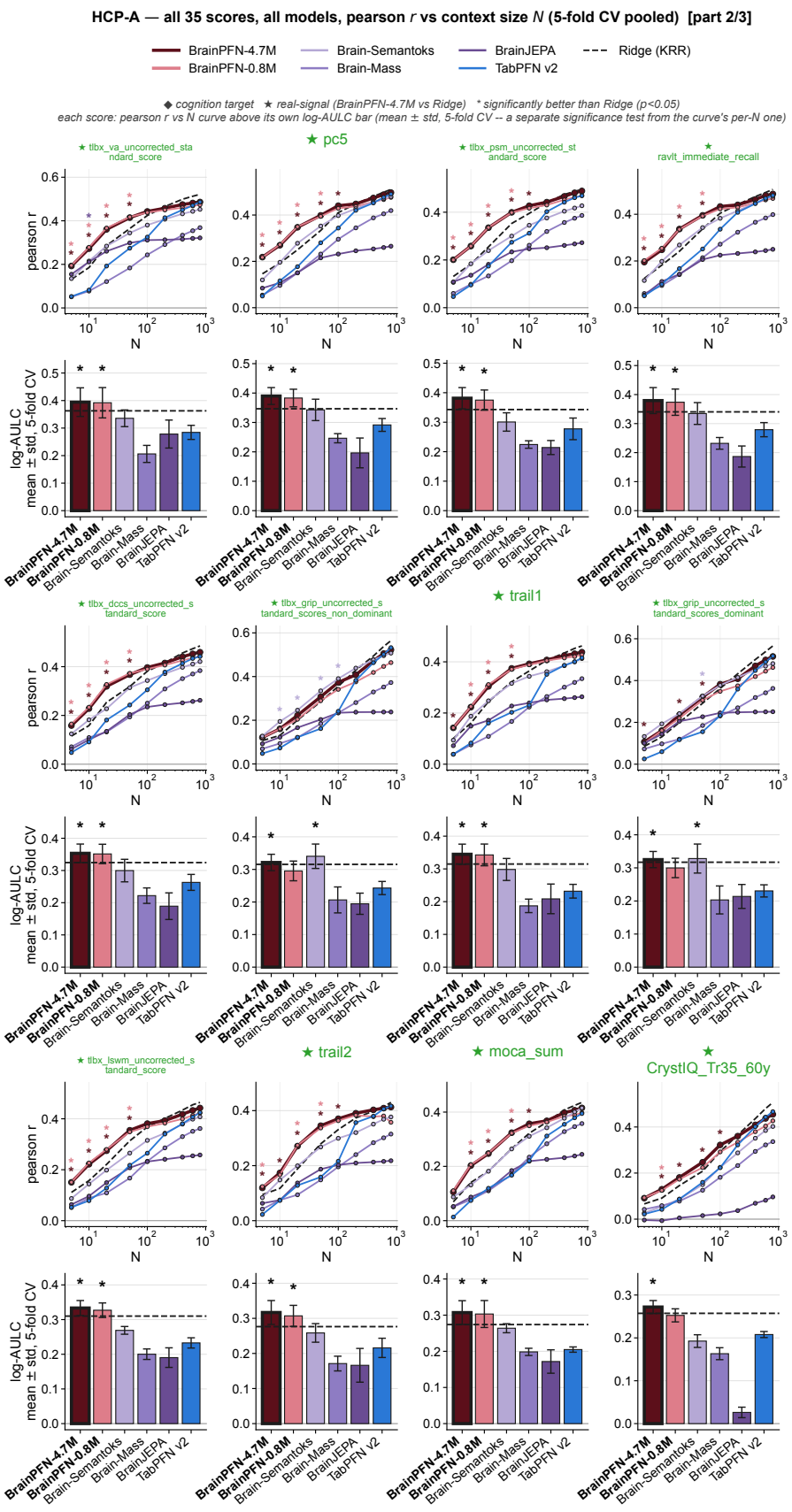


Figure 23: All 7 models vs. shared Ridge reference, all scores, HCP-A, 5-fold CV (2/3).

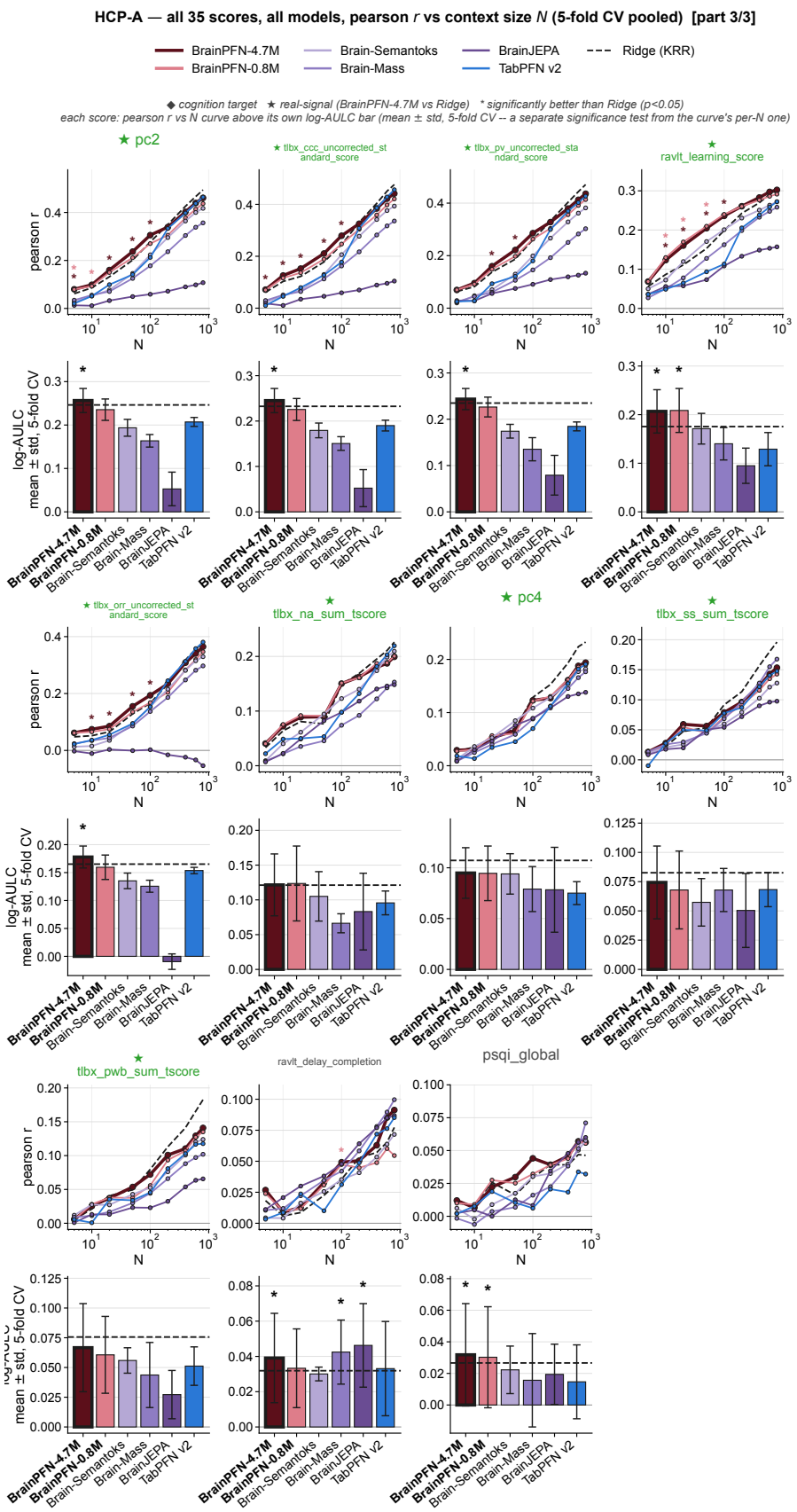


Figure 24: All 7 models vs. shared Ridge reference, all scores, HCP-A, 5-fold CV (3/3).

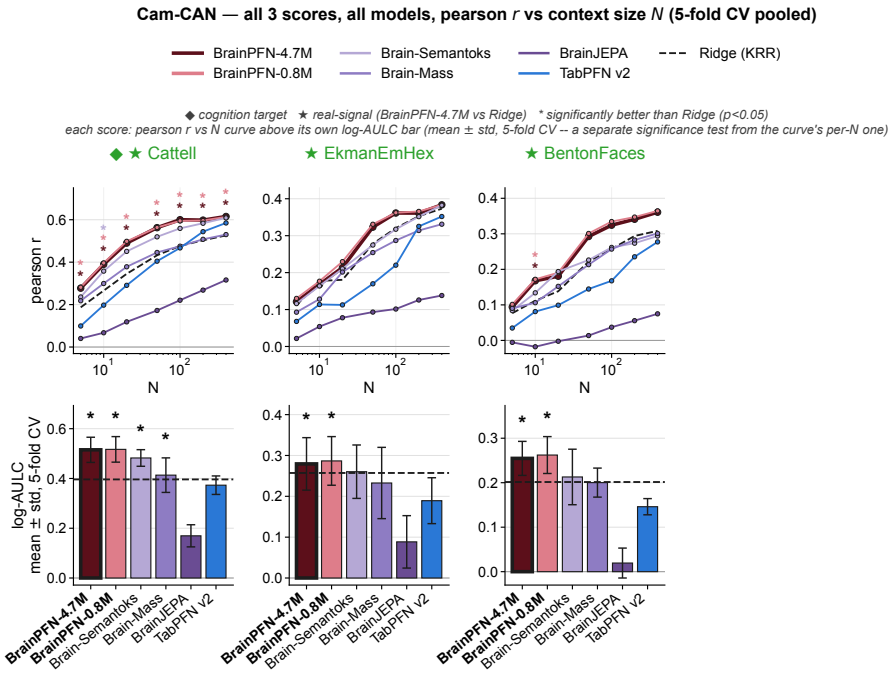


Figure 25: All 7 models vs. shared Ridge reference, all scores, Cam-CAN, 5-fold CV.

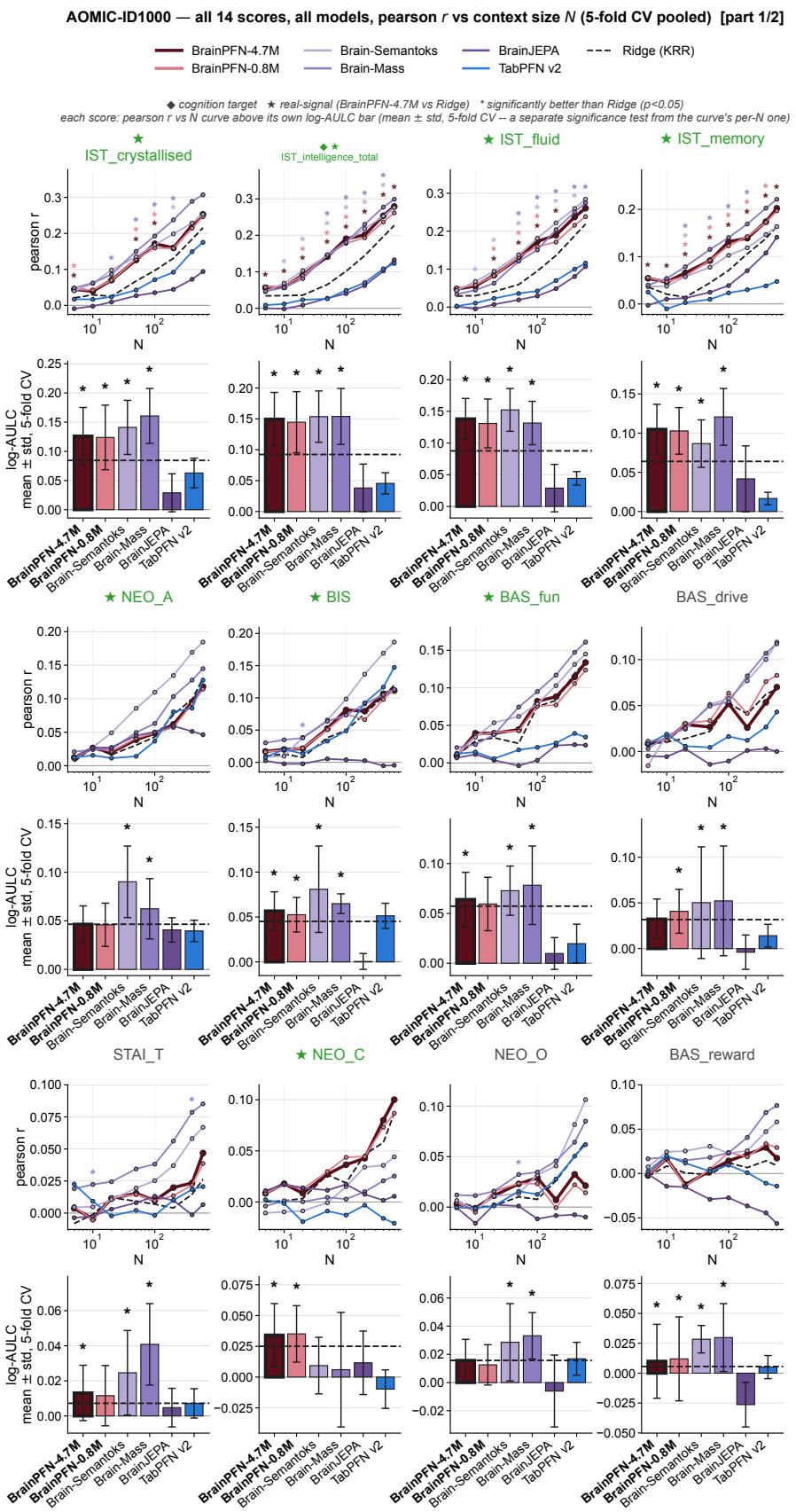


Figure 26: All 7 models vs. shared Ridge reference, all scores, AOMIC-ID1000, 5-fold CV (1/2).

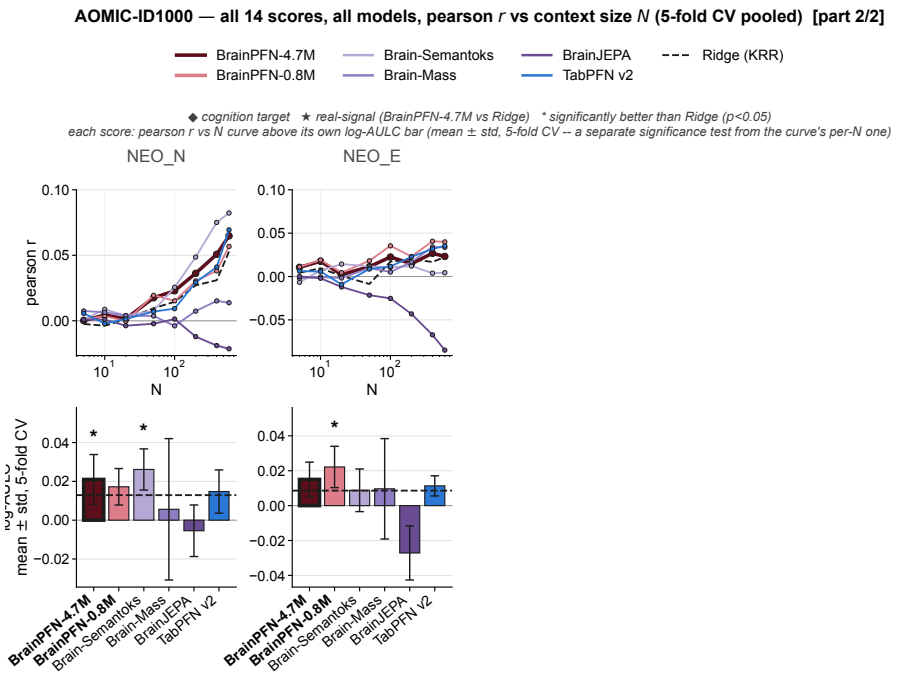


Figure 27: All 7 models vs. shared Ridge reference, all scores, AOMIC-ID1000, 5-fold CV (2/2).

D.2 PER-SCORE LEARNING CURVES, ALL 5-FOLD BENCHMARKS

For completeness, this section reports the full per-score breakdown (all ∼60 behavioral scores, Pearson r vs. context size N, pooled across the 5-fold CV) underlying the aggregate results of Figure 2, for every model and every dataset that has a complete 5-fold run – i.e. every (model, dataset) pair that contributes to `5fold_model_comparison_avg_curve.pdf`: BrainPFN-4.7M, BrainPFN-0.8M, Brain-Semantoks, Brain-Mass, BrainJEPA and TabPFN v2, on HCP-YA, HCP-A, Cam-CAN, and AOMIC-ID1000. HCP-YA (60 scores) is split across 3 pages and HCP-A across 2; every other panel fits on a single page.

D.2.1 BRAINPFN-4.7M

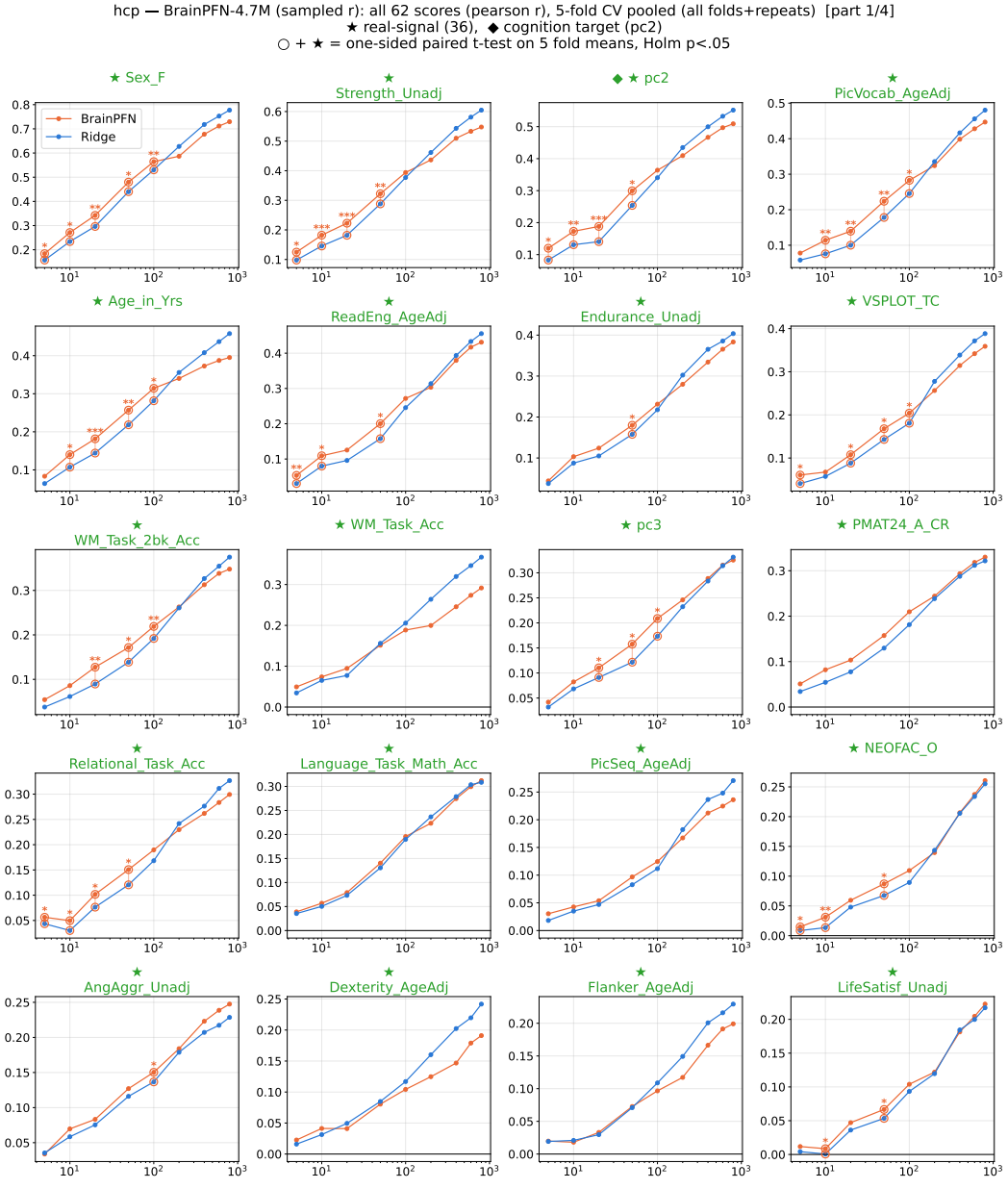


Figure 28: BrainPFN-4.7M, all scores, HCP-YA, 5-fold CV (1/3).

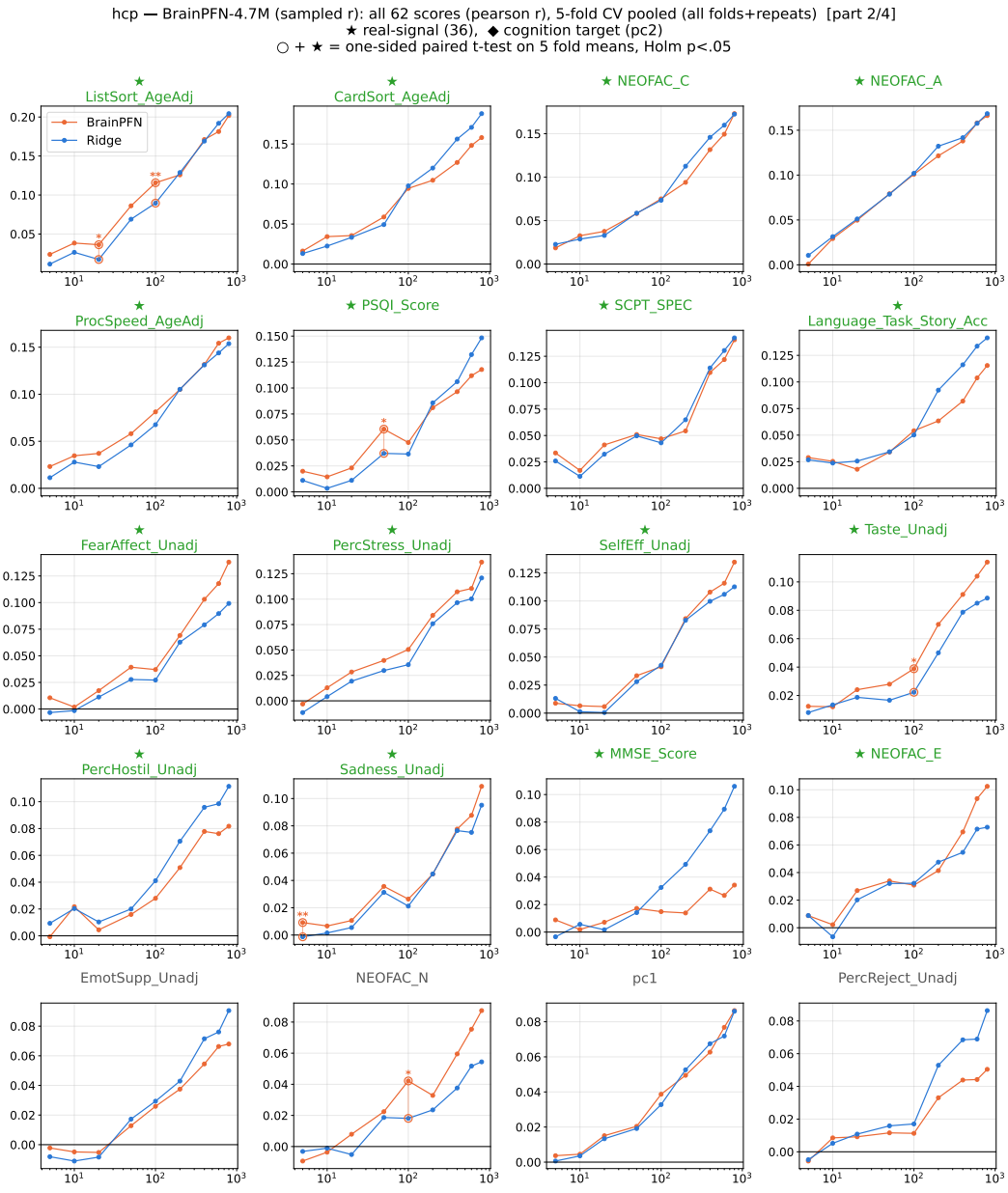


Figure 29: BrainPFN-4.7M, all scores, HCP-YA, 5-fold CV (2/3).

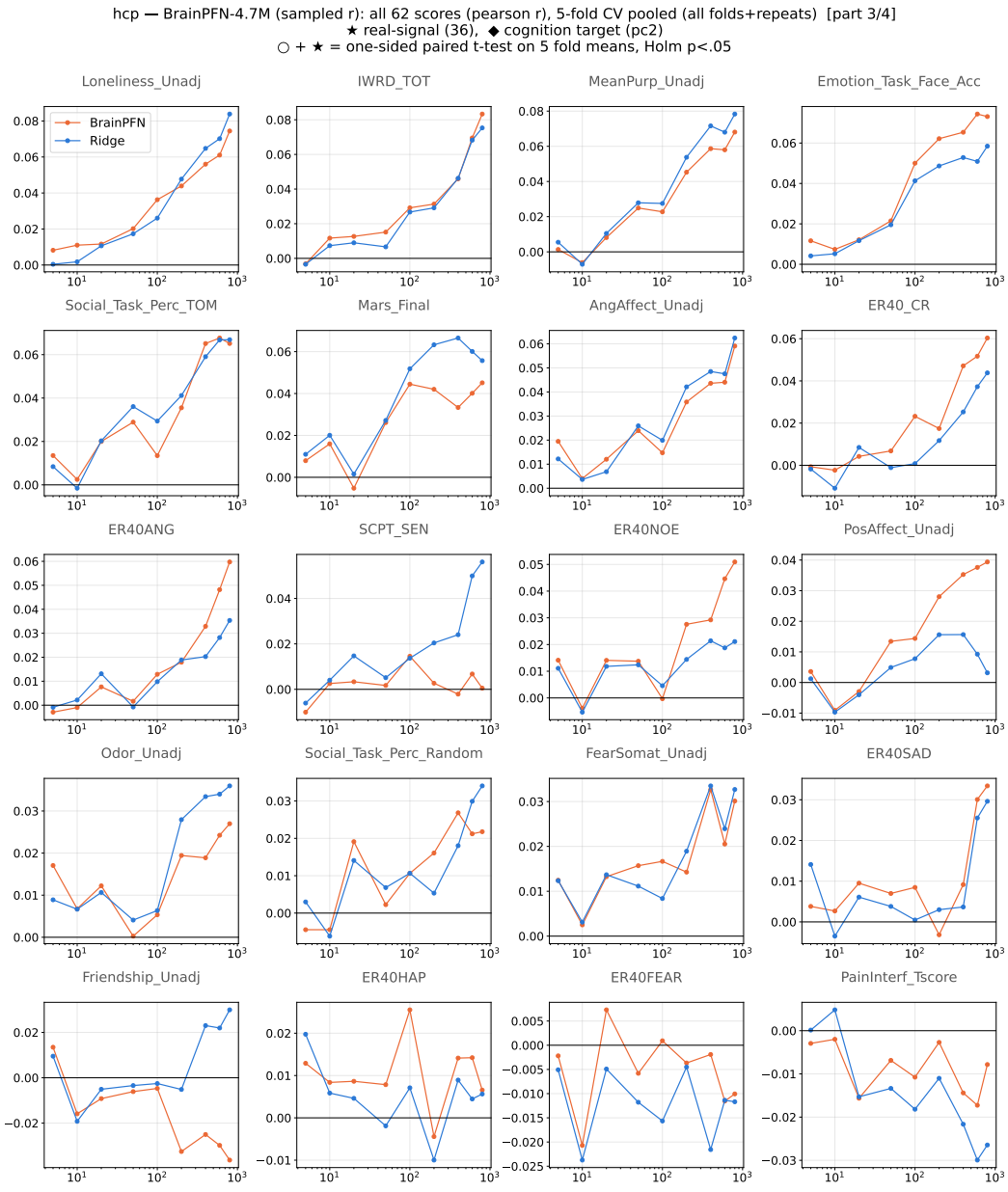


Figure 30: BrainPFN-4.7M, all scores, HCP-YA, 5-fold CV (3/3).

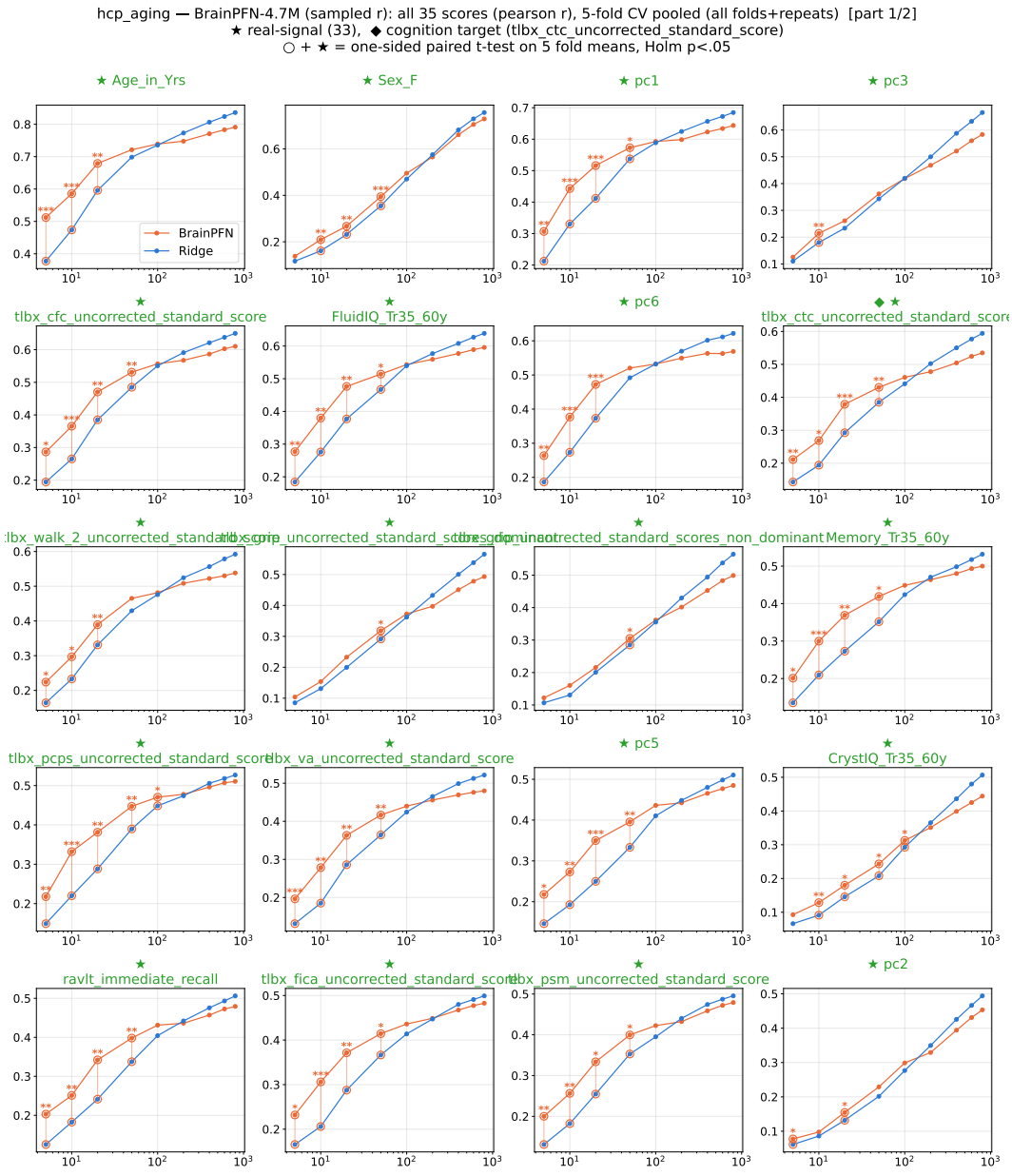


Figure 31: BrainPFN-4.7M, all scores, HCP-A, 5-fold CV (1/2).

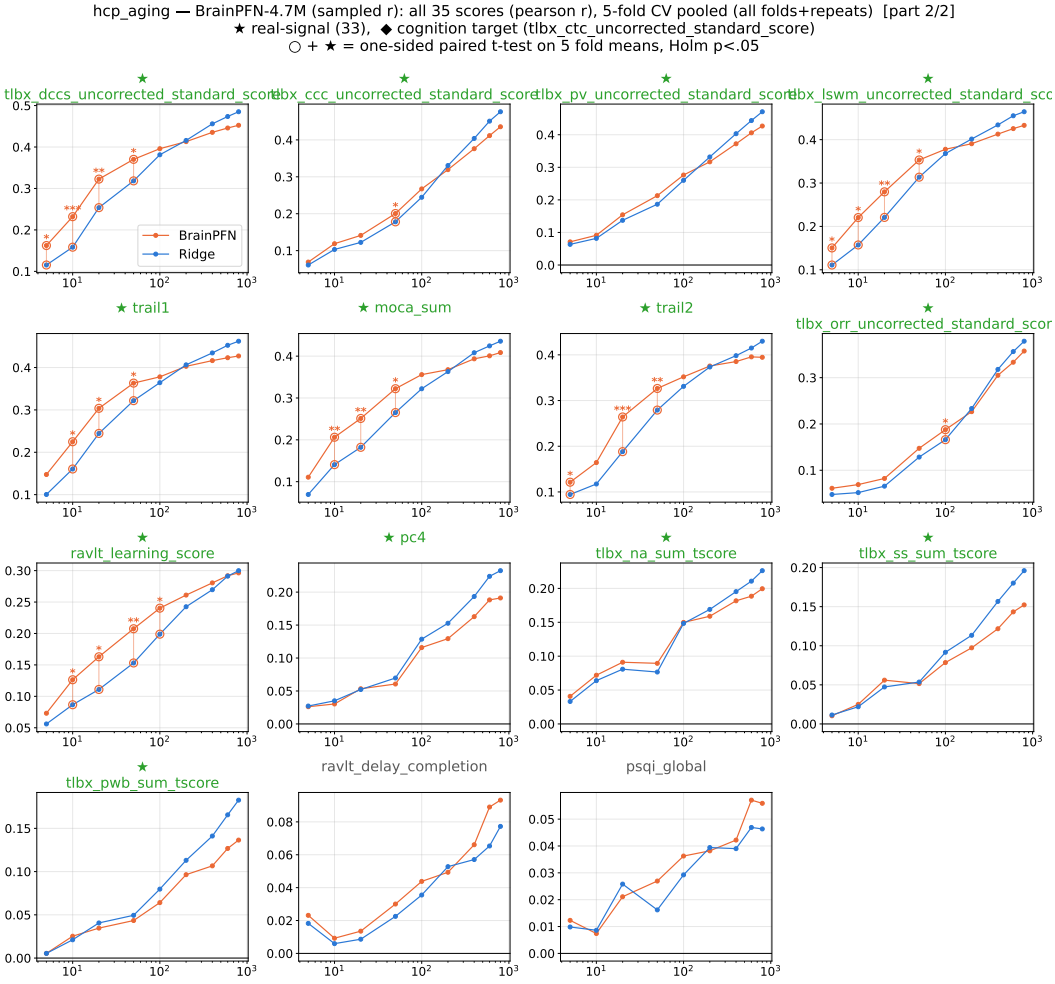


Figure 32: BrainPFN-4.7M, all scores, HCP-A, 5-fold CV (2/2).

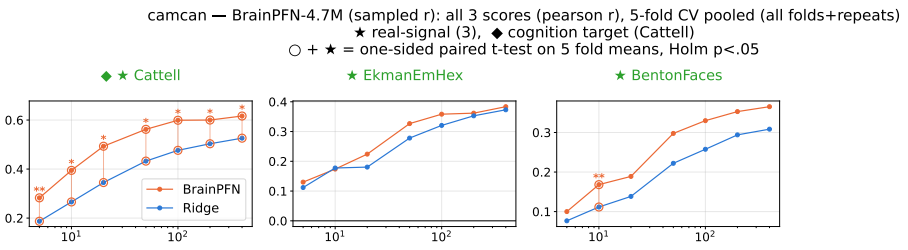


Figure 33: BrainPFN-4.7M, all scores, Cam-CAN, 5-fold CV.

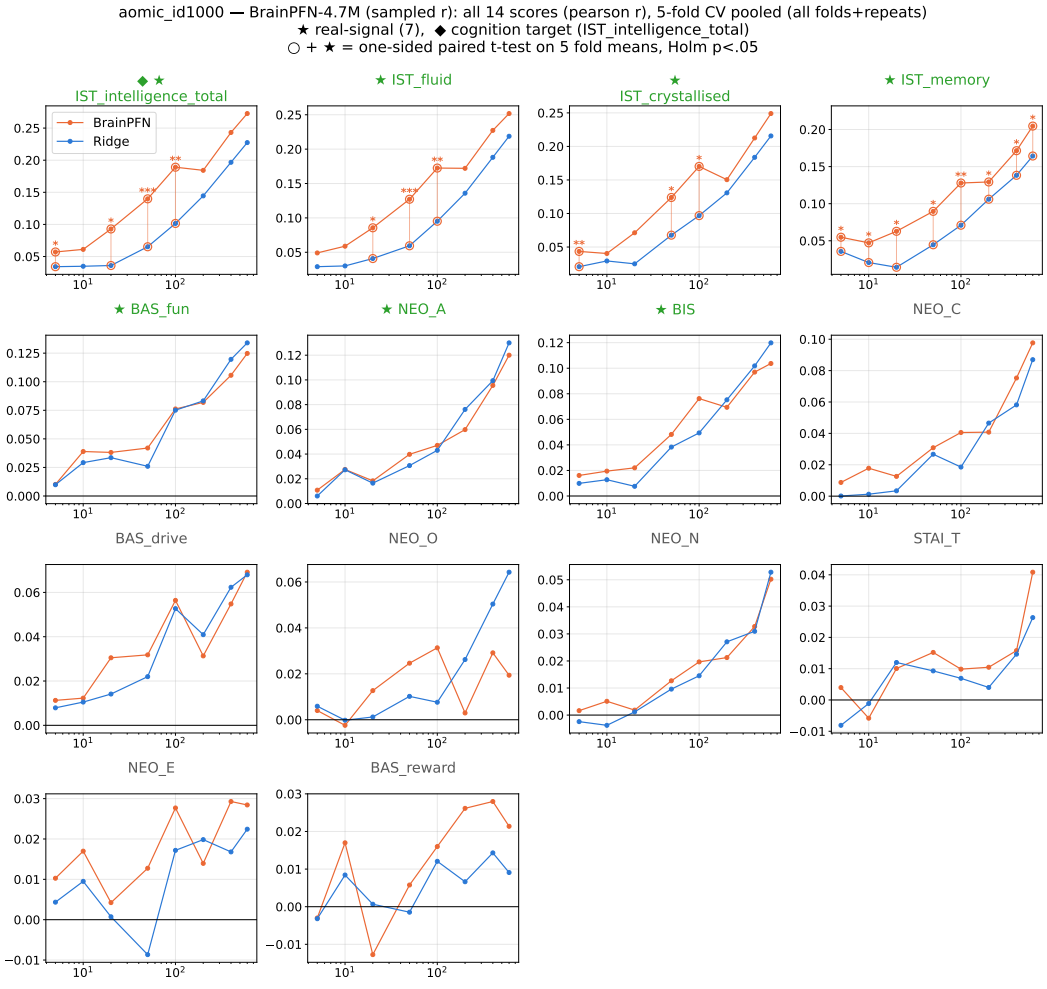


Figure 34: BrainPFN-4.7M, all scores, AOMIC-ID1000, 5-fold CV.

D.2.2 BRAINPFN-0.8M

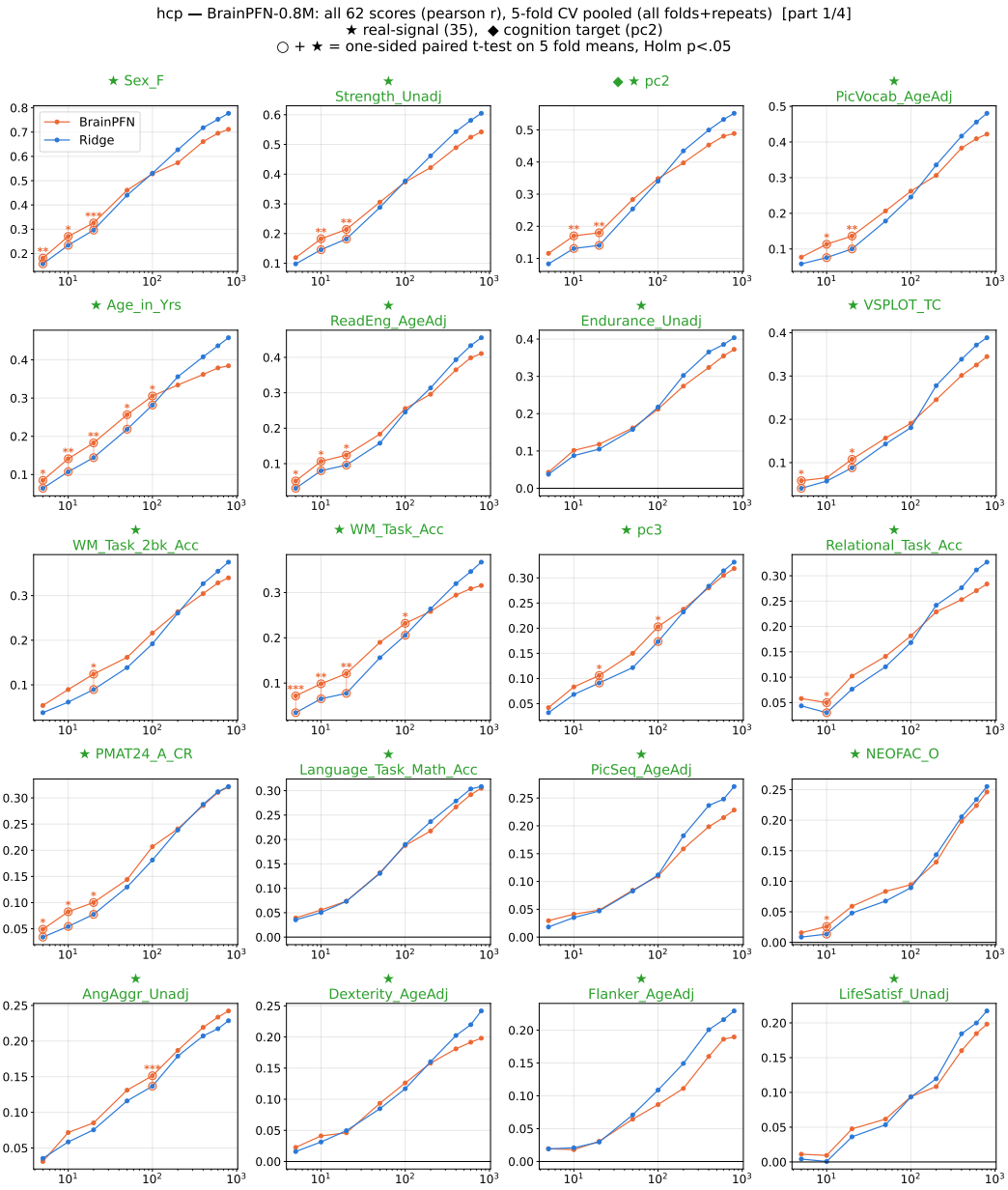


Figure 35: BrainPFN-0.8M, all scores, HCP-YA, 5-fold CV (1/3).

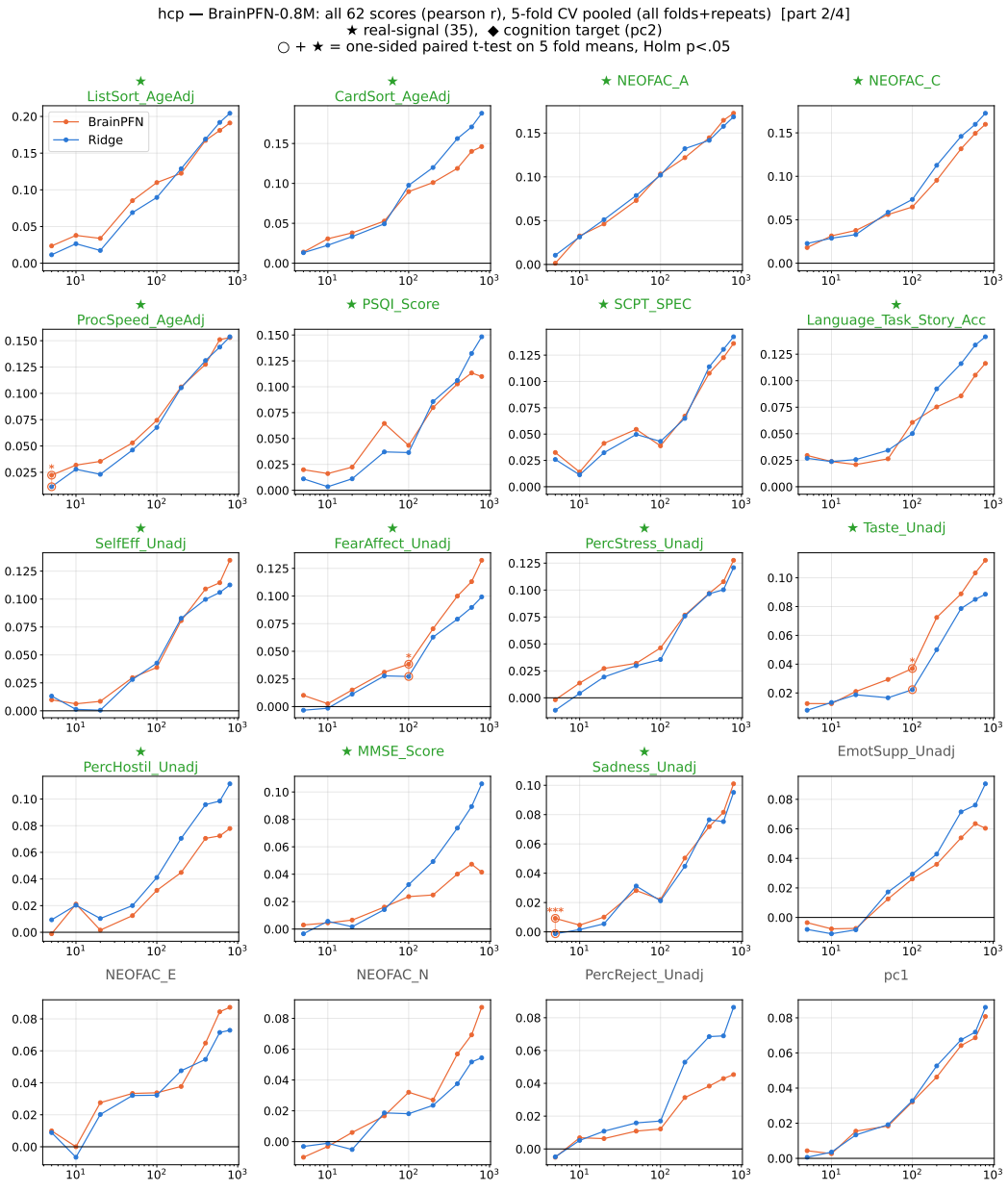


Figure 36: BrainPFN-0.8M, all scores, HCP-YA, 5-fold CV (2/3).



Figure 37: BrainPFN-0.8M, all scores, HCP-YA, 5-fold CV (3/3).

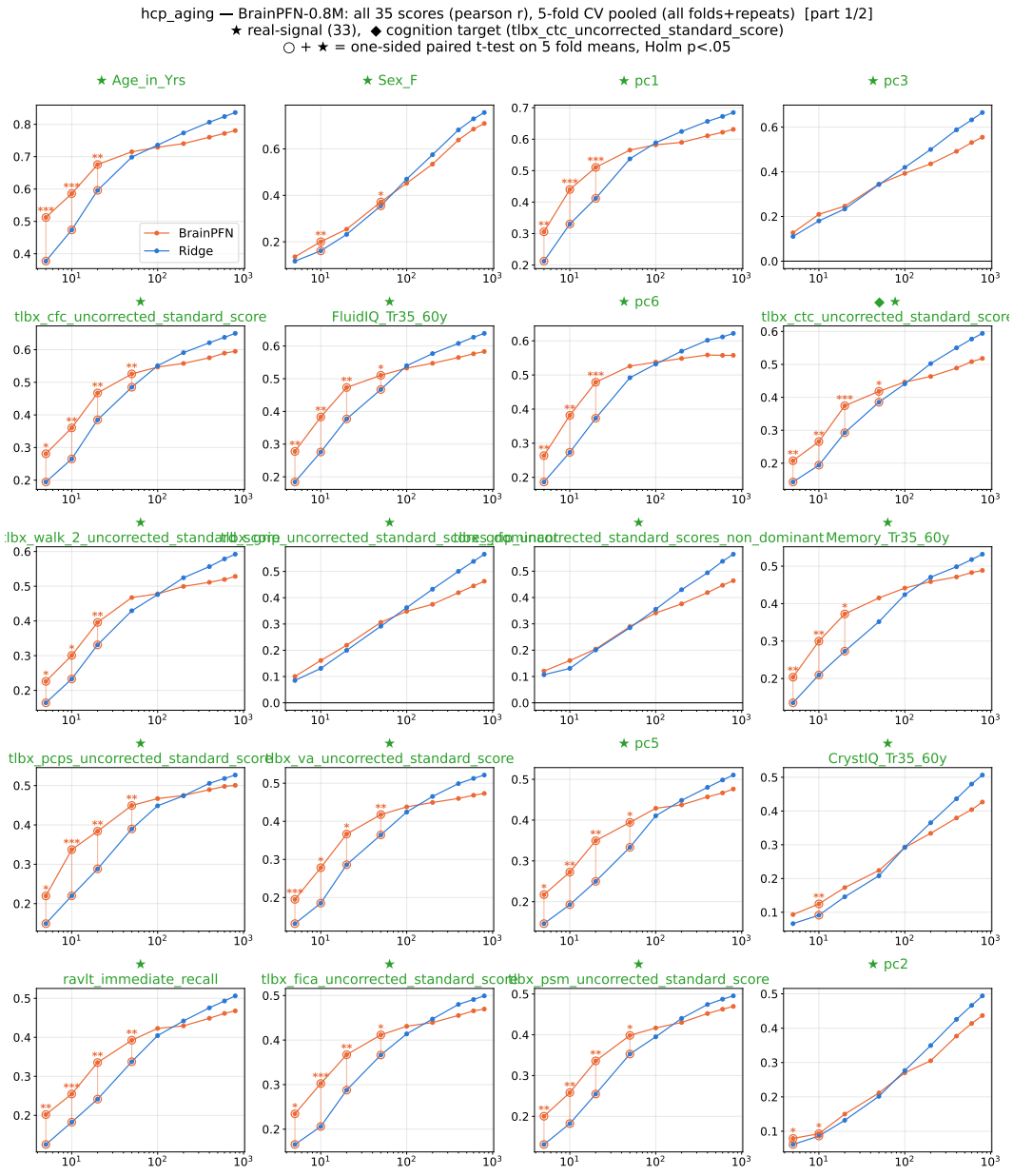


Figure 38: BrainPFN-0.8M, all scores, HCP-A, 5-fold CV (1/2).

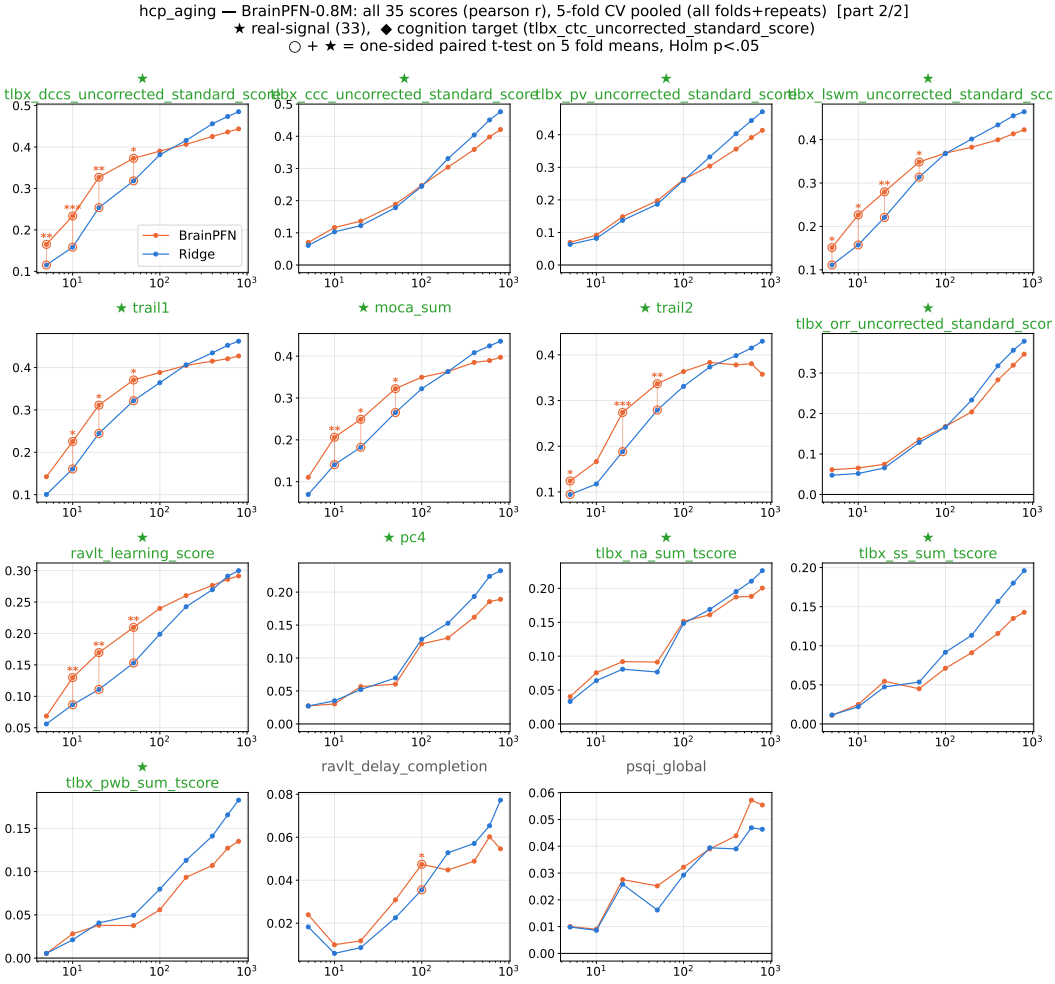


Figure 39: BrainPFN-0.8M, all scores, HCP-A, 5-fold CV (2/2).

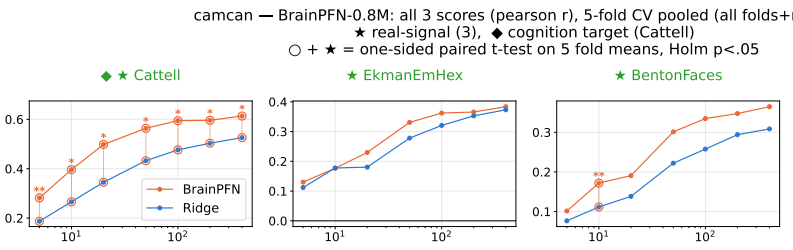


Figure 40: BrainPFN-0.8M, all scores, Cam-CAN, 5-fold CV.

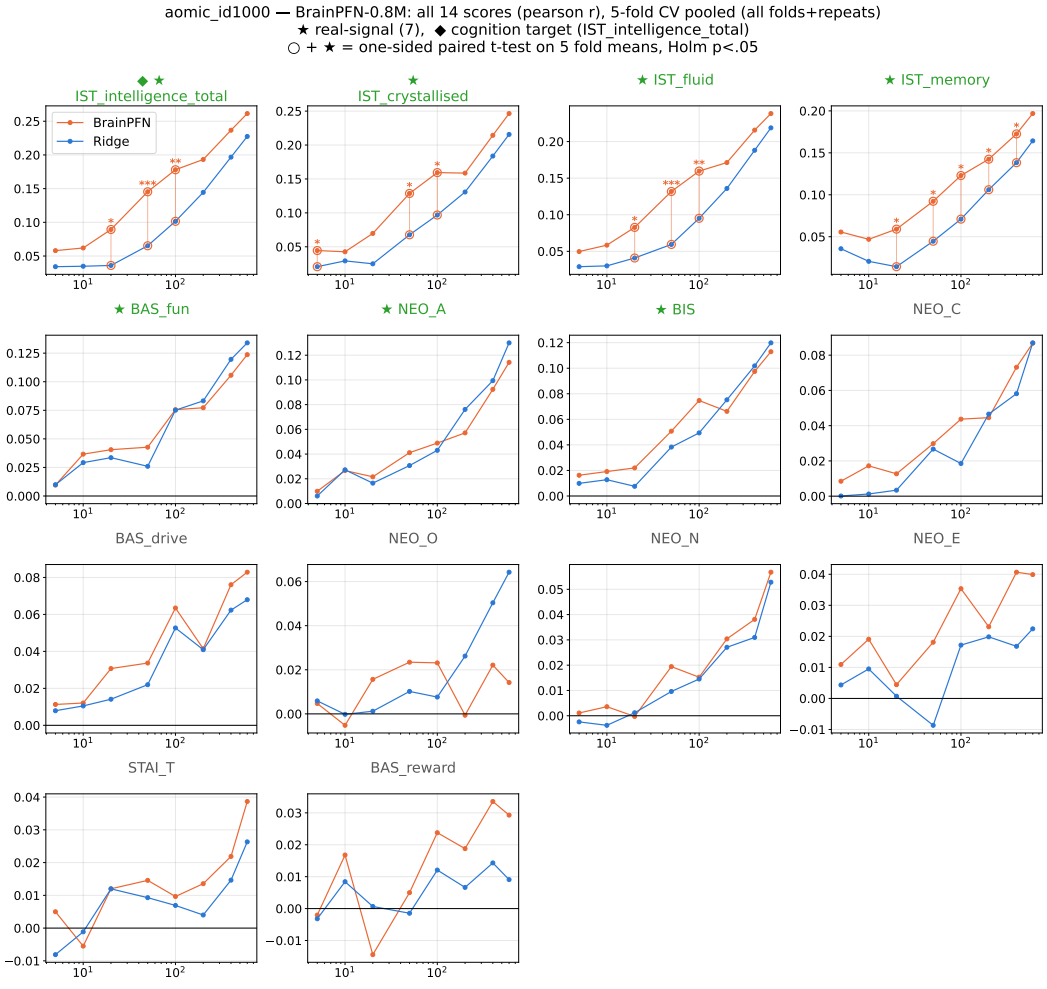


Figure 41: BrainPFN-0.8M, all scores, AOMIC-ID1000, 5-fold CV.

D.2.3 BRAIN-SEMANTOKS

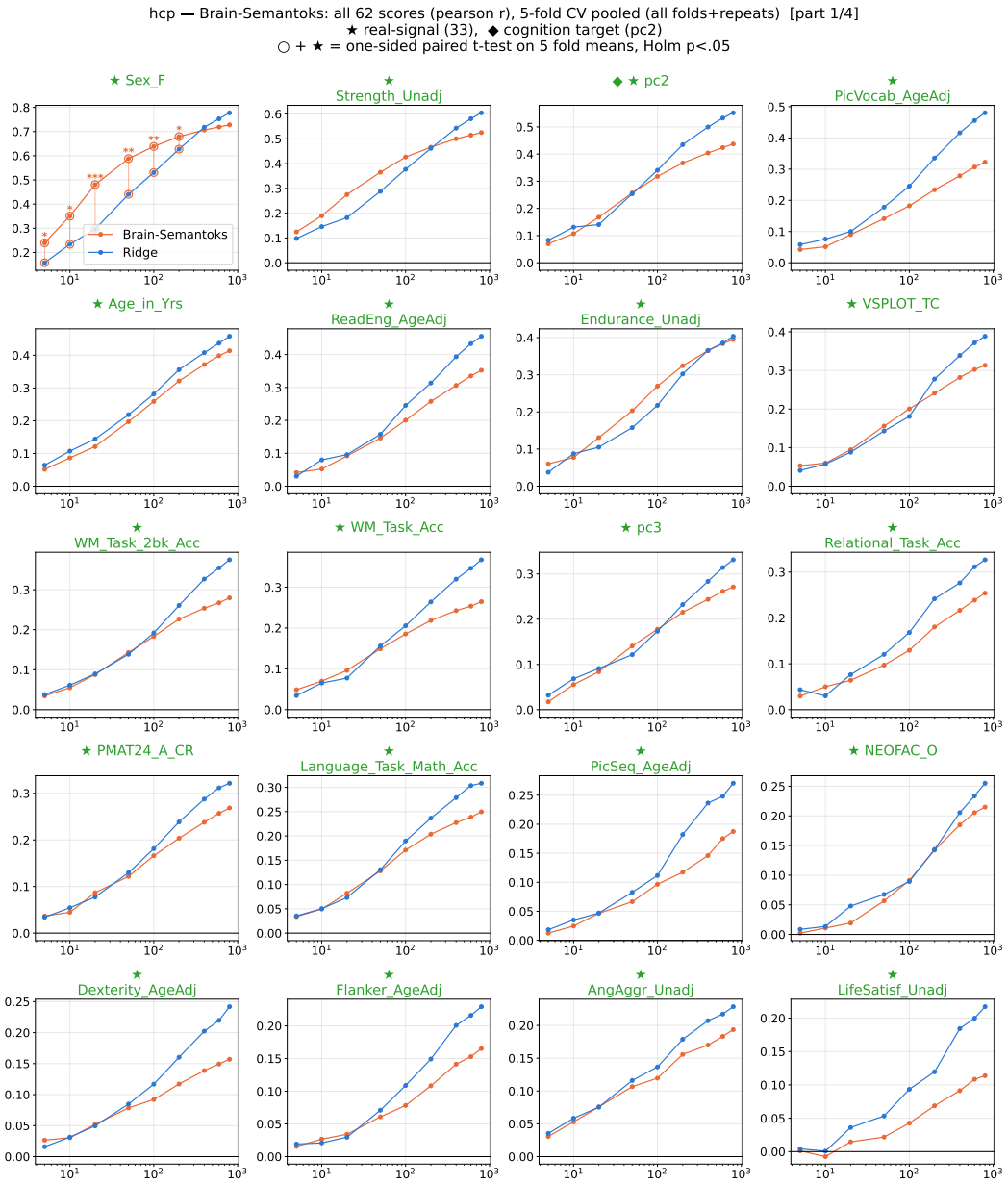


Figure 42: Brain-Semantoks, all scores, HCP-YA, 5-fold CV (1/3).

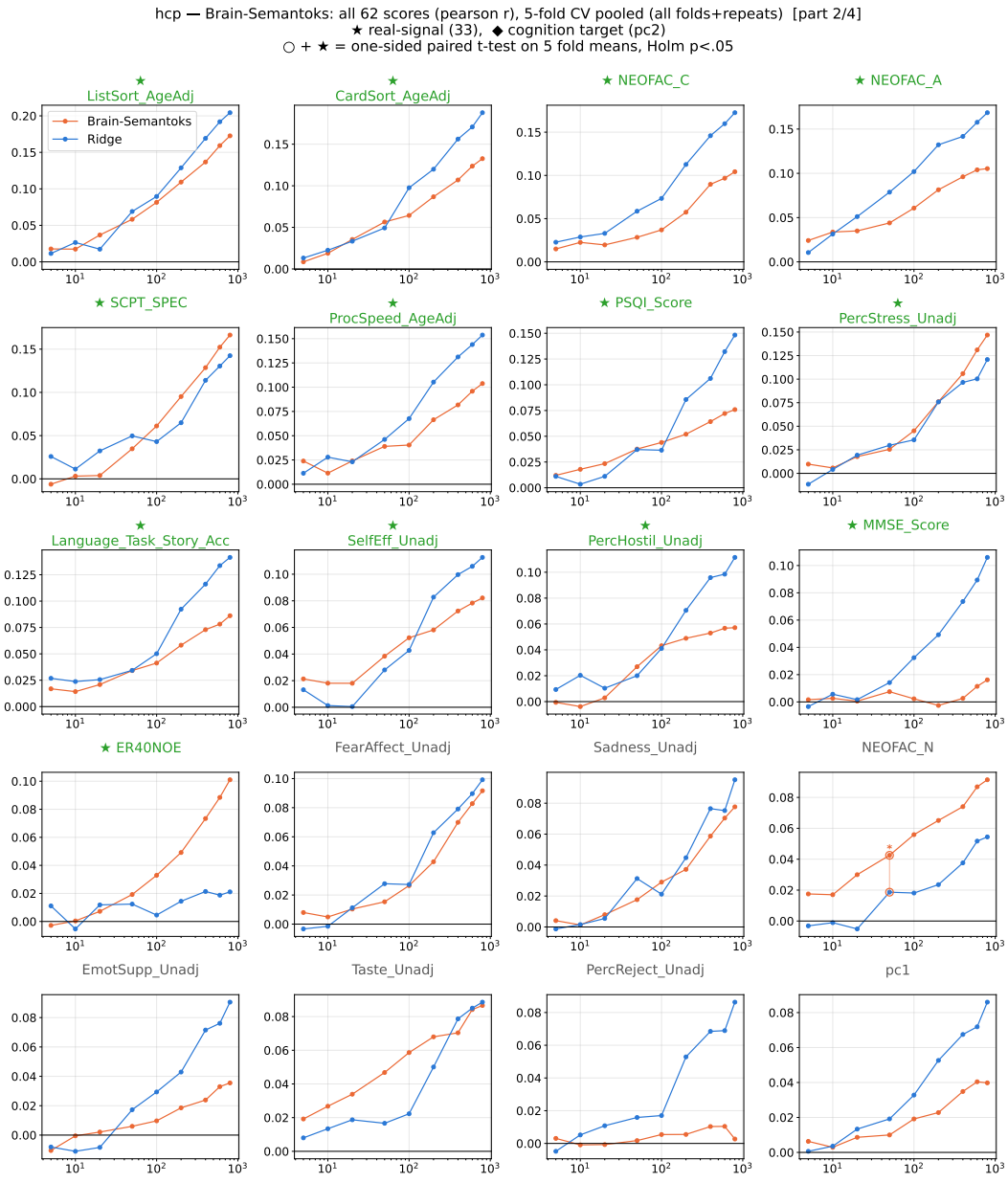


Figure 43: Brain-Semantoks, all scores, HCP-YA, 5-fold CV (2/3).

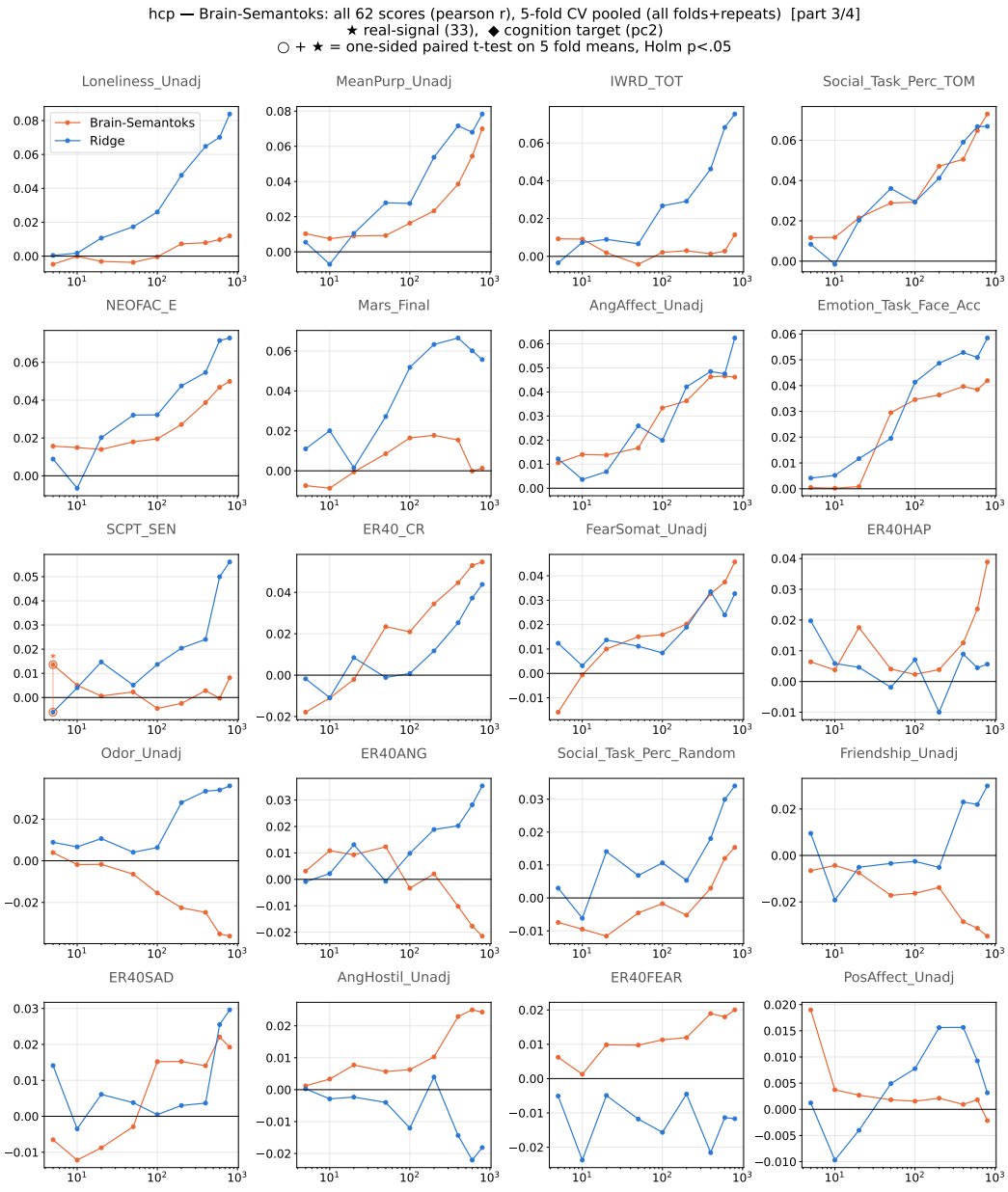


Figure 44: Brain-Semantoks, all scores, HCP-YA, 5-fold CV (3/3).

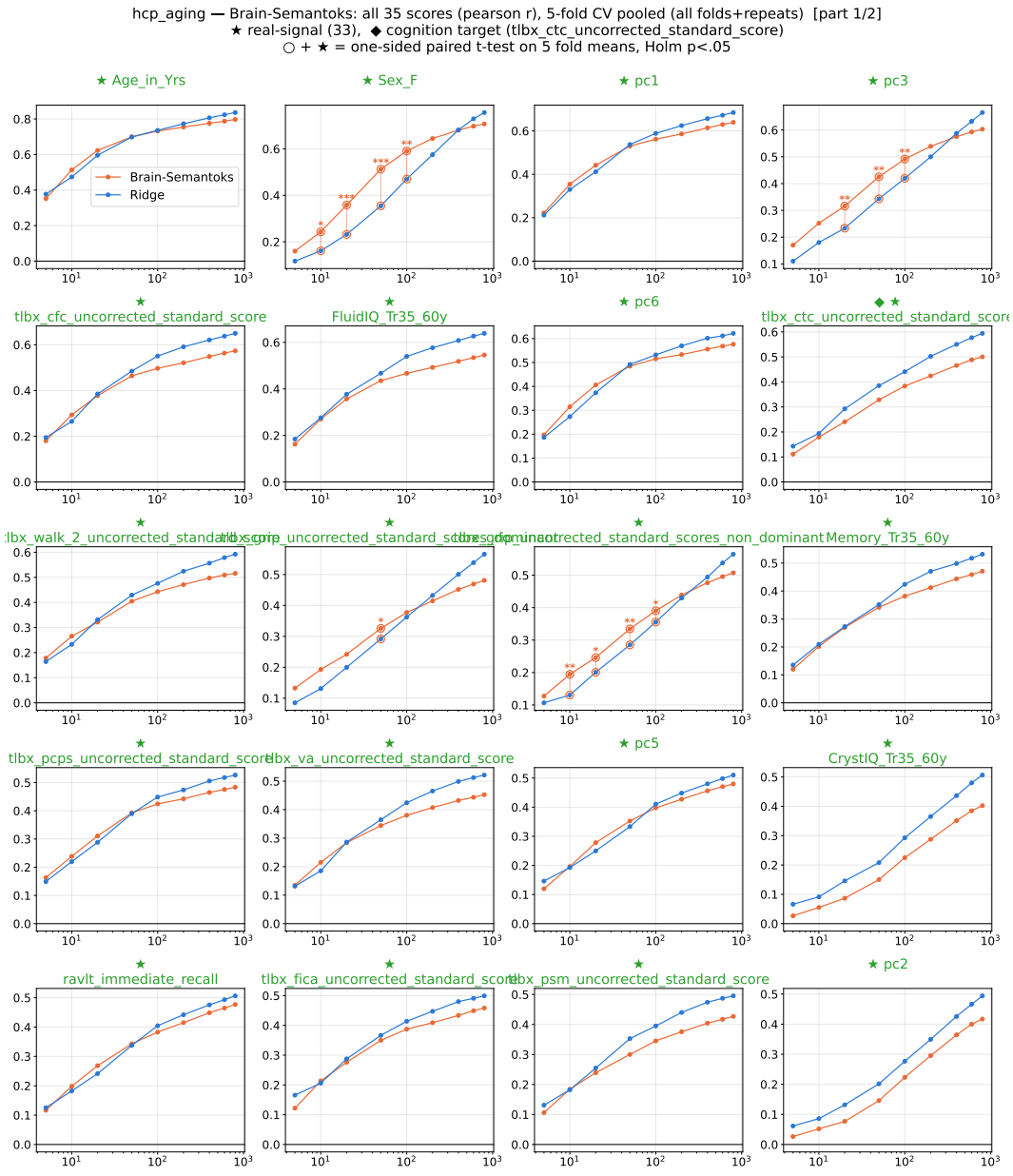


Figure 45: Brain-Semantoks, all scores, HCP-A, 5-fold CV (1/2).

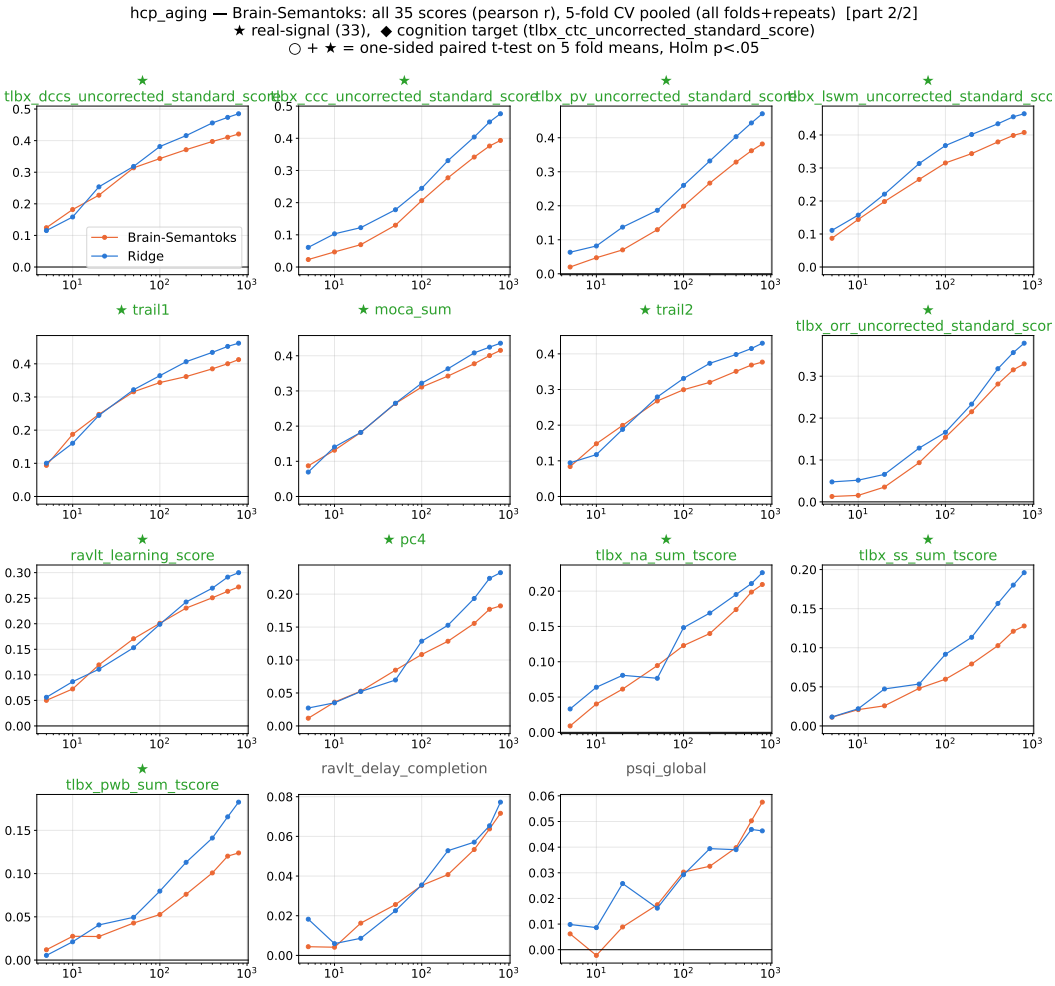


Figure 46: Brain-Semantoks, all scores, HCP-A, 5-fold CV (2/2).

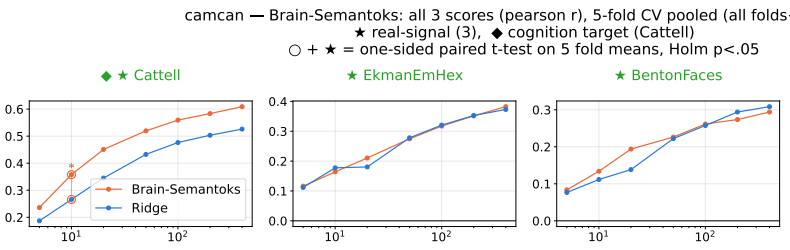


Figure 47: Brain-Semantoks, all scores, Cam-CAN, 5-fold CV.

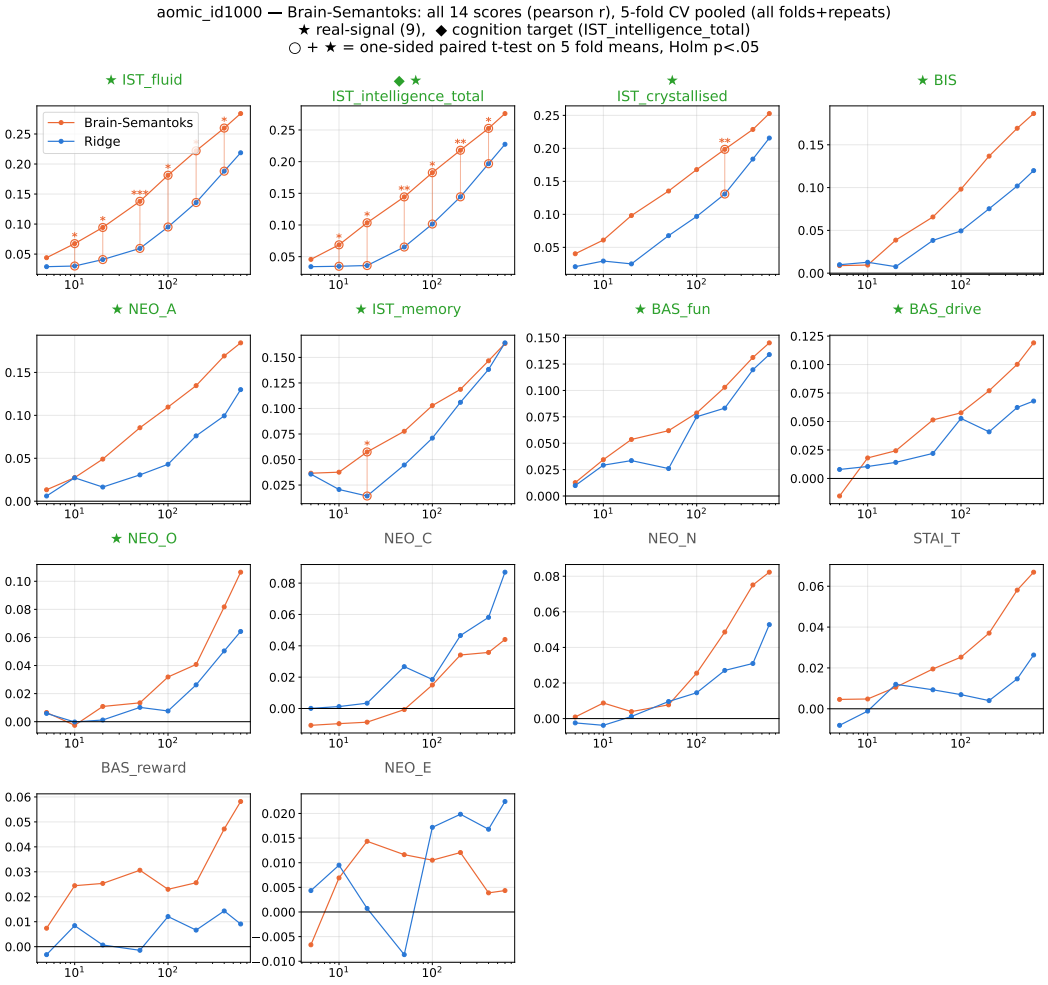


Figure 48: Brain-Semantoks, all scores, AOMIC-ID1000, 5-fold CV.

D.2.4 BRAIN-MASS

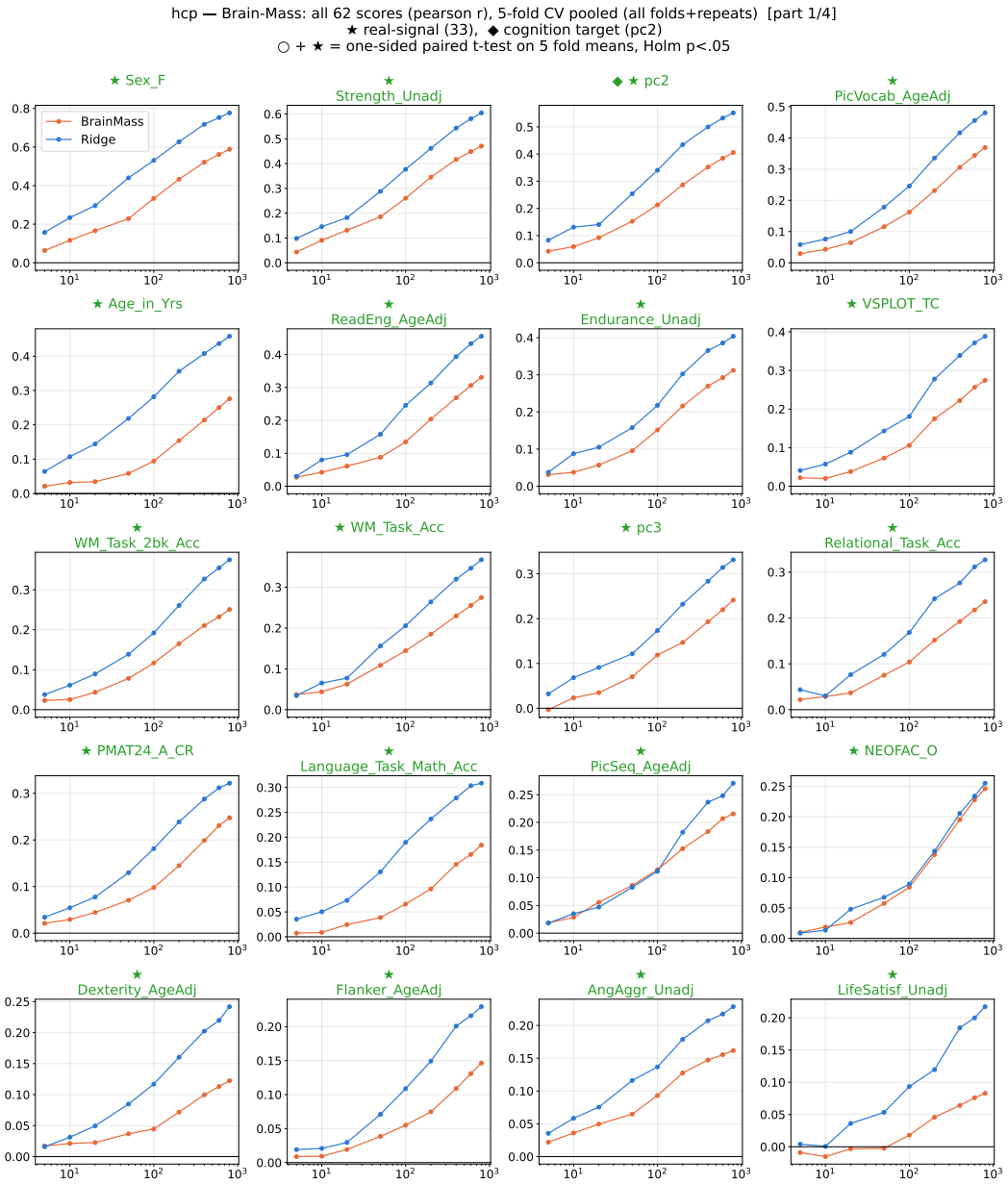


Figure 49: Brain-Mass, all scores, HCP-YA, 5-fold CV (1/3).

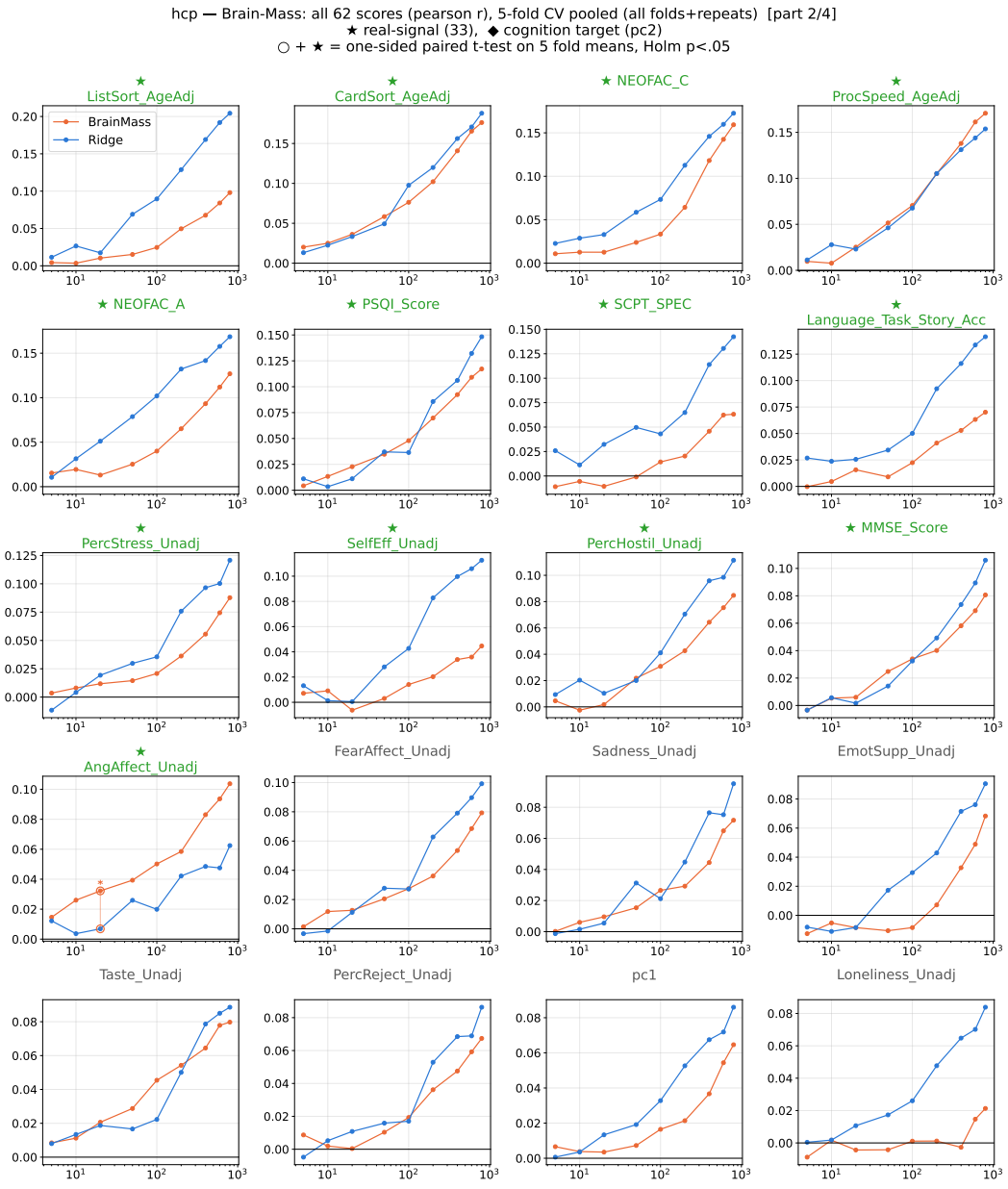


Figure 50: Brain-Mass, all scores, HCP-YA, 5-fold CV (2/3).



Figure 51: Brain-Mass, all scores, HCP-YA, 5-fold CV (3/3).

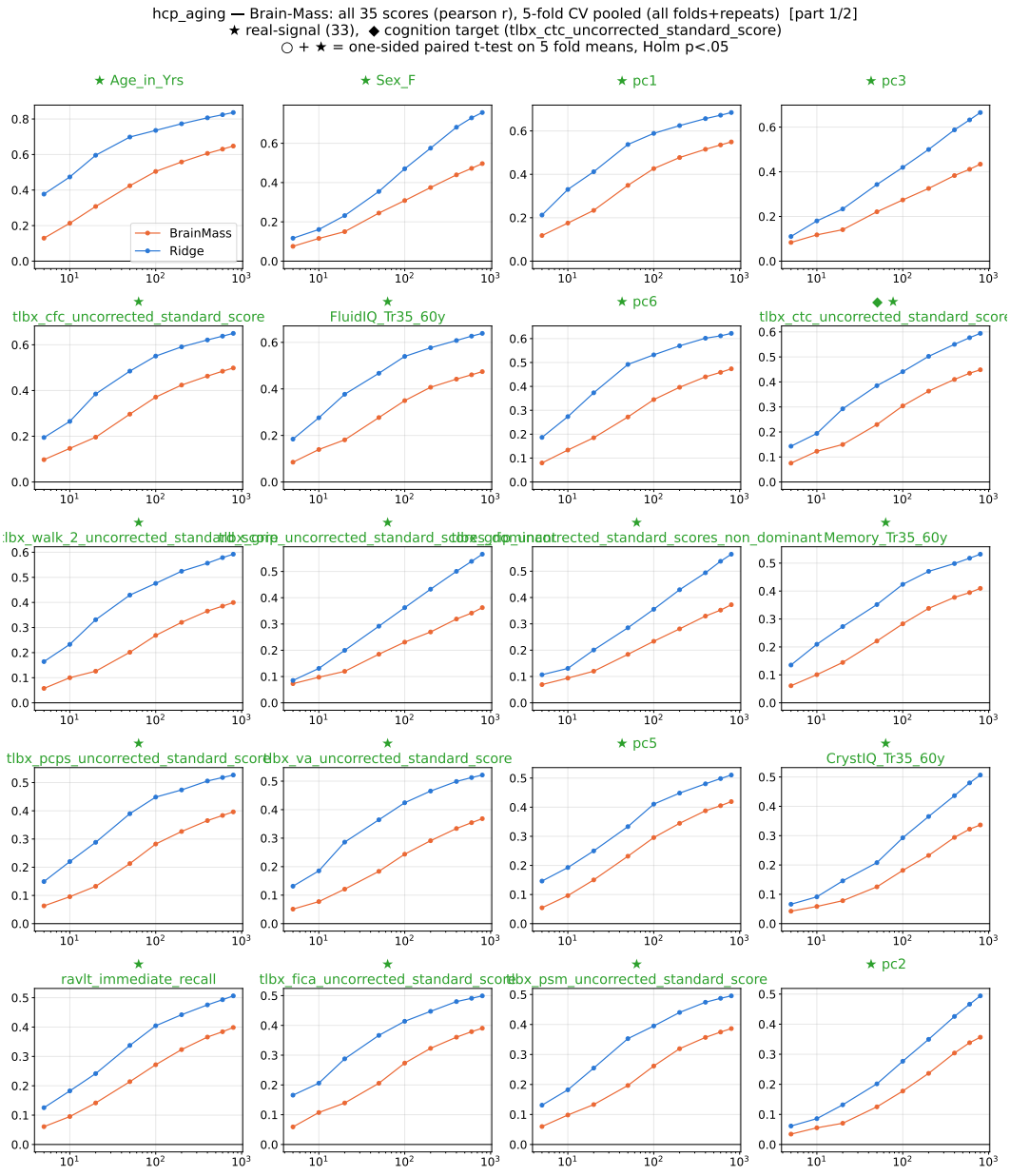


Figure 52: Brain-Mass, all scores, HCP-A, 5-fold CV (1/2).

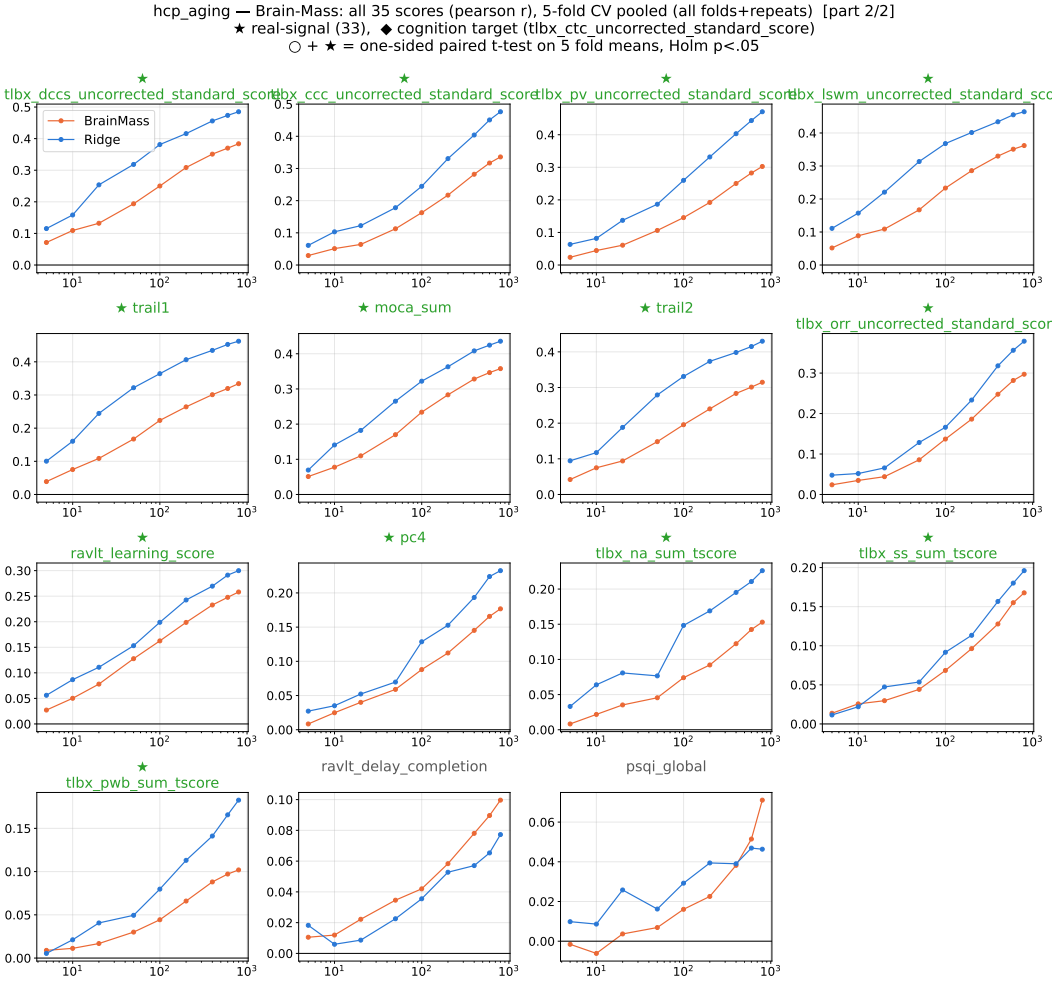


Figure 53: Brain-Mass, all scores, HCP-A, 5-fold CV (2/2).

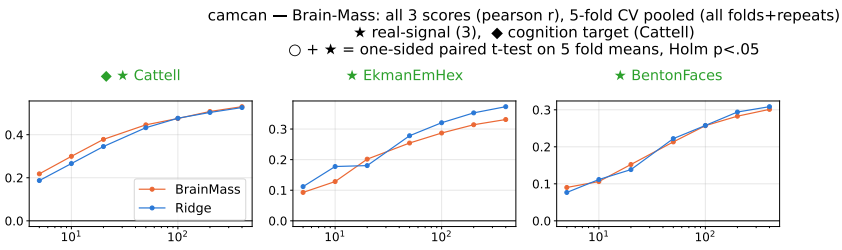


Figure 54: Brain-Mass, all scores, Cam-CAN, 5-fold CV.

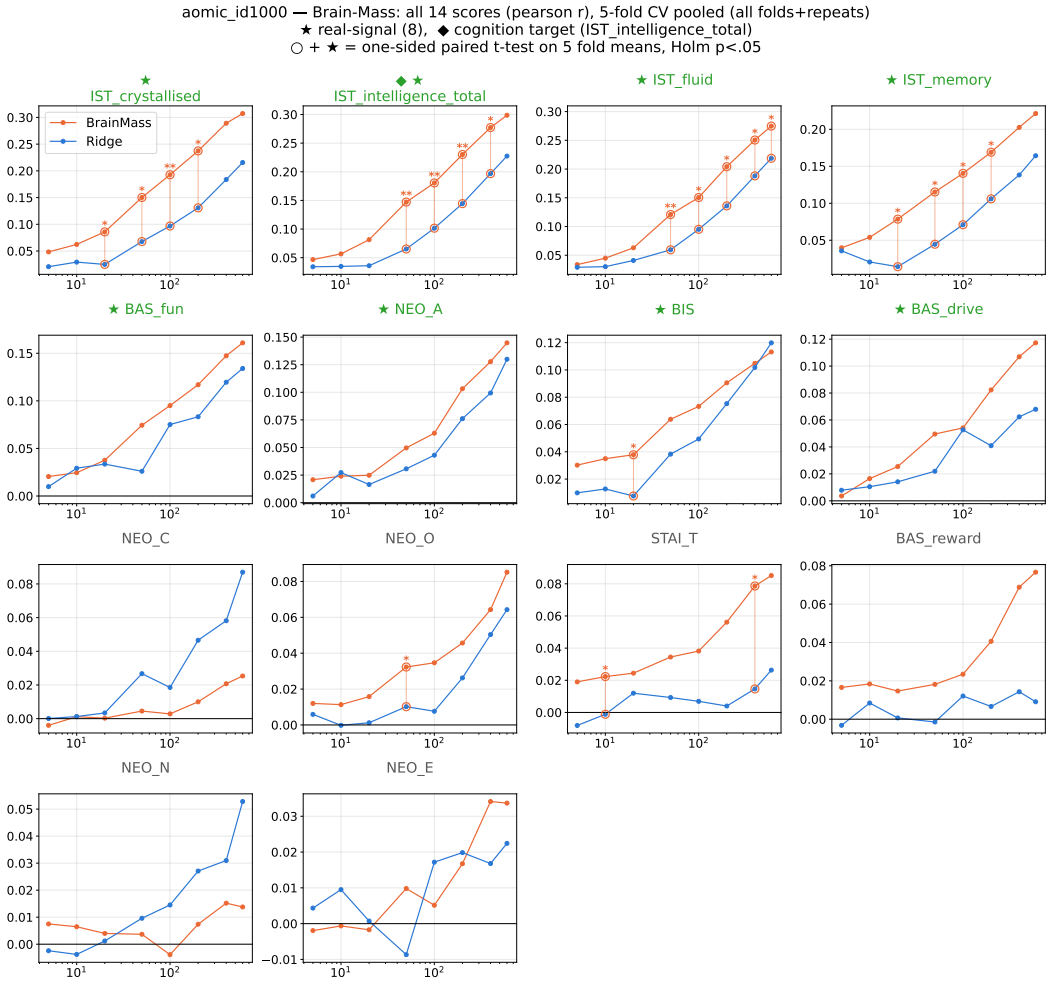


Figure 55: Brain-Mass, all scores, AOMIC-ID1000, 5-fold CV.

D.2.5 BRAINJEPA

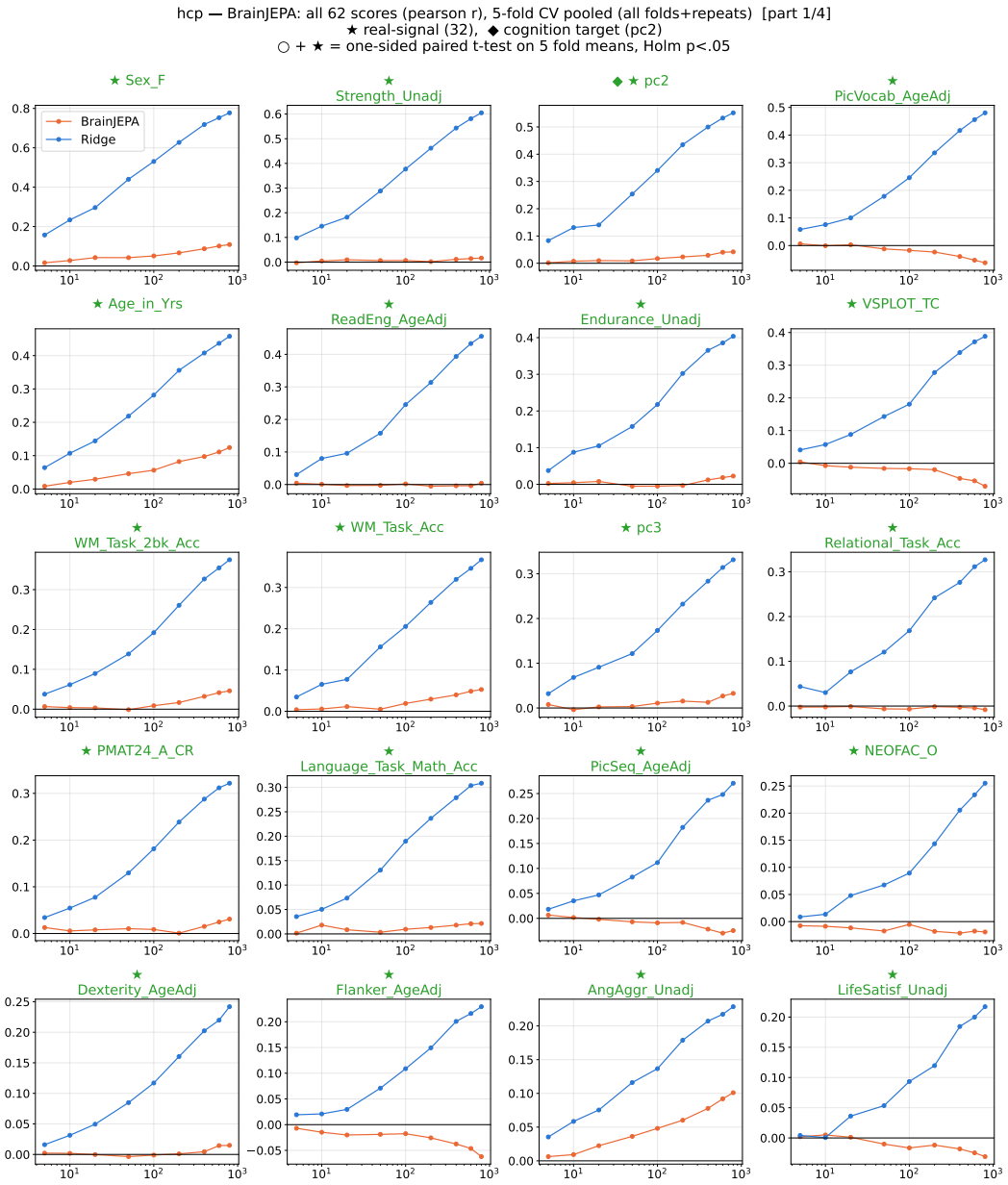


Figure 56: BrainJEPA, all scores, HCP-YA, 5-fold CV (1/3).

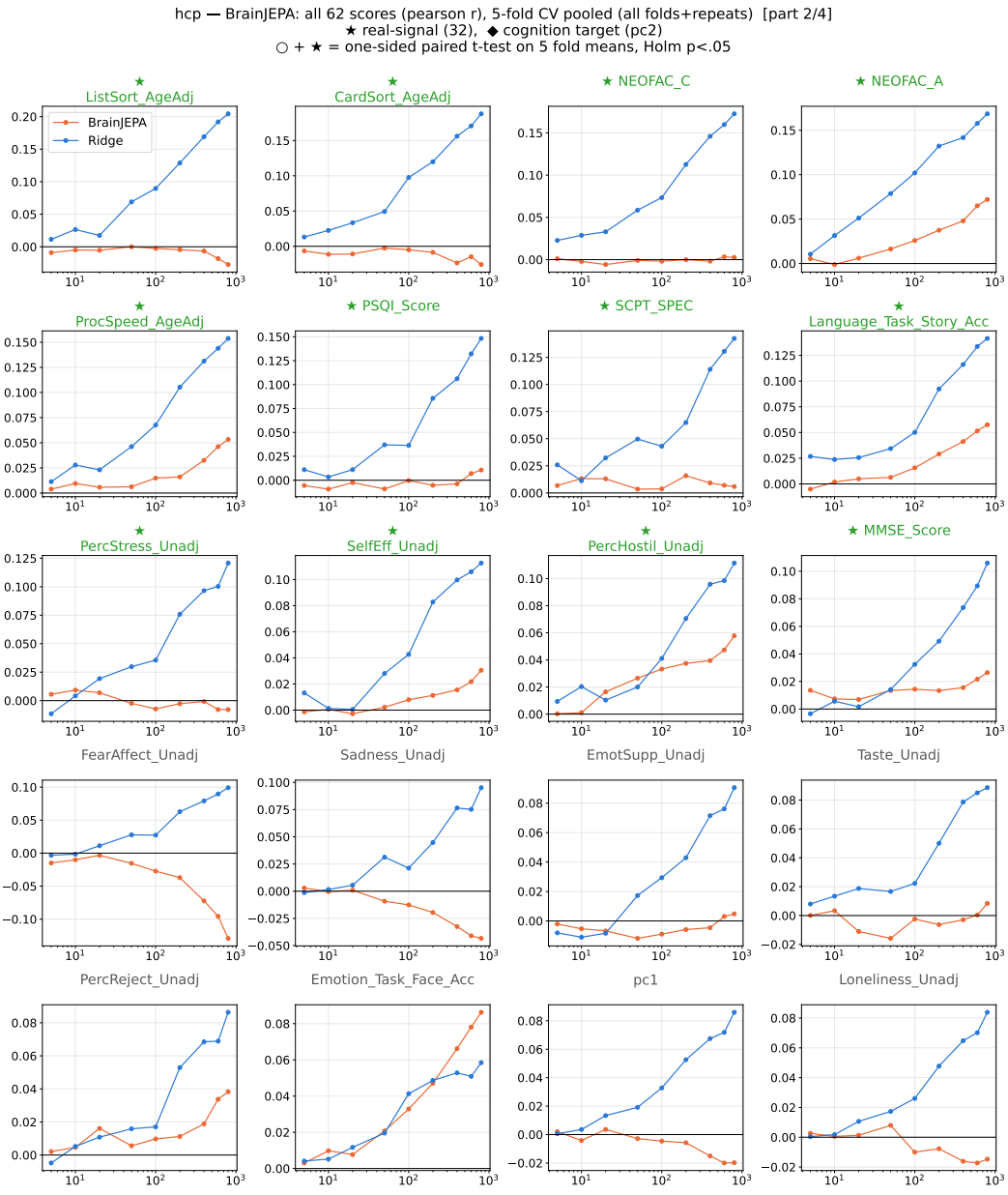


Figure 57: BrainJEPA, all scores, HCP-YA, 5-fold CV (2/3).



Figure 58: BrainJEPA, all scores, HCP-YA, 5-fold CV (3/3).

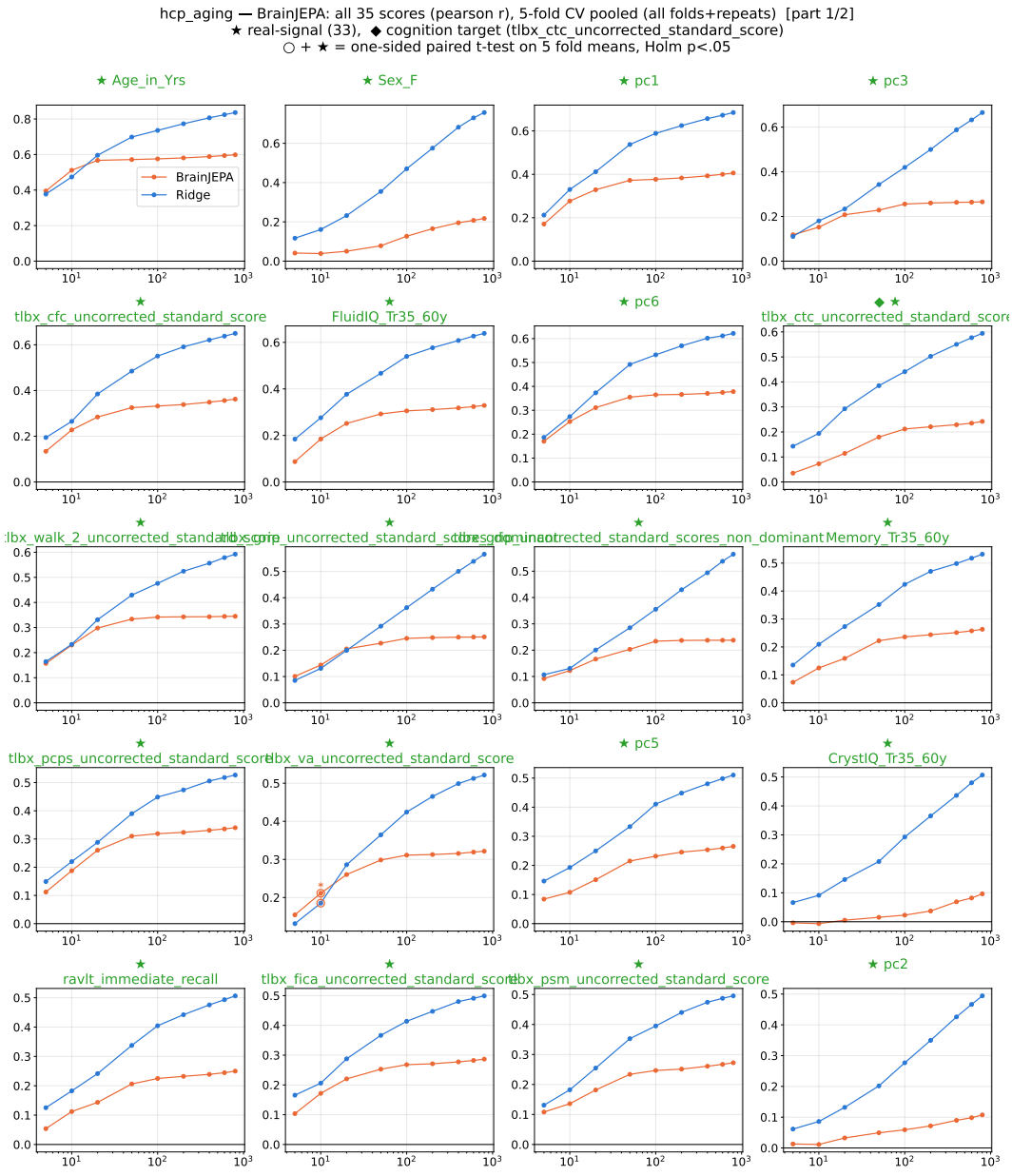


Figure 59: BrainJEPA, all scores, HCP-A, 5-fold CV (1/2).

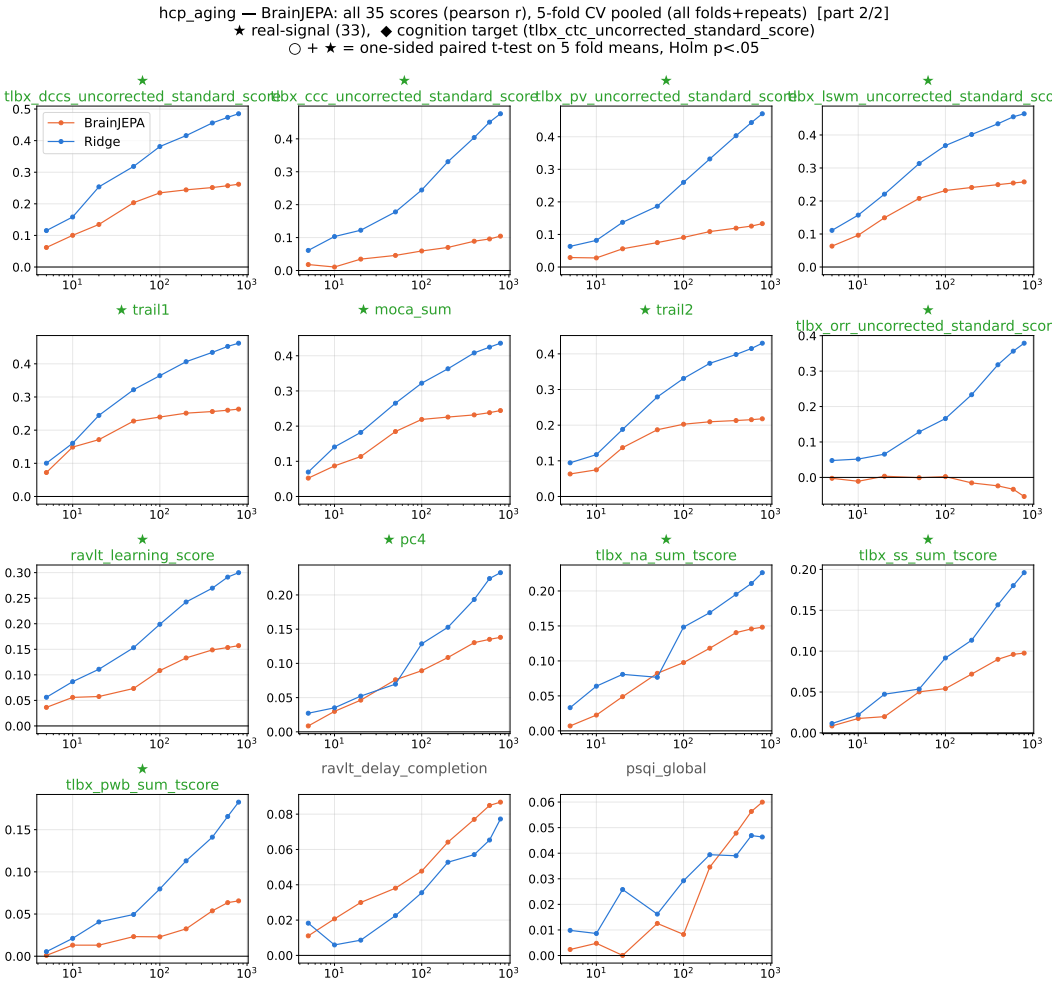


Figure 60: BrainJEPA, all scores, HCP-A, 5-fold CV (2/2).

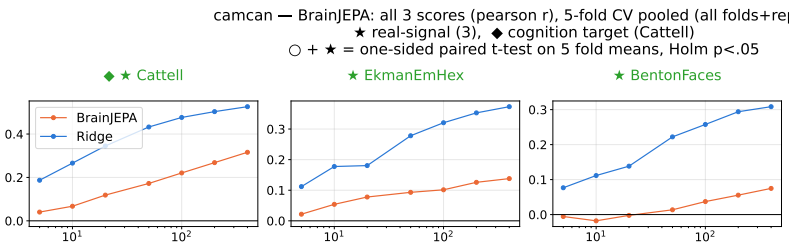


Figure 61: BrainJEPA, all scores, Cam-CAN, 5-fold CV.

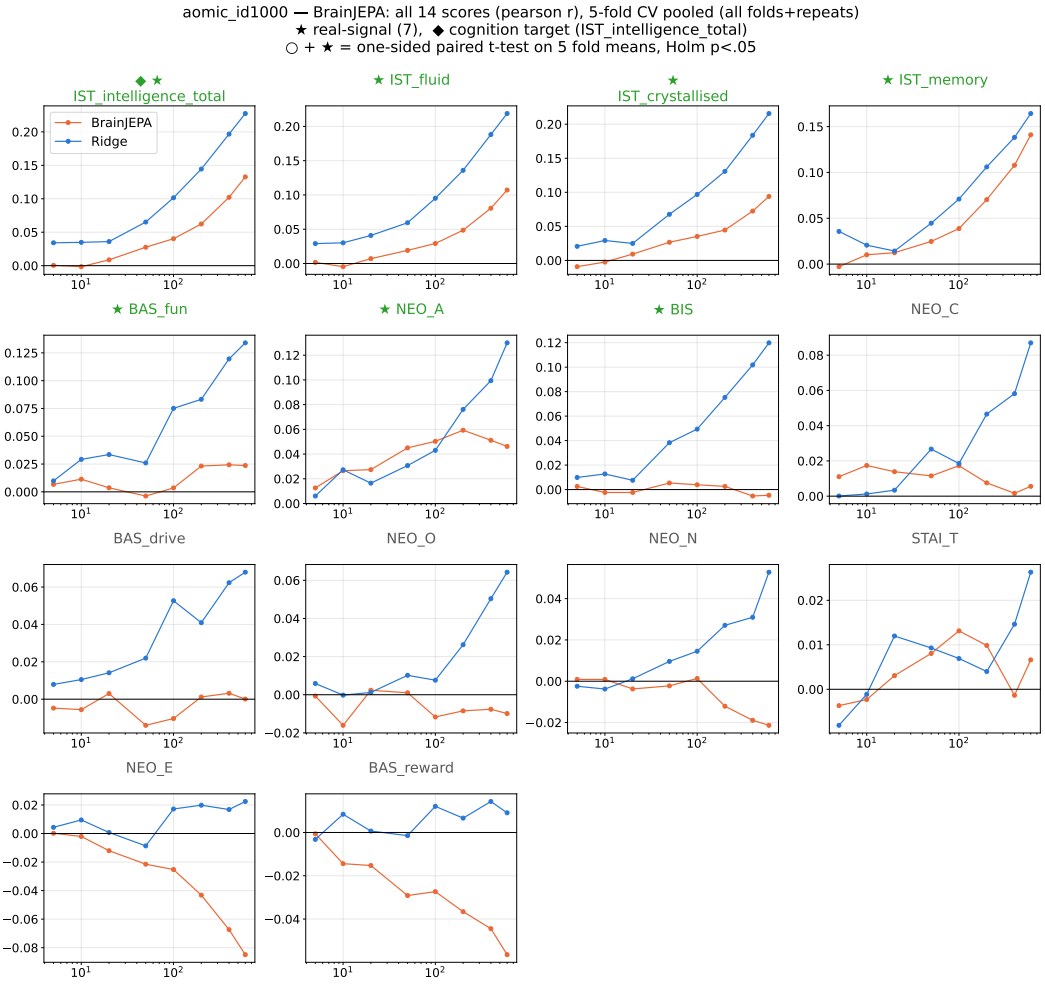


Figure 62: BrainJEPA, all scores, AOMIC-ID1000, 5-fold CV.

D.2.6 TABPFN V2

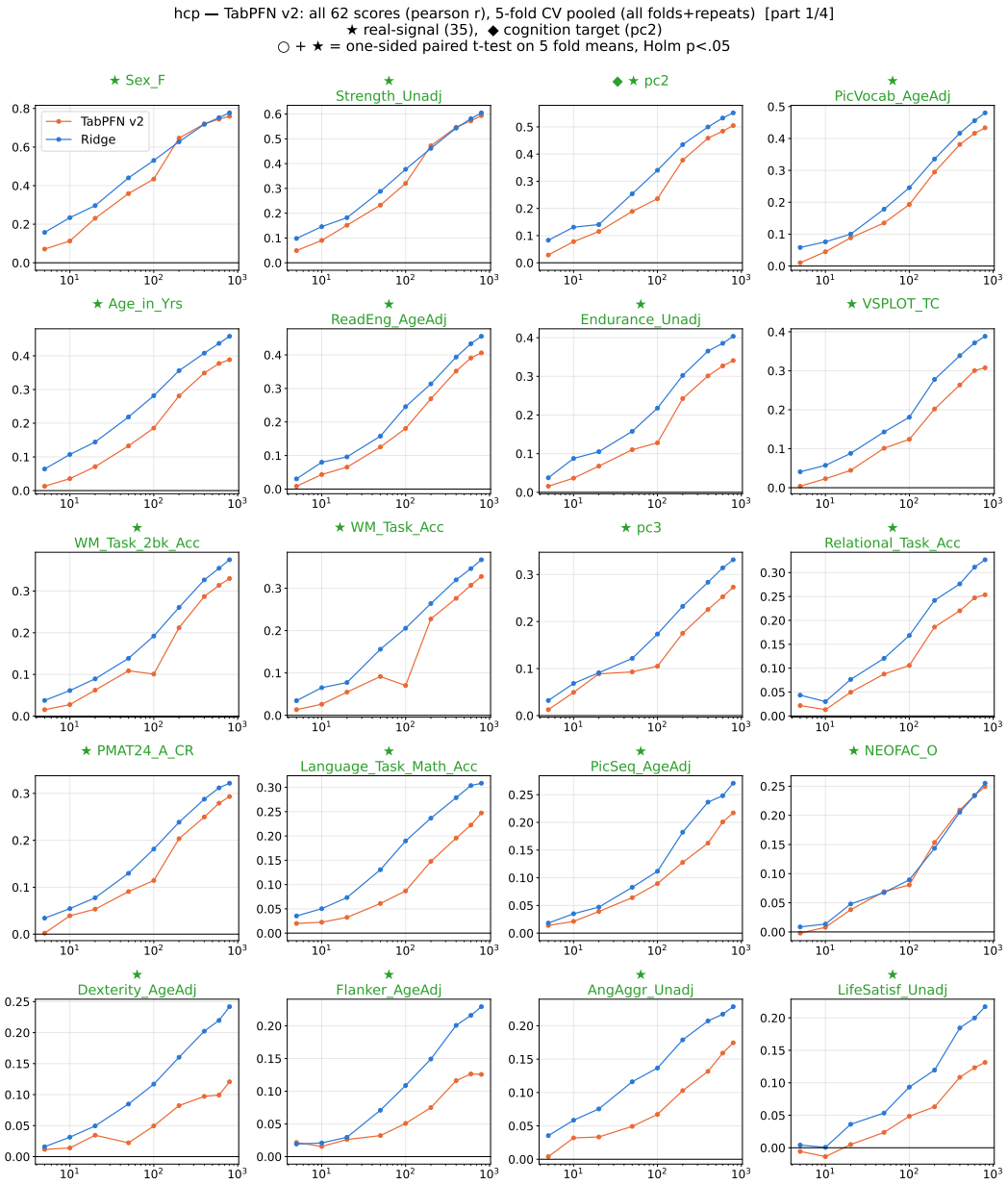


Figure 63: TabPFN v2, all scores, HCP-YA, 5-fold CV (1/3).

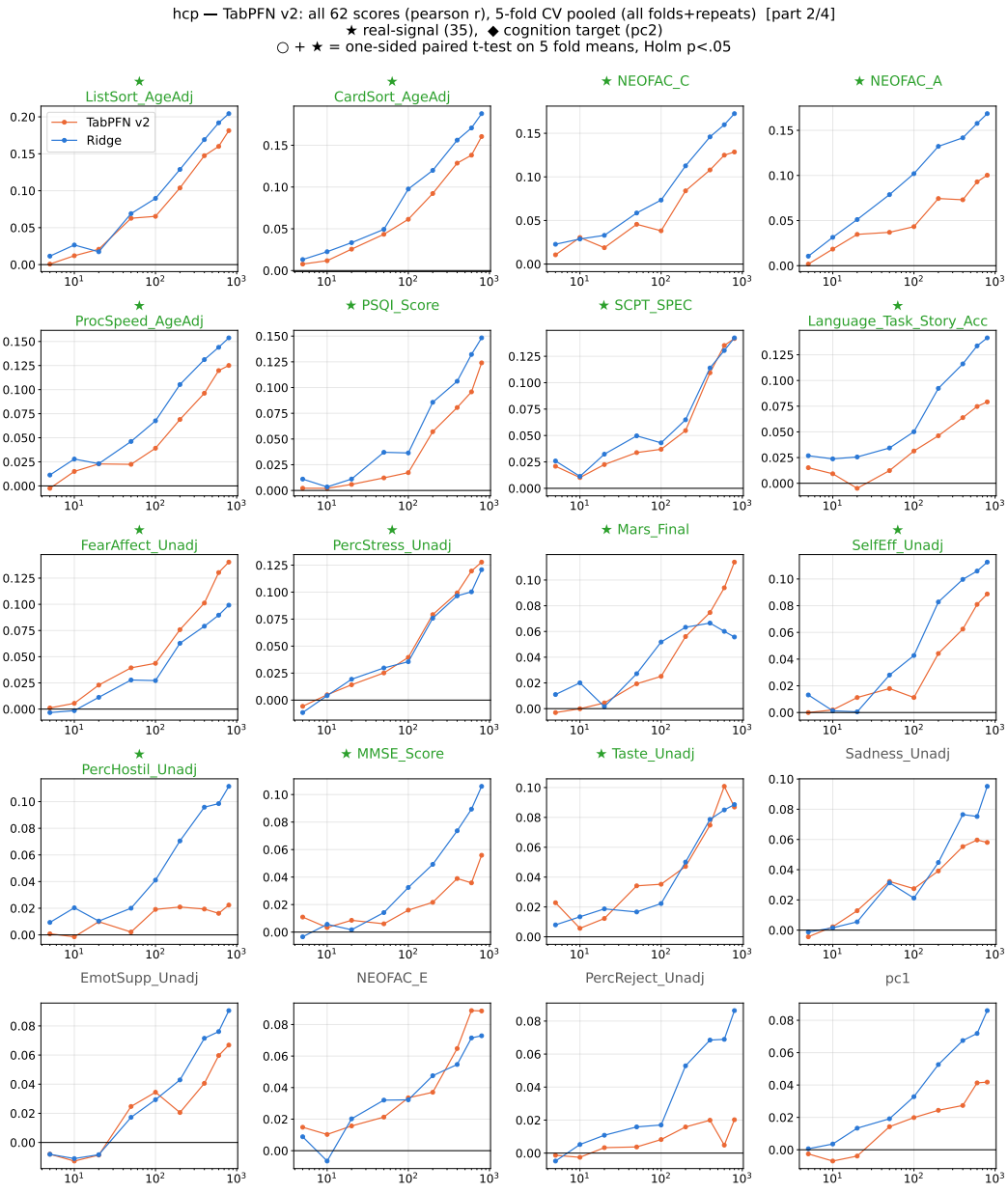


Figure 64: TabPFN v2, all scores, HCP-YA, 5-fold CV (2/3).



Figure 65: TabPFN v2, all scores, HCP-YA, 5-fold CV (3/3).

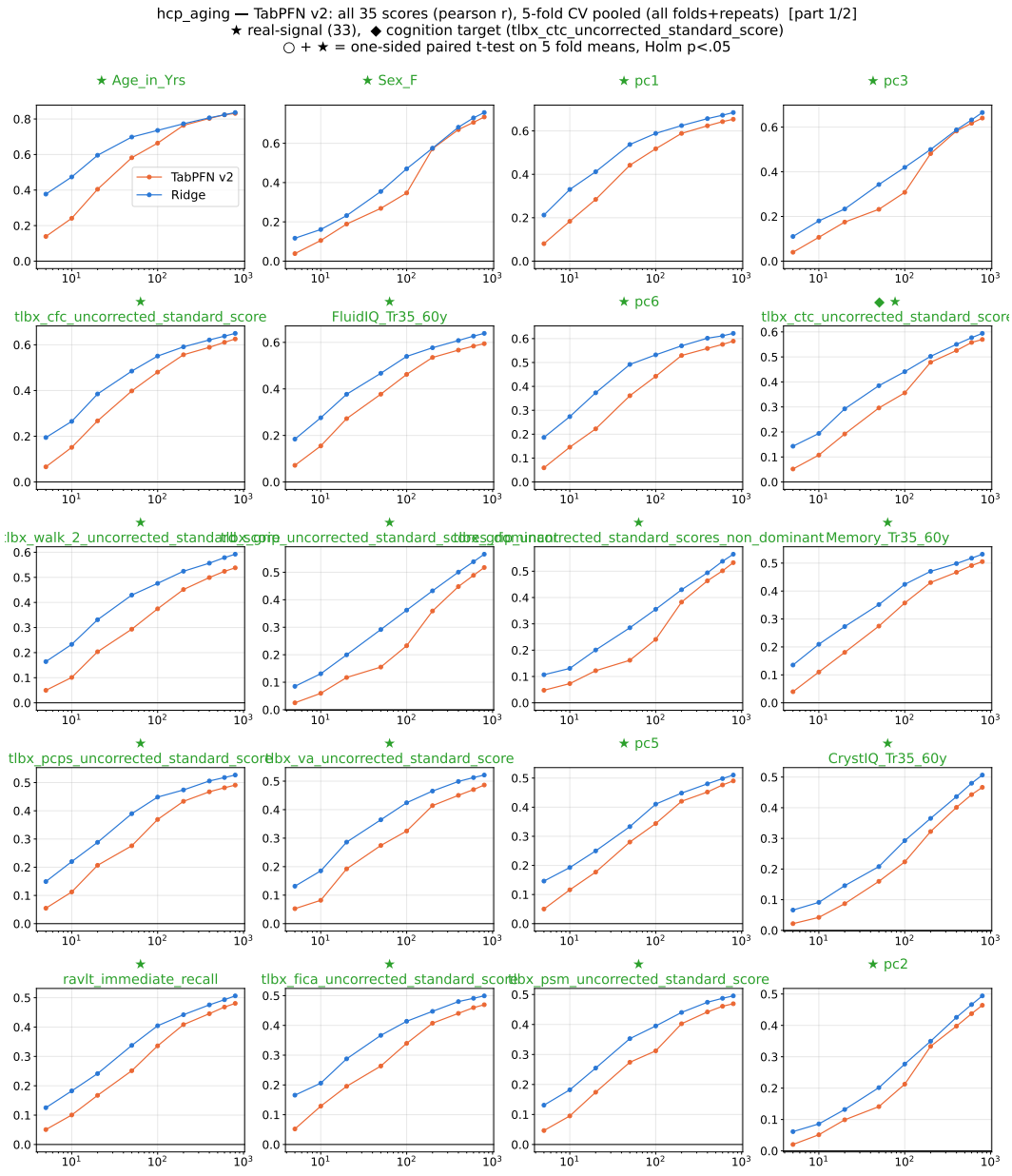


Figure 66: TabPFN v2, all scores, HCP-A, 5-fold CV (1/2).

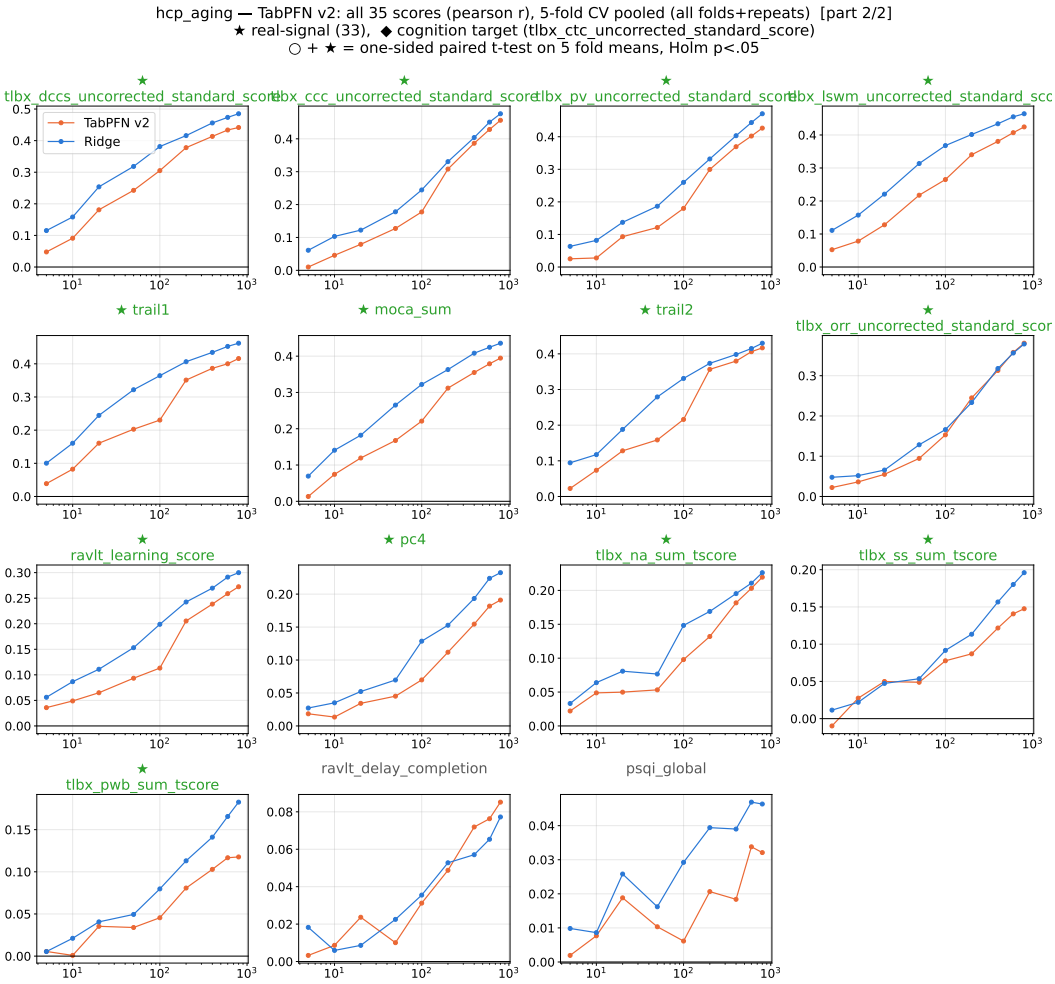


Figure 67: TabPFN v2, all scores, HCP-A, 5-fold CV (2/2).

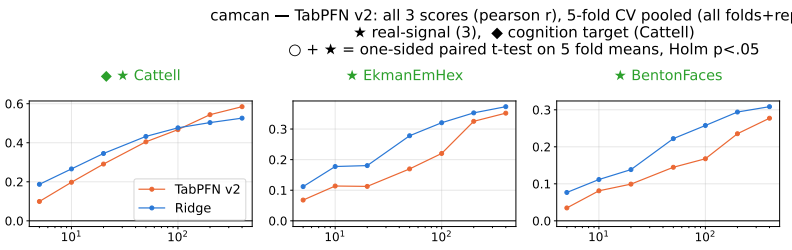


Figure 68: TabPFN v2, all scores, Cam-CAN, 5-fold CV.

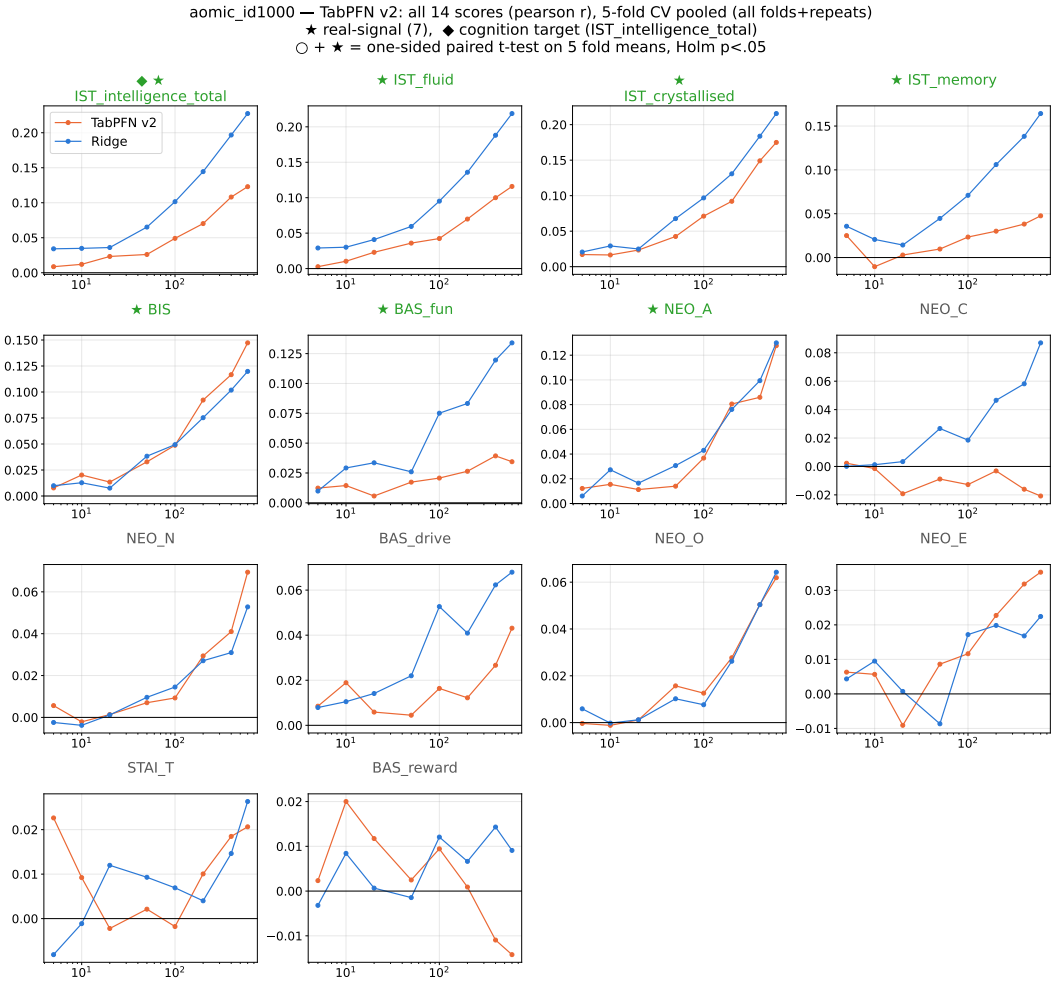


Figure 69: TabPFN v2, all scores, AOMIC-ID1000, 5-fold CV.

E LIST OF DATASETS

Table 8: OpenNeuro datasets included in the meta-training pool.

#	Accession	Name
1	ds000001	Balloon Analog Risk-taking Task
2	ds000002	Classification learning
3	ds000003	Rhyme judgment
4	ds000005	Mixed-gambles task
5	ds000006	ds000006
6	ds000007	Stop-signal task with spoken & manual responses
7	ds000008	Stop-signal task with unconditional and conditional stopping
8	ds000009	The generality of self-control
9	ds000011	Classification learning and tone-counting
10	ds000017	ds000017
11	ds000030	UCLA Consortium for Neuropsychiatric Phenomics LA5c Study
12	ds000101	Simon task
13	ds000102	Flanker task (event-related)
14	ds000107	Word and object processing
15	ds000108	Prefrontal-Subcortical Pathways Mediating Successful Emotion Regulation
16	ds000109	Social cognitive tasks (animate movement task, moral judgment task, false belief task)
17	ds000110	Incidental encoding task (Posner Cueing Paradigm)
18	ds000113	Forrest Gump
19	ds000114	A test-retest fMRI dataset for motor, language and spatial attention functions.
20	ds000115	Working memory in healthy and schizophrenic individuals
21	ds000116	Auditory and Visual Oddball EEG-fMRI
22	ds000119	Maturational Changes in Anterior Cingulate and Frontoparietal Recruitment Support the Development of Error Processing and Inhibitory Control (Antistate)
23	ds000120	Developmental changes in brain function underlying the influence of reward processing on inhibitory control (Slot Reward)
24	ds000121	Immaturities in Reward Processing and Its Influence on Inhibitory Control in Adolescence (Ring Reward)
25	ds000133	Modafinil alters intrinsic functional connectivity of the right posterior insula: a pharmacological resting state fMRI study
26	ds000140	Distinct brain systems mediate the effects of nociceptive input and self-regulation on pain
27	ds000144	Preschool Anxiety Disorders
28	ds000148	Who can afford self-control? The neural efficiency mechanism explains effective self-regulation of behavior
29	ds000149	Scene Perception
30	ds000157	Block design food and nonfood picture viewing task
31	ds000158	The human Voice Areas: spatial organisation and inter-individual variability in temporal and extra-temporal cortices
32	ds000164	Stroop Task
33	ds000168	Offline Processing in Associative Learning
34	ds000171	Neural Processing of Emotional Musical and Nonmusical Stimuli in Depression
35	ds000172	Physiological Contribution in Spontaneous Oscillations: An Approximate Quality - Assurance Index for Resting-State fMRI Signals
36	ds000200	Pre-adolescents Exposure to Manganese
37	ds000205	Affective Videos

#	Accession	Name
38	ds000208	Brain connectivity predicts placebo response across chronic pain clinical trials
39	ds000210	Multi-echo fMRI replication sample of autobiographical memory, prospection and theory of mind reasoning tasks
40	ds000212	Moral judgments of intentional and accidental moral violations across Harm and Purity domains
41	ds000213	Neural mechanism underlying appearance social comparison
42	ds000214	EUPD Cyberball
43	ds000218	Flavour Pleasantness (Oral Nutritional Supplements)
44	ds000219	Flavour Pleasantness (Regular Products)
45	ds000220	Cost Analysis TBI
46	ds000221	MPI-Leipzig_Mind-Brain-Body
47	ds000223	Magnitude Effect
48	ds000228	MRI data of 3-12 year old children and adults during viewing of a short animated film
49	ds000233	Neural responses to naturalistic clips of behaving animals in two different task contexts
50	ds000243	Washington University 120
51	ds000245	ds000245_R1.0.0
52	ds000249	Value generalization in human avoidance learning
53	ds000253	Female action video game players.
54	ds000254	Multiband Multi-Echo Simultaneous ASL/BOLD for task-induced functional MRI
55	ds000256	Behavioral interventions for reducing head motion during MRI scans in children
56	ds000258	Multi-echo Cambridge
57	ds001021	NKI-sample
58	ds001131	Milky-Vodka
59	ds001132	Sherlock
60	ds001145	Twilight Zone Movie Watching Dataset
61	ds001226	BTC_preop
62	ds001229	Neural Evidence of the Strategic Choice Between Working Memory and Episodic Memory in Prospective Remembering.
63	ds001232	Competition between items in working memory leads to forgetting
64	ds001235	Human-Object Interaction
65	ds001242	Examining effects of arousal on responses to salient and non-salient stimuli in younger and older adults
66	ds001297	Personally familiar faces
67	ds001302	Inverse Base Rate
68	ds001338	ParanoiaStory
69	ds001344	phfMRI_Control_Tasks
70	ds001345	Decoding of multisensory semantics and memories in low-level visual cortex
71	ds001353	Information Filtering in Electronic Networks of Practice: An fMRI Investigation of Expectation [Dis]confirmation
72	ds001357	NameFacePriming
73	ds001379	Loci case study
74	ds001399	multibandACQtests
75	ds001408	rsfMRI_single_session_EEG_NF
76	ds001419	thoughtExperiment
77	ds001430	Neural Overlap in Item Representations Across Episodes Impairs Context Memory
78	ds001439	A role for the medial temporal lobe subsystem in guiding prosociality: the effect of episodic processes on willingness to help others

#	Accession	Name
79	ds001454	Human hippocampal replay during rest prioritizes weakly learned information and predicts memory performance
80	ds001491	Valence processing differs across stimulus modalities (Multi-echo)
81	ds001521	Lingering representations of stimuli influence recall organization
82	ds001534	Calorie-labeled food cues
83	ds001545	Learning Naturalistic Temporal Structure in the Posterior Medial Network
84	ds001553	100 runs at 3T
85	ds001554	Neural pattern change during encoding of a narrative predicts retrospective duration estimates
86	ds001566	test
87	ds001576	Refresh my memory: Episodic memory reinstatements intrude on working memory maintenance
88	ds001607	Reinstated episodic context guides sampling-based decisions for reward
89	ds001608	Prestimulus Exp2
90	ds001612	Offline replay supports planning in human reinforcement learning
91	ds001621	Neural representations of events arise from temporal community structure
92	ds001634	Visual working memory: Study one Task fMRI and Behavioural response
93	ds001635	Visual working memory: Study two Task fMRI and Behavioural response
94	ds001715	Dissociable neural mechanisms track evidence accumulation for selection of attention versus action
95	ds001722	A role for the medial temporal lobe subsystem in guiding prosociality: the effect of episodic processes on willingness to help others (Experiment 2)
96	ds001728	Test-GEIER
97	ds001740	Brain activity during reciprocal social interaction investigated using conversational robots as control condition
98	ds001745	A neural signature of contextually mediated intentional forgetting
99	ds001747	Exploring the Resting State Neural Activity of Monolinguals and Late and Early Bilinguals
100	ds001748	Differentiation of functional networks during long-term memory retrieval in children and adolescents
101	ds001762	Dorsal anterior cingulate and ventromedial prefrontal cortex have inverse roles in both foraging and economic choice
102	ds001775	Neural Differentiation Tracks Improved Recall of Competing Memories Following Interleaved Study and Retrieval Practice
103	ds001796	Bilingualism and the brain
104	ds001838	Handedness and Symbolic Number Representation
105	ds001840	Prestimulus Exp1
106	ds001847	A Probability Distribution over Latent Causes, in the Orbitofrontal Cortex
107	ds001912	S-cone_MotionInDepth_fMRI
108	ds001921	Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value
109	ds001923	Anterior cingulate engagement in a foraging context reflects choice difficulty, not foraging value (2)
110	ds001927	The physiological effects of non-invasive brain stimulation fundamentally differ across the human cortex

#	Accession	Name
111	ds001928	Functional Connectivity of Music-Induced Analgesia in Fibromyalgia
112	ds002011	Locus coeruleus activity strengthens prioritized memories under arousal
113	ds002014	Language Production fMRI
114	ds002041	Mesolimbic dopamine D2 receptors and neural representations of subjective value
115	ds002080	BTC_postop
116	ds002149	Conditional Visual Associative Learning Task
117	ds002232	Caltech rsfMRI Dataset
118	ds002250	When my wrongs are worse than yours: behavioral and neural asymmetries in first-person and third-person perspectives of accidents
119	ds002270	Viewing_food_cues
120	ds002322	The Alice Dataset: fMRI Dataset to Study Natural Language Comprehension in the Brain
121	ds002345	Narratives
122	ds002363	Repetition of Computer Security Warnings Results in Differential Repetition Suppression Effects as Revealed with Functional MRI
123	ds002522	Contrast Response Functions
124	ds002547	SharedStates
125	ds002603	MB-EPI-comparison
126	ds002609	RPN-signature_Study2
127	ds002643	Can we have a second helping? A replication study on the neurobiological mechanisms underlying self-control
128	ds002711	Smiley
129	ds002715	AVScenes_Blind
130	ds002717	Unexplained Repeated Pregnancy Loss is Associated with Altered Perceptual and Brain Responses to Men's Body-Odor
131	ds002731	Multiple interactive memory representations underlie the induction of false memory
132	ds002732	Unnamed Dataset
133	ds002734	Evidence Accumulation in Value-Based decisions
134	ds002737	multibandCFtests
135	ds002748	Resting state with closed eyes for patients with depression and healthy participants
136	ds002750	YanDataBIDS
137	ds002785	AOMIC-PIOP1
138	ds002790	AOMIC-PIOP2
139	ds002793	ToM following strong vs weak priors
140	ds002797	eLife39497
141	ds002835	Prospection
142	ds002837	Naturalistic Neuroimaging Database
143	ds002841	Intuitive physics with fMRI
144	ds002843	Cognitive Training
145	ds002848	mriquality
146	ds002872	Aging
147	ds002896	Resting State fMRI study of non bothersome tinnitus
148	ds002938	Task-Dependent Effort-Induced Connectivity
149	ds003007	Two sessions of resting state with closed eyes for patients with depression in treatment course (NFB, CBT or No treatment groups)
150	ds003011	Social Processes Initiative in Neurobiology of the Schizophrenia(s) Traveling Human Phantoms
151	ds003012	action-in-auctions
152	ds003037	SUDMEX_TMS

#	Accession	Name
153	ds003076	Brain Correlates of Deductive Reasoning in Adults
154	ds003083	Brain Correlates of Math Processing in Adults
155	ds003085	Temporal Dynamics of Emotional Music
156	ds003096	Valuation of peers' safe choices is associated with substance-naivete in adolescents'
157	ds003103	Consensus-seeking and conflict-resolving: an fMRI study on college couples' shopping interaction
158	ds003114	How ovarian hormones influence the behaviroal activation and inhibition system through the dopamine pathway
159	ds003126	Reading-related functional activity in children with isolated spelling deficits and dyslexia
160	ds003146	vocal task
161	ds003171	Modeling an auditory stimulated brain under altered states of consciousness using the generalized ising model
162	ds003346	SUDMEX_CONN: The Mexican dataset of cocaine use disorder patients.
163	ds003354	EmpaToM with PPI-R, TAS, and EQ
164	ds003381	Estimation of in-scanner head pose changes during structural MRI using a convolutional neural network trained on eye tracker video
165	ds003382	Monash vis-fPET-fMRI
166	ds003425	Mental imagery can generate and regulate acquired differential fear conditioned reactivity
167	ds003427	Effects of Phase Regression on High-Resolution Functional MRI of the Primary Visual Cortex
168	ds003436	Agreeableness personality trait and social information encoding
169	ds003459	Component processes of word reading in adults and children
170	ds003466	Haptic three-dimensional curved surface exploration fMRI dataset
171	ds003469	Speech disfluencies: Neurophysiological aspect in normal population
172	ds003495	dStreamUpgrade
173	ds003500	Response inhibition and selective attention in adults and children with and without ADHD
174	ds003542	Adult language learners
175	ds003545	Adolescent language learners
176	ds003548	Emotion Category and Face Perception Task Optimized for Multivariate Pattern Analysis
177	ds003550	Visser, Scholte, Beemsterboer, & Kindt (2013) Nature Neuroscience
178	ds003553	Visser, Scholte & Kindt (2011) Journal of Neuroscience
179	ds003554	Visser et al. 2015 Psychoneuroendocrinology
180	ds003574	Reward biases spontaneous neural reactivation during sleep
181	ds003592	Neurocognitive aging data release with behavioral, structural, and multi-echo functional MRI measures
182	ds003688	Open multimodal iEEG-fMRI dataset from naturalistic stimulation with a short audiovisual film
183	ds003709	NIMH-CompyPsych MMI
184	ds003763	Interoception during aging: The heartbeat detection task
185	ds003770	Depression Treatment by Medial Prefrontal Real-Time fMRI Neurofeedback
186	ds003772	Functionally dissociable influences on learning rate in a dynamic environment
187	ds003789	Dynamic changes in neural representations underlie the repetition effect on false memory
188	ds003831	Cognitive Control Theoretic Mechanisms of Real-time fMRI-Guided Neuromodulation (CTM)

#	Accession	Name
189	ds003835	PE-Update
190	ds003849	Individual differences in frontoparietal plasticity in humans
191	ds003858	Monetary Incentive Delay task
192	ds003871	Resting-state for 34 younger and 28 older adults
193	ds003892	Loss and enhancement of layer-selective signals in geniculostriate and corticotectal pathways of adult human amblyopia
194	ds003965	NI-edu-data-minimal
195	ds003972	Body Motor
196	ds003990	Fast Event-Related Mapping of Population Fingertip Tuning Properties in Human Sensorimotor Cortex at 7T
197	ds003999	Pre-Post rehabilitation fMRI data of post-stroke patients.
198	ds004021	Visuomotor Rotation Adaptation Experiment
199	ds004101	Resting-state fMRI data acquired before and after a course of infra-low frequency neurofeedback
200	ds004128	Neural representations of others' traits predict social decisions
201	ds004134	Experience Sampling in Resting State
202	ds004141	Targeting Dynamic Facial Processing Mechanisms in Superior Temporal Sulcus Using fMRI Neurofeedback
203	ds004142	Exploring the Neural Correlates of Feedback-Related Reward Saliency and Valence During Real-Time fMRI-Based Neurofeedback
204	ds004144	A behavioral, clinical and brain imaging dataset with focus on emotion regulation of females with fibromyalgia
205	ds004182	Magic, Memory, and Curiosity (MMC) fMRI Dataset
206	ds004196	Bimodal dataset on Inner speech
207	ds004213	BrainBeats
208	ds004219	Litebook_Alertness_Study
209	ds004302	Brain correlates of speech perception in schizophrenia patients with and without auditory hallucinations
210	ds004392	Parkinson's disease, functional connectivity, and cognition
211	ds004400	Number comparison task
212	ds004440	Stanford Child and Adult Checkerboard Retinotopy Dataset
213	ds004450	XCP_WALKTHROUGH
214	ds004467	MRI data from 20 adults in response to videos of dialogue and monologue from Sesame Street
215	ds004469	Temporal Lobe Epilepsy - UNAM
216	ds004498	Perinatal Stroke
217	ds004544	An fMRI Investigation of the Neurocognitive Processing of Strategies and Measures
218	ds004557	The role of superstition of cognitive control during neurofeedback training - part 2
219	ds004562	MVPA for motor adaptation
220	ds004564	Demo
221	ds004594	Brain surface fingerprints
222	ds004604	Cerebrovascular Reactivity Normative Dataset
223	ds004627	Characterization and Treatment of Adolescent Depression (CAT-D)
224	ds004648	Sympathovagal quotient and functional connectivity of control networks are related to gut Ruminococcaceae abundance in healthy men
225	ds004656	Neural reactivity to visual food stimuli is reduced in some areas of the brain during evening hours compared to morning hours: an fMRI study in women
226	ds004662	Denoising task-correlated head motion from motor-task fMRI data with multi-echo ICA
227	ds004697	Food and Brain Study

#	Accession	Name
228	ds004718	Le Petit Prince Hong Kong: Naturalistic fMRI and EEG dataset from older Cantonese speakers
229	ds004720	Descriptive name for this dataset
230	ds004737	CS-DSI
231	ds004765	language fMRI
232	ds004775	Vicarious Punishment of Moral Violations in Naturalistic Drama Narratives Predicts Cortical Synchronization
233	ds004783	Functional MRI dataset for evaluating brain synchrony among dance spectators
234	ds004798	fMRI data for: Multimodal single-neuron, intracranial EEG, and fMRI brain responses during movie watching in human patients
235	ds004835	Processing of frequency and location in human subcortical auditory structures
236	ds004848	Game of Thrones - A naturalistic viewing dataset
237	ds004873	BOLD signal changes can oppose oxygen metabolism across the human cortex
238	ds004956	Braille letters - fMRI
239	ds005027	Michigan Longitudinal Study
240	ds005040	Political Moralization (PMM)
241	ds005069	test
242	ds005072	test
243	ds005075	Simultaneous Brain and Spinal Cord at rest: an fMRI dataset
244	ds005128	Deception Signaling Task
245	ds005134	Sequence test for multiband DWI and magnetization transfer sequence
246	ds005166	A framework for leveraging technology to probe mechanism: A proof of concept study of inhibitory control
247	ds005230	Reinforcement-Learning Generalization
248	ds005239	motor-fmri
249	ds005270	BOLD variability during cognitive control for an adult lifespan sample
250	ds005339	Yale_reading_dataset
251	ds005355	BSC_BIDS_071724
252	ds005365	Neurobiological substrates of Altered State of Consciousness Induced by High Ventilation Breathwork
253	ds005418	Cerebrovascular Reactivity: CO2 inhalation and resting state comparisons
254	ds005454	fMRIPrep boot camp – 7T training dataset
255	ds005455	An fMRI dataset for investigating language control and cognitive control in bilinguals
256	ds005479	Monetary Incentive Delay task - structural and functional images of 37 men; study of associations between circadian characteristics (eveningness, distinctness) and affective processing
257	ds005504	Psychosocial adversity and inhibitory control: an fMRI study of children growing up in extreme poverty
258	ds005518	Deciphering the neural responses to a naturalistic persuasive message
259	ds005525	Circadian misalignment and energy balance
260	ds005530	Depotentiation of emotional reactivity using TMR during REM sleep
261	ds005533	High-quality multimodal MRI with simultaneous EEG using conductive ink and polymer-thick film nets
262	ds005573	A comparison of resting state functional magnetic resonance imaging to invasive electrocortical stimulation for sensorimotor mapping in pediatric patients
263	ds005588	SoDec - Responsibility fMRI experiment dataset
264	ds005596	A naturalistic fMRI dataset in response to puclic speaking

#	Accession	Name
265	ds005598	Visuomotor Rotation Learning and Reward-based motor learning
266	ds005603	Mitchell_Hacker_2013
267	ds005623	test
268	ds005625	Connectomix test dataset 1
269	ds005699	Connectomix test dataset 2
270	ds005754	Multivariate Assessment of Inhibitory Control in Youth: Links with Psychopathology and Brain Function Dataset
271	ds005783	Representation of navigational affordances and ego-motion in the occipital place area - dataset
272	ds005795	MULTI-CLARID (Multimodal Category Learning and Resting-state Imaging Data)
273	ds005850	Spoken narrative comprehension in Chinese
274	ds005882	Integration of contextual reward motivation with category representations
275	ds005883	CoSpine Database_pain_dataset
276	ds005884	CoSpine Database_motor_dataset
277	ds005892	Resting State MRI data from healthy control (HC), Parkinson's disease with normal cognition (PD-NC), and Parkinson's disease with mild cognitive impairment (PD-MCI) cohorts
278	ds006105	Skewed Gambling Task: Deconstructing neural predictors of risky choice
279	ds006156	Executive Functioning-Related Neural Processes in Emerging Adults with Type 1 Diabetes: A Functional Magnetic Resonance Imaging Study
280	ds006193	PAFIN: PennLINC AFfective INstability (Test Dataset)
281	ds006211	Brain bases for navigating acoustic features - fMRI dataset
282	ds006559	Attitude Change Study 1 Persuade
283	ds006568	Attitude Change Study 2 Debate
284	ds006572	RSfMRI_Vision
285	ds006592	The First Comprehensive Study of Early-Emerging Prosopometamorphopsia
286	ds006594	RSfMRI_Function
287	ds006729	R01_Neuromuscular_Signature_Pilot
288	ds006893	The interplay of central insulin and menstrual cycle on functional brain networks and neural food cue reactivity in women
289	ds006944	One-person corticospinal fMRI dataset for Stanford Spinal Cord workshop 2025
290	ds006945	Dataset: T1-Weighted Structural MRI and fMRI of Participants Viewing Self-Avatar Exoskeleton Walking (11 SWS Cycles)
291	ds006952	PSYCH-REST
292	ds006971	Shoulder abduction motor-task fMRI in healthy participants
293	ds007048	Whole-Brain Auditory Mapping with 3T Multi-Echo fMRI in healthy participants
294	ds007115	Neural Mechanisms of Cognitive Conflict: Processing COVID-19 Vaccine Misinformation
295	ds007126	An open fMRI dataset examining the effects of online social and non-social information distraction on global-local attention
296	ds007135	Investigating Neuroepigenetic Alterations in Chronic Low Back Pain with PET/MR
297	ds007185	Neural Responses to Curiosity in Young Children
298	ds007313	Brain and spinal cord fMRI and qMRI - Single participant
299	ds007345	Multimodal Brain MRI and Clinical Data in Olfactory Groove Meningioma: A Prospective Data Report
300	ds007441	Human Dignity
301	ds007534	Neuroil