PROBABILISTIC MIXED-EFFECTS GRAPH AUTOENCODING FOR INDIVIDUALIZED MULTI-PARADIGM FUNCTIONAL CONNECTOME MODELING

Anonymous authors
Paper under double-blind review

ABSTRACT

Repeated observations of the same individual are common in biomedical data, and multi-paradigm fMRI is a prominent example: every scan reflects the current task while sharing a stable individual signature with the participant's other scans. Most representation-learning methods treat scans as independent samples or leave this shared component implicit. We formulate functional connectome reconstruction under repeated measurements as a conditional neural mixed-effects problem and propose P-MEGA, a probabilistic mixed-effects graph autoencoder that decomposes each functional connectome into a scan-specific population-level prediction and a low-rank participant-specific random effect shared across scans, which we refer to as the participant's brain fingerprint. A signed multi-window graph encoder and a shared edge decoder form the population-level pathway, while a permutation-invariant set encoder parameterizes an amortized Gaussian posterior over the participant-specific effect from an arbitrary set of support scans, enabling inference for unseen participants with all model parameters shared across the cohort. We evaluate P-MEGA on two large multi-paradigm cohorts widely used to study individual differences and brain development: the Human Connectome Project (HCP; 865 young adults, 9 paradigms) and the Philadelphia Neurodevelopmental Cohort (PNC; 1,272 youths aged 8–21, 3 paradigms). P-MEGA attains the best out-of-fold reconstruction among the evaluated baselines. A support-swap test, which fixes the target scan and changes only whose scans are used to infer the fingerprint, shows that fingerprints inferred from the same participant's other paradigms explain the target's residual connectivity better than those from demographically matched or random participants in both cohorts. The inferred fingerprints are reproducible across training seeds and folds and show network-level associations with cognition and age.

1 INTRODUCTION

Many fMRI studies acquire several rest and task scans from each participant, and each scan yields a functional connectome, the matrix of functional connectivity (FC) between all pairs of brain regions. An observed functional connectome reflects a common network architecture, the current task and acquisition, and stable features of the individual. FC changes between rest and cognitive tasks (Smith et al., 2009; Cole et al., 2014), yet each individual's connectivity profile persists across sessions and conditions and acts as a fingerprint that identifies the person within a large group (Finn et al., 2015; Gordon et al., 2017). These connectivity profiles also predict cognition (Finn et al., 2015), and task scans can sharpen such brain–behavior associations (Greene et al., 2018). Pooling rest with task data further improves the reliability of individual FC estimates (Elliott et al., 2019). Multi-paradigm datasets therefore have a repeated-measures structure: scans differ in condition but remain linked by participant.

Separating these sources of variation is central to studying individual differences: paradigm-dependent connectivity describes how the brain reconfigures for a task, whereas the participant-level component is the trait-like signal that brain–behavior studies seek to relate to cognition. Most fMRI representation models leave this separation implicit. Training on scans independently discards the participant link, and a flexible encoder that preserves identity may spread it throughout the represen-

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.

tation, so it is hard to tell whether a learned component reflects a transferable individual deviation, a scan-specific pattern, or extra model capacity.

General FC pools resting and task acquisitions into a more reliable estimate of shared individual structure (Elliott et al., 2019), connectome-to-connectome models learn transformations between designated states (Yoo et al., 2022), and graph neural networks, including dynamic models such as STAGIN, learn scan-level representations from static or window based dynamic FC (dFC) graphs (Li et al., 2021; Kim et al., 2021); none of these isolates a participant component shared across paradigms. Mixed-effects models supply the complementary statistical principle: a population mapping plus unit-specific random deviations accounts for dependence among repeated measurements (Laird & Ware, 1982; Lindstrom & Bates, 1990), and recent work carries this structure into deep networks (Simchoni & Rosset, 2021; 2023) and latent-variable models (Simchoni & Rosset, 2024). A whole-connectome formulation that combines nonlinear scan encoding with a participant effect shared across paradigms and amortized from data remains open.

We address this gap with a conditional neural mixed-effects formulation, instantiated as a *probabilistic mixed-effects graph autoencoder* (P-MEGA). Each connectome is decomposed into a scan-conditioned population-shared component and an additive low-rank participant effect shared across repeated scans. A permutation-invariant set encoder amortizes the posterior over the participant effect from any subset of available scans, enabling inference for previously unseen participants. This construction admits a direct mixed-effects interpretation: marginalizing the participant latent recovers the between-scan covariance structure of a linear mixed model, while the training objective is a variational lower bound on its marginal likelihood with a neural fixed component. Participant specificity is induced by parameter sharing across repeated scans.

To directly test whether the inferred participant effect carries transferable person-specific information, we introduce a *target-excluded support-swap test*. The target scan and shared prediction are fixed, while only the identity of the participant providing the support scans is changed. If the participant effect captures transferable individual information, support from the same participant should explain the held-out target residual better than support from matched or random participants.

Our contributions are threefold: **(i)** a neural mixed-effects graph autoencoder that, to our knowledge, is the first to equip a graph autoencoder with an explicit participant random effect, with amortized inference for unseen participants from any subset of their scans; **(ii)** an identity-controlled support-swap test that changes only whose scans inform the participant effect, separating person-specific information from generic corrections and added capacity; and **(iii)** a decomposition of individualized gains on two multi-paradigm cohorts comprising 2,137 participants and 11,601 scans, showing that the participant pathway combines a correction shared across people with a person-specific component that survives demographic matching and is reproducible across seeds and folds.

2 Related Work

Graph learning for functional connectivity. Graph-learning studies commonly treat regions as nodes and FC values as edges. BrainGNN uses ROI-aware convolution and pooling (Li et al., 2021); ContrastPool learns a contrastive hierarchy (Xu et al., 2024). Dynamic methods build a sequence of windowed graphs. STAGIN combines graph-isomorphism layers with a Transformer (Kim et al., 2021), and dFCExpert uses modularity experts and soft state prototypes (Chen et al., 2025). Window length, sampling, and noise all shape these fluctuations (Hutchison et al., 2013; Lurie et al., 2020). P-MEGA uses windowed graphs as within-scan evidence and connects repeated scans through a participant latent variable.

Individualized and multi-paradigm functional connectivity. FC fingerprints persist across sessions and tasks (Finn et al., 2015), and intensive sampling reveals stable individual factors within a dominant common architecture (Gordon et al., 2017; Gratton et al., 2018). General FC pools rest and task data to estimate reliable individual differences (Elliott et al., 2019); related methods extract state-unspecific modes (Takagi et al., 2019) or transform connectivity between cognitive states (Yoo et al., 2022). P-MEGA connects these objectives: each scan retains its own reconstruction target, the scan-conditioned pathway captures paradigm-related variation, and only a low-rank participant effect is shared across that participant's observations. This decomposition enables participant-specific information to be tested directly through support substitution across paradigms.

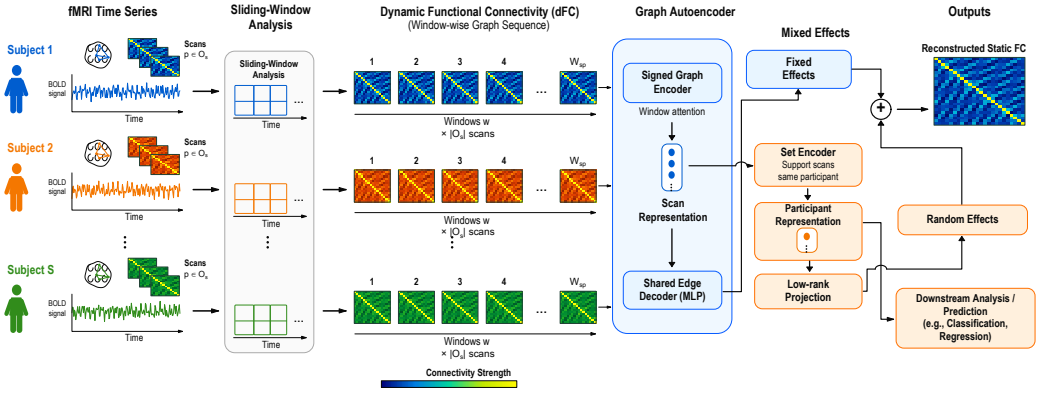

Figure 1: Overview of P-MEGA. Regional fMRI time series are converted into sliding-window functional connectivity graphs and encoded into scan-level representations. Connectome reconstruction is decomposed into a population-shared component and a participant-specific random effect inferred from repeated scans. Their sum reconstructs the static functional connectome and provides participant-level representations for downstream analyses.

Mixed-effects and set-conditioned latent-variable models. Mixed-effects models combine population terms with unit-level random deviations to handle repeated observations (Laird & Ware, 1982; Lindstrom & Bates, 1990). Neural extensions include LMMNN (Simchoni & Rosset, 2023) and LMM-VAE (Simchoni & Rosset, 2024). Set-conditioned latent models, including the Neural Statistician, multi-level VAE, and neural processes, infer a group latent through permutation-invariant aggregation (Edwards & Storkey, 2016; Bouchacourt et al., 2018; Garnelo et al., 2018a;b). P-MEGA shares amortized set-conditioned inference with the latter, but places an additive low-rank participant effect directly in connectome space, giving it an explicit mixed-effects covariance interpretation, while retaining a scan-conditioned population pathway. Its participant-level interpretation is evaluated through identity-controlled support swaps rather than inferred from reconstruction performance alone. A particularly relevant statistical precursor is the nonlinear mixed model for multi-paradigm fMRI of Li et al. (2026), which models population covariates together with subject-specific random intercepts on selected edges. P-MEGA instead performs graph-based whole-connectome reconstruction and replaces fitted subject-specific coefficients with an amortized posterior that can be inferred for previously unseen participants.

3 Method

P-MEGA represents an observed connectome as the sum of a scan-conditioned population prediction and a participant-level random deviation shared across repeated scans (Figure 1). We first state the decomposition statistically, then describe its two neural components and the learning objective.

3.1 Problem Setup

Let $s \in \{1, \dots, S\}$ index participants and $p \in \{1, \dots, P\}$ index paradigms (resting-state and task conditions). Each participant contributes at most one scan per paradigm, so the subscript pair sp identifies a single scan: the scan of participant s acquired under paradigm p. Let $\mathcal{O}_s \subseteq \{1, \dots, P\}$ denote the paradigms observed for participant s. Scan sp provides regional time series $\mathbf{Y}_{sp} \in \mathbb{R}^{N \times T_{sp}}$, where N is the number of brain regions and T_{sp} is the number of time points, which varies across participants and paradigms. From each scan we compute its static functional connectome, i.e., the Fisher-transformed Pearson correlations between all pairs of regions, vectorized as $\mathbf{x}_{sp} \in \mathbb{R}^M$ over the $M = N(N-1)/2$ unique edges.

Input and output. For a target scan sp, the model takes two inputs: (i) the target time series $\mathbf{Y}_{sp}$, and (ii) a support set of scans from the same participant, $\{\mathbf{Y}_{sq}\}_{q \in \mathcal{S}_s}$ with $\mathcal{S}_s \subseteq \mathcal{O}_s$. It produces two outputs: (i) a reconstruction of the target connectome as the sum of a population prediction, computed from the target scan, and a participant effect (the participant's brain fingerprint), inferred from the support set alone; and (ii) a posterior distribution over this fingerprint, whose mean serves

as the participant-level representation in downstream analyses. All parameters are shared across participants, and data are split by participant, so every test participant is unseen during training and their fingerprint is inferred without refitting.

Support settings. The *complete* setting uses $\mathcal{S}_s = \mathcal{O}_s$, matching the training configuration. The *held-out-support* setting removes the target paradigm, $\mathcal{S}_s = \mathcal{O}_s \setminus \{p\}$, so the fingerprint must be inferred from other paradigms. An empty support set yields the prior-mean (zero) participant effect and serves as the *population-only* condition.

3.2 Conditional neural mixed-effects model

Repeated scans of a participant are not independent: they share a stable individual component across sessions and tasks (Finn et al., 2015; Gratton et al., 2018). Mixed-effects models handle such repeated measurements by pairing a population (fixed) part with a participant-level random deviation pooled over that participant's scans (Laird & Ware, 1982), which we use to separate paradigm-dependent connectivity from each participant's brain fingerprint. In a linear mixed-effects model (LMM), the fixed part is a linear function of observed covariates and the random deviation is a Gaussian coefficient shared within each cluster of observations; integrating out the random deviation couples observations from the same cluster and shrinks their deviations toward a common population distribution (Laird & Ware, 1982).

P-MEGA extends this construction to a multivariate nonlinear setting:

$$\mathbf{x}_{sp} \Leftarrow \underbrace{\mathcal{D}_\theta(\mathbf{u}_{sp})}_{\mathbf{m}_{sp}^{\text{pop}}} + \underbrace{\mathbf{W}_b^\top \mathbf{a}_s}_{\mathbf{b}_s} + \boldsymbol{\varepsilon}_{sp}, \qquad \mathbf{a}_s \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{d_b}), \quad \boldsymbol{\varepsilon}_{sp} \sim \mathcal{N}(\mathbf{0}, \sigma_x^2 \mathbf{I}_M). \tag{1}$$

Here $\boldsymbol{\varepsilon}_{sp}$ is the scan-level residual with a shared observation variance σ_x^2, and θ collects the decoder parameters. The deterministic code $\mathbf{u}_{sp} \in \mathbb{R}^{d_u}$ is computed from the observed scan, and $\mathcal{D}_\theta$ maps it to the population-shared prediction $\mathbf{m}_{sp}^{\text{pop}}$. The participant latent variable $\mathbf{a}_s \in \mathbb{R}^{d_b}$ is projected through a shared basis $\mathbf{W}_b \in \mathbb{R}^{d_b \times M}$ to give the connectome-space random effect $\mathbf{b}_s$, which is the same for every scan of participant s, whereas $\mathbf{m}_{sp}^{\text{pop}}$ changes with the scan and paradigm. In matrix form, let $R = \sum_s |\mathcal{O}_s|$ be the number of scans, $\mathbf{X}_{\text{FC}} \in \mathbb{R}^{R \times M}$ stack the FC vectors, $\mathbf{U} \in \mathbb{R}^{R \times d_u}$ stack the scan codes, $\mathbf{Z}_{\text{subj}} \in \{0,1\}^{R \times S}$ be the participant-incidence matrix, and $\mathbf{A} \in \mathbb{R}^{S \times d_b}$ stack one row $\mathbf{a}_s^\top$ per participant:

$$\mathbf{X}_{\text{FC}} \Leftarrow \mathcal{D}_\theta(\mathbf{U}) + \mathbf{Z}_{\text{subj}} \mathbf{A} \mathbf{W}_b + \mathbf{E} \tag{2}$$

where $\mathbf{E} \in \mathbb{R}^{R \times M}$ stacks the residuals $\boldsymbol{\varepsilon}_{sp}^\top$. The linear fixed-effect term of the LMM becomes the nonlinear population mapping $\mathcal{D}_\theta(\mathbf{U})$, and its random-effect term becomes the low-rank participant term $\mathbf{Z}_{\text{subj}} \mathbf{A} \mathbf{W}_b$. Here "population-shared" describes the common mapping; its output still depends on the individual scan. Writing $\mathbf{b}_s = \mathbf{W}_b^\top \mathbf{a}_s$ gives $\mathbf{b}_s \sim \mathcal{N}(\mathbf{0}, \boldsymbol{\Sigma}_b)$ with $\boldsymbol{\Sigma}_b = \mathbf{W}_b^\top \mathbf{W}_b$ and $\operatorname{rank}(\boldsymbol{\Sigma}_b) \le d_b$. Two scans from one participant share covariance $\boldsymbol{\Sigma}_b$, while scans from different participants carry separate deviations.

3.3 Population pathway: signed multi-window graph encoder and edge decoder

Targets and dynamic graphs. Each ROI time series is standardized within scan. With $\mathbf{C}_{sp}^{(0)} \in [-1,1]^{N \times N}$ the full-scan Pearson correlation matrix, the target $\mathbf{x}_{sp}$ is the strict upper triangle of $\operatorname{atanh}(\mathbf{C}_{sp}^{(0)})$, with correlations clipped slightly inside $(-1,1)$ for numerical stability. A window of length L moved with stride Δ yields $J_{sp} = \lfloor (T_{sp} - L)/\Delta \rfloor + 1$ window correlation matrices $\mathbf{C}_{sp}^{(w)}$, $w = 1, \ldots, J_{sp}$. Static and window-specific connectivity are fused into the weighted graph used for message passing,

$$\widetilde{\mathbf{C}}_{sp}^{(w)} = \alpha_p \mathbf{C}_{sp}^{(0)} + (1 - \alpha_p) \mathbf{C}_{sp}^{(w)}, \tag{3}$$

where $\alpha_p \in (0,1)$ is a learnable weight for paradigm p that balances full-scan and window-specific connectivity, initialized at 0.5; it is the model's only paradigm-indexed parameter. The initial feature of region i in window w is $\mathbf{h}_{sp,i}^{(w,0)} = [\phi_y(\widetilde{\mathbf{y}}_{sp,i}^{(w)}) \,\|\, \phi_c(\mathbf{c}_{sp,i}^{(0)})]$, which concatenates projections of its standardized window segment $\widetilde{\mathbf{y}}_{sp,i}^{(w)} \in \mathbb{R}^L$ and its full-scan connectivity profile $\mathbf{c}_{sp,i}^{(0)}$ (row i of $\mathbf{C}_{sp}^{(0)}$ with a zeroed diagonal); ϕ_y, ϕ_c are learned projections followed by normalization and a nonlinearity.

Signed propagation and window aggregation. Positive and negative correlations are propagated separately. Suppressing scan and window subscripts and omitting biases, one layer computes

$$\widetilde{\mathbf{C}}_+ = \mathrm{ReLU}(\widetilde{\mathbf{C}}), \quad \widetilde{\mathbf{C}}_- = \mathrm{ReLU}(-\widetilde{\mathbf{C}}), \quad \mathbf{H}_\pm^{(\ell)} = \mathrm{diag}(\widetilde{\mathbf{C}}_\pm \mathbf{1})^{-1}\, \widetilde{\mathbf{C}}_\pm \mathbf{H}^{(\ell-1)} \mathbf{W}_\pm^{(\ell)}. \tag{4}$$

Here $\ell = 1, 2$ indexes the layer, $\mathbf{H}^{(0)}$ stacks the initial features $\mathbf{h}_{sp,i}^{(w,0)}$ over regions, $\mathbf{1}$ is the all-ones vector so that $\mathrm{diag}(\widetilde{\mathbf{C}}_\pm \mathbf{1})$ is the degree matrix, and $\mathbf{W}_\pm^{(\ell)}$ are learnable weights. The two messages are concatenated, linearly fused, normalized, and passed through a nonlinearity to form $\mathbf{H}^{(\ell)}$. Mean readout over regions yields a window token $\mathbf{z}_{sp}^{(w)} \in \mathbb{R}^{d_h}$, and attention pools these tokens,

$$e_{sp}^{(w)} = \mathbf{v}^\top \tanh(\mathbf{W}_z \mathbf{z}_{sp}^{(w)}), \qquad \omega_{sp}^{(w)} = \frac{\exp(e_{sp}^{(w)})}{\sum_{r=1}^{J_{sp}} \exp(e_{sp}^{(r)})}, \qquad \mathbf{g}_{sp} = \sum_{w=1}^{J_{sp}} \omega_{sp}^{(w)} \mathbf{z}_{sp}^{(w)}, \tag{5}$$

giving a scan-level summary $\mathbf{g}_{sp}$, where $\mathbf{W}_z \in \mathbb{R}^{d_h \times d_h}$ and $\mathbf{v} \in \mathbb{R}^{d_h}$ are learnable attention parameters and d_h is the encoder width. The attention weights windows adaptively but does not model temporal state transitions.

Edge decoder. A shared MLP compresses the scan summary to the code $\mathbf{u}_{sp} = \mathrm{MLP}_\phi(\mathbf{g}_{sp}) \in \mathbb{R}^{d_u}$. The decoder learns an embedding $\boldsymbol{\rho}_i \in \mathbb{R}^{d_r}$ for every ROI and, for each pair $i < j$, forms the descriptor $\boldsymbol{\kappa}_{ij} = [\boldsymbol{\rho}_i \| \boldsymbol{\rho}_j \| \boldsymbol{\rho}_i \odot \boldsymbol{\rho}_j \| |\boldsymbol{\rho}_i - \boldsymbol{\rho}_j|]$, which is symmetric in i and j up to ordering. A shared edge function then gives

$$m_{sp,ij}^{\mathrm{pop}} = f_\theta([\mathbf{u}_{sp} \| \boldsymbol{\kappa}_{ij}]), \qquad \mathcal{D}_\theta(\mathbf{u}_{sp}) = \left[m_{sp,ij}^{\mathrm{pop}}\right]_{1 \le i < j \le N} \in \mathbb{R}^M. \tag{6}$$

Because the ROI embeddings and f_θ are shared across all scans and participants, the population term varies with the observed scan while its mapping is common to the cohort. Outputs stay on the unbounded Fisher-z scale.

3.4 Amortized probabilistic participant effects

Let $\mathcal{S}_s \subseteq \mathcal{O}_s$ be the support scans used to infer the random effect for participant s, with $K_s = |\mathcal{S}_s|$. A shared network R_ψ maps the scan summary $\mathbf{g}_{sp}$ of each support scan to an evidence vector; these are averaged in DeepSets style (Zaheer et al., 2017) into $\bar{\mathbf{e}}_s$, which is concatenated with $\log K_s$ to form the participant feature $\mathbf{v}_s$. Two heads h_μ and h_σ then map $\mathbf{v}_s$ to a diagonal Gaussian posterior:

$$\begin{aligned} \bar{\mathbf{e}}_s &= \frac{1}{K_s} \sum_{p \in \mathcal{S}_s} R_\psi(\mathbf{g}_{sp}), \qquad \mathbf{v}_s = [\bar{\mathbf{e}}_s \| \log K_s], \\ q_\psi(\mathbf{a}_s \mid \mathcal{S}_s) &= \mathcal{N}\big(\boldsymbol{\mu}_s, \mathrm{diag}(\boldsymbol{\sigma}_s^2)\big), \qquad \boldsymbol{\mu}_s = h_\mu(\mathbf{v}_s), \qquad \boldsymbol{\sigma}_s = \mathrm{softplus}(h_\sigma(\mathbf{v}_s)) + \sigma_{\min}. \end{aligned} \tag{7}$$

Mean pooling controls the aggregate scale, $\log K_s$ carries the amount of evidence, and the floor $\sigma_{\min} > 0$ guards against numerical degeneracy. During training, $\mathcal{S}_s$ contains all retained scans of participant s in the batch. One reparameterized sample $\mathbf{a}_s = \boldsymbol{\mu}_s + \boldsymbol{\sigma}_s \odot \boldsymbol{\epsilon}_s$, $\boldsymbol{\epsilon}_s \sim \mathcal{N}(\mathbf{0}, \mathbf{I}_{d_b})$, is shared across those scans, giving $\widehat{\mathbf{x}}_{sp} = \mathcal{D}_\theta(\mathbf{u}_{sp}) + \mathbf{W}_b^\top \mathbf{a}_s$. Validation and test reconstruction use the posterior mean. We write $\boldsymbol{\mu}_\psi(\mathcal{S}_s)$ and $\boldsymbol{\sigma}_\psi(\mathcal{S}_s)$ for $\boldsymbol{\mu}_s$ and $\boldsymbol{\sigma}_s$ when the dependence on the support set matters. Shared inference functions apply the same rule to training and unseen participants. Orthogonal rotations of $\mathbf{a}_s$ and $\mathbf{W}_b$ preserve their product, and the two pathways have overlapping capacity. Our analyses center on the projected effect $\mathbf{b}_s$, its network summaries, and its held-out-support behavior.

3.5 Learning objective and inference

Training batches are organized by participant. For a batch $\mathcal{B}$ of K participants with all their retained scans, the Gaussian reconstruction negative log-likelihood under (1) and the per-participant KL divergence to the standard-normal prior are

$$\mathcal{L}_{\mathrm{NLL}} = \frac{1}{2} \sum_{s \in \mathcal{B}} \sum_{p \in \mathcal{O}_s} \sum_{m=1}^{M} \left[\frac{(x_{spm} - \widehat{x}_{spm})^2}{\sigma_x^2} + \log(2\pi\sigma_x^2) \right], \quad \mathcal{L}_{\mathrm{KL},s} = \frac{1}{2} \sum_{k=1}^{d_b} \left(\mu_{sk}^2 + \sigma_{sk}^2 - 1 - \log \sigma_{sk}^2 \right), \tag{8}$$

where x_{spm} is the m-th edge of $\mathbf{x}_{sp}$, μ_{sk} and σ_{sk} are the k-th entries of $\boldsymbol{\mu}_s$ and $\boldsymbol{\sigma}_s$, and σ_x is learned jointly with the network. The complete objective is

$$\mathcal{L}(\mathcal{B}) = \frac{\mathcal{L}_{\mathrm{NLL}} + \beta \sum_{s \in \mathcal{B}} \mathcal{L}_{\mathrm{KL},s}}{KM} + \lambda_U \|\mathbf{U}_{\mathcal{B}}\|_{\mathrm{F}}^2 + \lambda_W \|\mathbf{W}_b\|_{\mathrm{F}}^2, \tag{9}$$

where $\mathbf{U}_{\mathcal{B}}$ stacks the scan codes in the batch and $\|\cdot\|_{\mathrm{F}}$ is the Frobenius norm. The first term is a scaled negative evidence lower bound for the participant latent variable (Kingma & Welling, 2013), in which β weights the KL term; λ_U and λ_W are regularization weights on the scan codes and the random-effect basis. We set $\beta = 1$, which recovers the standard evidence lower bound. The KL term acts as the participant-level shrinkage penalty and is counted once per participant.

Inference for unseen participants. After training, all global parameters stay fixed. For a target scan (s, p) of a new participant and any support set $\mathcal{S}_s \subseteq \mathcal{O}_s$, the reconstruction is $\widehat{\mathbf{x}}_{sp}^{(\mathcal{S}_s)} = \mathcal{D}_\theta(\mathbf{u}_{sp}) + \mathbf{W}_b^\top \boldsymbol{\mu}_\psi(\mathcal{S}_s)$ for $K_s > 0$ and $\mathcal{D}_\theta(\mathbf{u}_{sp})$ for $K_s = 0$. With $\mathcal{S}_s = \mathcal{O}_s$, the posterior mean uses the participant's complete scan set including the target; this full-set reconstruction is the primary setting and matches the training configuration. With an empty support set, the effect equals its prior mean, giving the population-only prediction. For held-out-support inference, all scans of the target paradigm are excluded from $\mathcal{S}_s$ and the effect is inferred from the remaining paradigms, while the target scan still enters the population pathway through $\mathcal{D}_\theta(\mathbf{u}_{sp})$. This setting tests whether a deviation inferred from other paradigms transfers to the target paradigm; target-blind prediction of an unobserved connectome is a separate problem. The posterior scale $\boldsymbol{\sigma}_\psi(\mathcal{S}_s)$ summarizes uncertainty in the latent effect as the amount of support evidence changes.

4 Experiments

All results below are out of fold (OOF), with the participant as the unit of aggregation and inference.

4.1 Setup

Cohorts and protocol. We use two multi-paradigm cohorts (Appendix Table 5): the Human Connectome Project (HCP; 865 participants, 9 paradigms, 7,785 scans), with repeated resting-state and task fMRI from young adults (Van Essen et al., 2013) in the 268-region Shen parcellation (Shen et al., 2013), and the Philadelphia Neurodevelopmental Cohort (PNC; 1,272 participants aged 8–21, 3 paradigms, 3,816 scans), with resting-state, emotion-identification, and working-memory scans (Satterthwaite et al., 2014) in the 264-region Power parcellation (Power et al., 2011). The cohorts are modeled independently, so raw errors are compared within cohort. Each participant and all of that participant's scans fall in a single outer fold of a sex-stratified five-fold cross-validation, with 12.5% of each development set reserved for validation. Splits are fixed with one seed, every model is trained with three seeds, hyperparameters are fitted on the training partition, validation data select the checkpoint with the lowest participant-averaged MSE, and each outer test fold is evaluated once. The encoder uses two signed graph-convolution layers of width 128 and 64-dimensional scan and participant codes; the decoder uses 32-dimensional ROI embeddings and a three-layer MLP of width 256; training uses Adam for at most 200 epochs with early stopping (Appendices A and B).

Baselines. The *paradigm template* assigns each test scan the mean FC vector of its paradigm, estimated from training participants only. *Graph-AE* and *Graph-VAE* are deterministic and variational graph autoencoders with scan-level codes (Kipf & Welling, 2016; Simonovsky & Komodakis, 2018). We reimplemented both models following their original formulations and retrained them from scratch on each cohort under the same training budget as P-MEGA. *LMM-VAE* (Simchoni & Rosset, 2024), with a 64-dimensional latent, is the closest probabilistic mixed-effects representation-learning baseline. All learned models share the partitions and targets. We also report *population only*, which sets $\mathbf{b}_s = \mathbf{0}$ in the trained P-MEGA model at inference, and *complete*, which uses the posterior mean from all scans of the participant (Appendix C). The comparison covers methods implemented under a common protocol and serves as a controlled reference set.

4.2 Connectome reconstruction

The complete P-MEGA model attains the lowest participant-averaged MSE and the highest edge-wise correlation in both cohorts (Table 1). In HCP, LMM-VAE is the strongest learned baseline,

Table 1: Five-fold OOF connectome reconstruction (mean ± SD over three training seeds). MSE is averaged over scans within participant; Pearson correlation is computed across FC edges per scan and averaged within participant. The paradigm template is deterministic and is reported without a seed SD.

	HCP		PNC	
Method	MSE ↓	Pearson ↑	MSE ↓	Pearson ↑
Paradigm template	0.03075	0.695	0.06531	0.502
Graph-AE	0.02501 ± 0.00135	0.698 ± 0.002	0.03858 ± 0.00012	0.596 ± 0.001
Graph-VAE	0.03268 ± 0.00115	0.668 ± 0.002	0.04178 ± 0.00024	0.561 ± 0.001
LMM-VAE ($d = 64$)	0.02116 ± 0.00045	0.705 ± 0.006	0.04139 ± 0.00128	0.560 ± 0.007
P-MEGA, population only	0.02972 ± 0.00009	0.563 ± 0.007	0.04847 ± 0.00033	0.407 ± 0.005
P-MEGA, complete	**0.01985 ± 0.00008**	**0.728 ± 0.001**	**0.03836 ± 0.00005**	**0.599 ± 0.001**

Table 2: Component ablations under a common 200-epoch training cap (five-fold OOF participant-averaged MSE; lower is better). The full model is retrained under the same cap.

Variant	HCP MSE ↓	PNC MSE ↓
P-MEGA (full)	**0.01968**	**0.03819**
No participant random effect	0.02527	0.03974
Deterministic participant effect	0.02041	0.03910
Direct high-dimensional effect ($\mathbf{b}_s \in \mathbb{R}^M$)	0.02029	0.03868
Static adjacency	0.02010	0.03927
Dynamic-focused encoder	0.03525	0.04571

and P-MEGA lowers MSE by a further 6.2% while raising Pearson correlation by 0.023. In PNC, Graph-AE is the strongest learned baseline; P-MEGA improves MSE by 0.6% and correlation by 0.003. Its gains over LMM-VAE are 7.3% in MSE and 0.039 in correlation. Compared with the population-only condition, the complete reconstruction lowers MSE by 33.2% in HCP and 20.9% in PNC and raises correlation by 0.165 and 0.192. Every paradigm shows an error reduction (Appendix Table 6), and all 30 main-model runs pass the prespecified posterior-degeneracy diagnostic.

4.3 Component ablations

We retrain the model under a common 200-epoch budget, changing one component at a time (Table 2; definitions in Appendix C). Removing the participant effect raises reconstruction error markedly in HCP and moderately in PNC. A deterministic low-rank code and a direct M-dimensional effect both stay close to the full model. The individualized pathway supplies most of the gain, and the probabilistic low-rank parameterization gives the lowest error in both cohorts. Static adjacency also stays close to the full model when window features and pooling remain active. The dynamic-focused encoder performs substantially worse in both cohorts. These results favor window-resolved features anchored by full-scan FC.

4.4 Participant specificity from held-out support scans

The central test holds the target scan and its population prediction fixed while changing only the scans used to infer $\boldsymbol{\mu}_s$. For each participant and target paradigm, *true support* consists of one or more of the participant's other paradigms; *matched support* uses the identical paradigms from a different participant matched on sex and nearest training-standardized age; *random support* uses one randomly selected different participant. A single donor supplies every scan in a pseudo-support set, so the three conditions are matched in size and paradigm composition; *population only* sets $\mathbf{b} = \mathbf{0}$, and *oracle* uses all scans of the true participant including the target. The design isolates the incremental information supplied by the random-effect pathway, as $\mathcal{D}_\theta(\mathbf{u}_{sp})$ receives the target time series in every condition. We compare the inferred $\widehat{\mathbf{b}}$ with the target residual $\mathbf{x}_{sp} - \mathbf{m}_{sp}^{\text{pop}}$ through residual correlation and MSE reduction relative to the population-only prediction (Appendix E).

True same-participant support predicts the held-out target residual more accurately than matched or random other-participant support (Table 3; Figure 2). The paired true-versus-matched difference

Table 3: Held-out-support validation of participant random effects. True, matched, and random support use identical support counts and paradigms. Values are participant-level estimates with 95% bootstrap CIs; MSE reduction is relative to the population-only prediction. Bold marks the best condition among those excluding the target scan; oracle includes the target as an upper reference.

	HCP		PNC	
Source of $\widehat{\mathbf{b}}$	Residual r ↑	MSE reduction ↑	Residual r ↑	MSE reduction ↑
Random support (other participant)	0.556 [0.549, 0.562]	0.00823 [0.00808, 0.00838]	0.404 [0.398, 0.410]	0.00683 [0.00656, 0.00709]
Matched support (other participant)	0.557 [0.551, 0.564]	0.00826 [0.00811, 0.00842]	0.410 [0.403, 0.416]	0.00712 [0.00686, 0.00739]
True support (same, target excluded)	**0.574 [0.565, 0.581]**	**0.00877 [0.00857, 0.00897]**	**0.445 [0.438, 0.453]**	**0.00887 [0.00860, 0.00915]**
Oracle (same, target included)	0.599 [0.592, 0.606]	0.00992 [0.00971, 0.01014]	0.475 [0.468, 0.482]	0.01041 [0.01013, 0.01070]

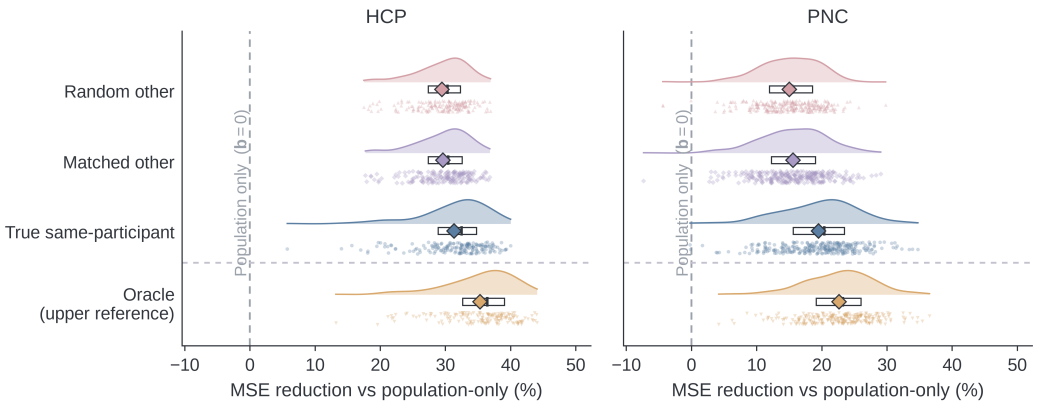

Figure 2: Held-out-support validation. MSE reduction relative to the population-only prediction for random-other, matched-other, true same-participant, and oracle support in HCP and PNC; diamonds mark mean estimates and boxplots summarize participant-level distributions. Paired participant-level differences are shown in Appendix Figure 3.

in residual correlation is 0.0164 (95% CI [0.0137, 0.0188]) in HCP and 0.0354 (95% CI [0.0320, 0.0389]) in PNC ($q < 0.001$ for both); the true-versus-random contrasts are 0.0179 (95% CI [0.0152, 0.0203]) and 0.0410 (95% CI [0.0375, 0.0446]). Demographic matching leaves this advantage intact, so it reflects participant information beyond age and sex. Other-participant support also lowers error relative to the population-only prediction, indicating that part of the pathway's contribution is a correction that transfers across participants; the additional gain from true support is the participant-specific component and constitutes the more specific evidence for individualization. Performance improves with the number of true support paradigms, and the mean posterior variance contracts steadily as support grows; the oracle condition provides an upper reference.

4.5 Phenotypic Utility

As a secondary test, we attach prediction heads to either the population representation $\mathbf{u}$ alone or the combined $[\mathbf{u}\|\boldsymbol{\mu}_s]$ and fine-tune the relevant encoder pathways end to end within each outer training fold (Appendix F). The contribution of the participant representation depends on the outcome (Table 4). For cognition, adding $\boldsymbol{\mu}_s$ yields a small improvement in HCP (fluid-reasoning RMSE 0.995 to 0.993) and a clearer one for reading achievement in PNC (RMSE 0.972 to 0.960; R^2 from 0.051 to 0.075). For sex prediction, $\mathbf{u}$ alone reaches the higher AUROC in both cohorts (0.846 versus 0.832 in HCP; 0.724 versus 0.699 in PNC), and P-MEGA's $\mathbf{u}$ exceeds every baseline representation. The two representations therefore carry complementary phenotypic information, with the participant representation adding the most for reading achievement in PNC.

Table 4: Out-of-fold downstream phenotypic prediction (mean ± SD over training seeds): sex AUROC and RMSE/MAE for standardized reading achievement (PNC) or fluid reasoning (HCP). Baselines use their learned representation $\mathbf{u}$; P-MEGA uses $\mathbf{u}$ alone or $[\mathbf{u}\|\boldsymbol{\mu}_s]$.

Cohort	Method (representation)	Sex AUROC ↑	Cognitive RMSE ↓	Cognitive MAE ↓
PNC	Graph-AE ($\mathbf{u}$)	0.679 ± 0.017	0.976 ± 0.002	0.760 ± 0.001
	Graph-VAE ($\mathbf{u}$)	0.703 ± 0.012	0.993 ± 0.004	0.773 ± 0.002
	LMM-VAE ($\mathbf{u}$)	0.689 ± 0.010	0.979 ± 0.011	0.761 ± 0.009
	P-MEGA ($\mathbf{u}$)	$\mathbf{0.724 \pm 0.018}$	0.972 ± 0.009	0.760 ± 0.010
	P-MEGA ($[\mathbf{u}\|\boldsymbol{\mu}_s]$)	0.699 ± 0.012	$\mathbf{0.960 \pm 0.006}$	$\mathbf{0.750 \pm 0.008}$
HCP	Graph-AE ($\mathbf{u}$)	0.775 ± 0.051	1.003 ± 0.004	0.840 ± 0.006
	Graph-VAE ($\mathbf{u}$)	0.535 ± 0.068	1.002 ± 0.000	0.844 ± 0.001
	LMM-VAE ($\mathbf{u}$)	0.788 ± 0.018	0.998 ± 0.010	0.833 ± 0.010
	P-MEGA ($\mathbf{u}$)	$\mathbf{0.846 \pm 0.013}$	0.995 ± 0.009	0.830 ± 0.014
	P-MEGA ($[\mathbf{u}\|\boldsymbol{\mu}_s]$)	0.832 ± 0.011	$\mathbf{0.993 \pm 0.010}$	$\mathbf{0.828 \pm 0.011}$

4.6 Connectome-space characterization of the two pathways

We finally ask whether the two components exhibit coherent organization in connectome space, using seed-averaged OOF representations grouped by cohort-specific network assignments (Appendix G). The population pathway shows widespread, reproducible task-dependent organization: within-participant task–rest contrasts of $\mathbf{m}_{sp}^{\text{pop}}$ survive FDR correction for 245 of 264 HCP and 179 of 210 PNC network-pair tests, with effect directions agreeing across all three seeds for 98.5% and 98.6% of tests. The projected participant effect $\mathbf{b}_s$ forms a reproducible individual-variation fingerprint: network-pair rankings of its across-participant variability are concordant across training seeds (mean pairwise Spearman $\rho = 0.983$ in HCP, 0.984 in PNC) and across outer folds ($\rho = 0.972$ and 0.957). Network summaries of $\mathbf{b}_s$ relate to phenotype after covariate adjustment (31 of 66 HCP network pairs with fluid reasoning, 75 of 105 PNC pairs with reading achievement, and 54 of 66 HCP and 83 of 105 PNC pairs with age), and adjusting for the global RMS of $\mathbf{b}_s$ retains all 31 HCP and 65 of the 75 PNC cognitive associations, so the phenotypic organization reflects more than overall fingerprint magnitude.

5 Discussion

P-MEGA turns repeated scans into an explicit modeling resource. The population pathway follows the observed scan, while the participant pathway pools evidence across paradigms. This division is useful for two reasons. First, the projected effect lives in connectome space, so it can be visualized and summarized by functional systems. Second, its meaning can be tested directly by swapping support identity, which separates the value of person-specific information from that of added model capacity. The support experiment reveals both shared and individual structure. Effects inferred from another participant already improve on the population-only prediction, showing that the participant pathway learns a broadly useful correction. True support adds a further gain in both cohorts, and this gain persists under demographic matching, isolating information that the same person carries across paradigms. The ablations identify the participant pathway as the main source of improvement, and the low-rank probabilistic form yields the best reconstruction in both cohorts while adding an explicit covariance interpretation. The phenotype analyses further indicate that the participant and scan representations carry complementary information.

Scope and future directions. P-MEGA uses a Gaussian participant effect that enters additively in connectome space. Nonlinear or paradigm-dependent random effects, richer priors, and additional cohorts and imaging modalities are natural extensions. The same amortized formulation also provides a starting point for target-blind prediction of unobserved paradigms.

Overall, P-MEGA improves whole-connectome reconstruction on HCP and PNC, and its identity-controlled support swap shows that same-participant scans carry transferable information about a held-out target residual. The resulting effects are reproducible across fits and organize into phenotype-related connectome patterns.

ETHICS STATEMENT

This study uses de-identified, publicly available neuroimaging and phenotypic data from the Human Connectome Project and the Philadelphia Neurodevelopmental Cohort, obtained under their respective data use agreements; no new data were collected. The models characterize individual variation in functional connectivity for research purposes, and the reported brain–behavior associations are correlational and population-level. Any clinical or individual-level use would require separate validation, including motion-aware and family-aware analyses, and careful attention to fairness across demographic groups.

REPRODUCIBILITY STATEMENT

Section 3 specifies the model and objective, Section 4.1 summarizes the evaluation protocol, and Appendices A and B give the cohort inclusion criteria, preprocessing pipelines, window construction, cross-validation design, seeds, architecture sizes, optimization settings, and software versions. Appendix C defines every baseline and ablation variant, Appendix E details the support-swap protocol, Appendix F the downstream probes, and Appendix G the network-level statistical analyses. Source code and configuration files reproducing the experiments will be provided as anonymized supplementary material. Code will be released publicly upon publication.

AI USE STATEMENT

Generative AI tools were used to assist with editing, condensing, and reformatting the manuscript text and LaTeX source for this submission. All AI-assisted text was reviewed and revised by the authors, and every reported number, table, and figure was produced by the authors' own analysis pipeline. Research ideas, experimental design, and the analyses are the authors' own work, and any AI assistance in writing analysis code was verified and tested by the authors. We take responsibility for the final content of this work, including text, claims, and artifacts produced with the aid of generative AI.

REFERENCES

Diane Bouchacourt, Ryota Tomioka, and Sebastian Nowozin. Multi-level variational autoencoder: Learning disentangled representations from grouped observations. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 32, 2018.

Tingting Chen, Hongming Li, Hao Zheng, and Yong Fan. dfcexpert: Learning dynamic functional connectivity patterns with modularity and state experts. *IEEE transactions on medical imaging*, 2025.

Michael W Cole, Jeremy R Reynolds, Jonathan D Power, Grega Repovs, Alan Anticevic, and Todd S Braver. Multi-task connectivity reveals flexible hubs for adaptive task control. *Nature neuroscience*, 16(9):1348–1355, 2013.

Michael W Cole, Danielle S Bassett, Jonathan D Power, Todd S Braver, and Steven E Petersen. Intrinsic and task-evoked network architectures of the human brain. *Neuron*, 83(1):238–251, 2014.

Michael W Cole, Takuya Ito, Douglas Schultz, Ravi Mill, Richard Chen, and Carrisa Cocuzza. Task activations produce spurious but systematic inflation of task functional connectivity estimates. *Neuroimage*, 189:1–18, 2019.

Harrison Edwards and Amos Storkey. Towards a neural statistician. *arXiv preprint arXiv:1606.02185*, 2016.

Maxwell L Elliott, Annchen R Knodt, Megan Cooke, M Justin Kim, Tracy R Melzer, Ross Keenan, David Ireland, Sandhya Ramrakha, Richie Poulton, Avshalom Caspi, et al. General functional connectivity: Shared features of resting-state and task fmri drive reliable and heritable individual differences in functional brain networks. *NeuroImage*, 189:516–532, 2019.

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.

Guillaume Flandin and Karl J Friston. Statistical parametric mapping (spm). *Scholarpedia*, 3(4), 2008.

Marta Garnelo, Dan Rosenbaum, Christopher Maddison, Tiago Ramalho, David Saxton, Murray Shanahan, Yee Whye Teh, Danilo Rezende, and SM Ali Eslami. Conditional neural processes. In *International conference on machine learning*, pp. 1704–1713. PMLR, 2018a.

Marta Garnelo, Jonathan Schwarz, Dan Rosenbaum, Fabio Viola, Danilo J Rezende, SM Eslami, and Yee Whye Teh. Neural processes. *arXiv preprint arXiv:1807.01622*, 2018b.

Matthew F Glasser, Stamatios N Sotiropoulos, J Anthony Wilson, Timothy S Coalson, Bruce Fischl, Jesper L Andersson, Junqian Xu, Saad Jbabdi, Matthew Webster, Jonathan R Polimeni, et al. The minimal preprocessing pipelines for the human connectome project. *Neuroimage*, 80:105–124, 2013.

Evan M Gordon, Timothy O Laumann, Adrian W Gilmore, Dillan J Newbold, Deanna J Greene, Jeffrey J Berg, Mario Ortega, Catherine Hoyt-Drazen, Caterina Gratton, Haoxin Sun, et al. Precision functional mapping of individual human brains. *Neuron*, 95(4):791–807, 2017.

Caterina Gratton, Timothy O Laumann, Ashley N Nielsen, Deanna J Greene, Evan M Gordon, Adrian W Gilmore, Steven M Nelson, Rebecca S Coalson, Abraham Z Snyder, Bradley L Schlaggar, et al. Functional brain networks are dominated by stable group and individual factors, not cognitive or daily variation. *Neuron*, 98(2):439–452, 2018.

Abigail S Greene, Siyuan Gao, Dustin Scheinost, and R Todd Constable. Task-induced brain state manipulation improves prediction of individual traits. *Nature communications*, 9(1):2807, 2018.

R Matthew Hutchison, Thilo Womelsdorf, Elena A Allen, Peter A Bandettini, Vince D Calhoun, Maurizio Corbetta, Stefania Della Penna, Jeff H Duyn, Gary H Glover, Javier Gonzalez-Castillo, et al. Dynamic functional connectivity: promise, issues, and interpretations. *Neuroimage*, 80: 360–378, 2013.

Byung-Hoon Kim, Jong Chul Ye, and Jae-Jin Kim. Learning dynamic graph representation of brain connectome with spatio-temporal attention. *Advances in Neural Information Processing Systems*, 34:4314–4327, 2021.

Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv:1312.6114*, 2013.

Thomas N Kipf and Max Welling. Variational graph auto-encoders. *arXiv preprint arXiv:1611.07308*, 2016.

Maki S Koyama, Adriana Di Martino, Xi-Nian Zuo, Clare Kelly, Maarten Mennes, Devika R Jutagir, F Xavier Castellanos, and Michael P Milham. Resting-state functional connectivity indexes reading competence in children and adults. *The Journal of Neuroscience*, 31(23):8617–8624, 2011.

Nan M Laird and James H Ware. Random-effects models for longitudinal data. *Biometrics*, pp. 963–974, 1982.

Xiaoxiao Li, Yuan Zhou, Nicha Dvornek, Muhan Zhang, Siyuan Gao, Juntang Zhuang, Dustin Scheinost, Lawrence H Staib, Pamela Ventola, and James S Duncan. Braingnn: Interpretable brain graph neural network for fmri analysis. *Medical image analysis*, 74:102233, 2021.

Xiaoxuan Li, Gemeng Zhang, Gang Qu, Anton Orlichenko, Zhengming Ding, Tony W Wilson, Julia M Stephen, Vince D Calhoun, and Yu-Ping Wang. Modeling complex effects and individual variability in multi-paradigm fmri with nonlinear mixed models. *bioRxiv*, pp. 2026–05, 2026.

Mary J Lindstrom and Douglas M Bates. Nonlinear mixed effects models for repeated measures data. *Biometrics*, pp. 673–687, 1990.

Daniel J Lurie, Daniel Kessler, Danielle S Bassett, Richard F Betzel, Michael Breakspear, Shella Kheilholz, Aaron Kucyi, Raphaël Liégeois, Martin A Lindquist, Anthony Randal McIntosh, et al. Questions and controversies in the study of time-varying functional connectivity in resting fmri. *Network neuroscience*, 4(1):30–69, 2020.

Sophia Mueller, Danhong Wang, Michael D Fox, BT Thomas Yeo, Jorge Sepulcre, Mert R Sabuncu, Rebecca Shafee, Jie Lu, and Hesheng Liu. Individual variability in functional connectivity architecture of the human brain. *Neuron*, 77(3):586–595, 2013.

Jonathan D Power, Alexander L Cohen, Steven M Nelson, Gagan S Wig, Kelly Anne Barnes, Jessica A Church, Alecia C Vogel, Timothy O Laumann, Fran M Miezin, Bradley L Schlaggar, et al. Functional network organization of the human brain. *Neuron*, 72(4):665–678, 2011.

Theodore D Satterthwaite, Mark A Elliott, Kosha Ruparel, James Loughead, Karthik Prabhakaran, Monica E Calkins, Ryan Hopson, Chad Jackson, Jack Keefe, Marisa Riley, et al. Neuroimaging of the philadelphia neurodevelopmental cohort. *Neuroimage*, 86:544–553, 2014.

Xilin Shen, Fuyuze Tokoglu, Xenios Papademetris, and R Todd Constable. Groupwise whole-brain parcellation from resting-state fmri data for network node identification. *Neuroimage*, 82:403–415, 2013.

Joshua S Siegel, Anish Mitra, Timothy O Laumann, Benjamin A Seitzman, Marcus Raichle, Maurizio Corbetta, and Abraham Z Snyder. Data quality influences observed links between functional connectivity and behavior. *Cerebral cortex*, 27(9):4492–4502, 2017.

Giora Simchoni and Saharon Rosset. Using random effects to account for high-cardinality categorical features and repeated measures in deep neural networks. *Advances in Neural Information Processing Systems*, 34:25111–25122, 2021.

Giora Simchoni and Saharon Rosset. Integrating random effects in deep neural networks. *Journal of Machine Learning Research*, 24(156):1–57, 2023.

Giora Simchoni and Saharon Rosset. Integrating random effects in variational autoencoders for dimensionality reduction of correlated data. *arXiv preprint arXiv:2412.16899*, 2024.

Martin Simonovsky and Nikos Komodakis. Graphvae: Towards generation of small graphs using variational autoencoders. In *International conference on artificial neural networks*, pp. 412–422. Springer, 2018.

Stephen M Smith, Peter T Fox, Karla L Miller, David C Glahn, P Mickle Fox, Clare E Mackay, Nicola Filippini, Kate E Watkins, Roberto Toro, Angela R Laird, et al. Correspondence of the brain's functional architecture during activation and rest. *Proceedings of the national academy of sciences*, 106(31):13040–13045, 2009.

Yu Takagi, Jun-ichiro Hirayama, and Saori C Tanaka. State-unspecific patterns of whole-brain functional connectivity from resting and multiple task states predict stable individual traits. *NeuroImage*, 201:116036, 2019.

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.

Jiaxing Xu, Qingtian Bian, Xinhang Li, Aihu Zhang, Yiping Ke, Miao Qiao, Wei Zhang, Wei Khang Jeremy Sim, and Balázs Gulyás. Contrastive graph pooling for explainable classification of brain networks. *IEEE Transactions on Medical Imaging*, 43(9):3292–3305, 2024.

Kwangsun Yoo, Monica D Rosenberg, Young Hye Kwon, Dustin Scheinost, R Todd Constable, and Marvin M Chun. A cognitive state transformation model for task-general and task-specific subsystems of the brain connectome. *NeuroImage*, 257:119279, 2022.

Manzil Zaheer, Satwik Kottur, Siamak Ravanbakhsh, Barnabas Poczos, Russ R Salakhutdinov, and Alexander Smola. Deep sets. *Advances in neural information processing systems*, 30, 2017.

A COHORTS AND PREPROCESSING

HCP. Data were drawn from the S1200 Young Adult release (Van Essen et al., 2013), acquired on the customized Siemens 3 T Connectome Skyra with a 32-channel receive coil. Whole-brain BOLD imaging used multiband gradient-echo EPI with TR/TE $= 720/33.1$ ms, flip angle $52°$, multiband factor 8, and 2-mm isotropic voxels. The time-series archive comprised two resting-state sessions and seven task domains (emotion, gambling, language, motor, relational processing, social cognition, and working memory), yielding nine paradigms. Left-to-right phase-encoding runs were retained to maintain a common acquisition direction. Images had undergone the HCP minimal preprocessing pipeline, including correction of gradient nonlinearities, head motion, and EPI distortion, registration to the structural image, transformation to MNI space, and intensity normalization (Glasser et al., 2013). Linear trends were then removed and the time series were band-pass filtered at 0.01–0.25 Hz, with no additional spatial smoothing. Regional signals were obtained by averaging voxel time courses within each node of the Shen 268-region atlas (Shen et al., 2013). HCP age is released in four ordered five-year bins (bin counts 192, 371, 294, and 8; reference midpoint 23.5); the age summary in Table 5 is the mean $\pm$ SD of bin midpoints.

PNC. Imaging was acquired on a single Siemens TIM Trio 3 T system with a 32-channel head coil (Satterthwaite et al., 2014). Resting-state, emotion-identification, and fractal n-back scans used whole-brain gradient-echo EPI (TR/TE $= 3000/32$ ms, flip angle $90°$, field of view 192×192 mm, 64×64 matrix, 46 contiguous 3-mm slices, 3-mm isotropic voxels); scan length ranged from 124 to 231 volumes across paradigms. All three paradigms were processed with the same SPM12-based workflow (Flandin & Friston, 2008): realignment for head motion, normalization to MNI space, and smoothing with a 3-mm FWHM Gaussian kernel. Six rigid-body motion parameters were regressed from the BOLD series, followed by temporal filtering at 0.01–0.1 Hz. Mean regional signals were extracted from 5-mm-radius spheres centered at the 264 Power coordinates (Power et al., 2011). To limit motion-related contamination, participants whose maximum translational displacement exceeded 1.5 mm were excluded during upstream preprocessing.

Table 5: Cohorts and analysis samples. Counts are participants entering the five-fold cross-validated experiments; every participant contributes all paradigms. HCP age is released in five-year bins, so its value is the mean $\pm$ SD of bin midpoints; PNC ages are exact.

Cohort	Participants	Paradigms	Scans	ROIs	FC edges	Sex (F/M)	Age (years)
HCP	865	9	7,785	268	35,778	457/408	28.8 ± 3.6
PNC	1,272	3	3,816	264	34,716	679/593	14.5 ± 3.4 (8–21)

Model inputs and targets. Each ROI series was standardized within scan. Static FC was calculated across the entire scan with Pearson correlation, Fisher z-transformed, and vectorized over the strict upper triangle. Dynamic graphs used 30-time-point windows with stride 5; at most 64 windows were retained for HCP and all PNC windows were used. Behavioral and demographic variables were kept outside the autoencoding objective. The two parcellations define different edge sets and the cohorts were modeled independently, so raw reconstruction errors are compared within cohort only.

B IMPLEMENTATION AND TRAINING DETAILS

Cross-validation. Each participant and all of that participant's scans were assigned to a single outer fold. Five-fold cross-validation was stratified by sex when class counts permitted. Within each outer development set, 12.5% of participants were reserved for validation, giving approximately 70%/10%/20% train/validation/test partitions. Split generation was fixed with seed 20260818, and every model was trained with seeds 13, 42, and 2026 under deterministic execution. Preprocessing estimators, PCA bases, probe standardizers, and hyperparameter choices were fitted on the training partition; validation data were used for model selection, and each outer test fold was evaluated once.

Architecture and optimization. The graph encoder used two signed graph-convolution layers with hidden width 128, a 64-dimensional scan representation, dropout 0.1, and a 64-dimensional participant random-effect code ($d_b = 64$). The edge decoder used 32-dimensional ROI embeddings and

a three-layer MLP with hidden width 256. Models were optimized with Adam at a learning rate of 10^{-3}, gradient-norm clipping at 5, and early stopping with patience 20. The HCP and PNC main models used a maximum budget of 200 epochs; component ablations used a common 200-epoch cap, including a separately trained full-model reference under the same cap. The checkpoint with the lowest validation participant-averaged MSE was retained. All models were trained on a single workstation with two NVIDIA GeForce RTX 5090 GPUs, using Python 3.14.4 and PyTorch 2.13.0 (CUDA 13.0, cuDNN 9.2), together with NumPy 2.5.2, SciPy 1.18.0, pandas 3.0.5, scikit-learn 1.9.0, and PyArrow 25.0.1.

Metrics. For each scan, MSE, MAE, and Pearson correlation were computed across FC edges and then averaged within participant. For each training seed, the five outer test folds were concatenated to obtain one OOF estimate, followed by the mean and standard deviation across the three seeds. Confidence intervals were obtained by resampling participants (10,000 bootstrap resamples), and permutation tests used 1,000 participant-level permutations. Benjamini–Hochberg FDR correction was applied within each prespecified family of tests. Missing phenotypes were handled by complete-case analysis. All tests were two-sided.

C Baselines and ablation variants

Baselines. The *paradigm template* assigns each test scan the mean FC vector of the corresponding paradigm, estimated exclusively from training participants; it has no trainable parameters and is reported without a seed SD. *Graph-AE* reconstructs the connectome through a deterministic scan representation (Kipf & Welling, 2016), and *Graph-VAE* introduces a scan-level stochastic latent code (Kipf & Welling, 2016; Simonovsky & Komodakis, 2018). We reimplemented both models from the original descriptions, adapted them to connectome reconstruction, and retrained them on HCP and PNC with the same seeds, early-stopping rule, and checkpoint selection used for P-MEGA. *LMM-VAE* (Simchoni & Rosset, 2024), implemented with a 64-dimensional latent representation, serves as the probabilistic mixed-effects representation-learning baseline. All learned models used the same participant-level partitions and FC reconstruction targets. The *population-only* condition sets $\mathbf{b}_s = \mathbf{0}$ in the trained P-MEGA model at inference, retaining the scan-conditioned nonlinear prediction while removing the inferred participant deviation; the *complete* condition uses the posterior mean inferred from all available scans of the participant. This is an autoencoding analysis, so the target scan may contribute to the participant posterior; reconstruction from other paradigms after excluding the target is examined by the held-out-support experiment.

Ablation variants. All variants were retrained under a common 200-epoch budget with one component changed at a time. *No participant random effect* removes the participant pathway. *Deterministic participant effect* keeps the architecture and uses a deterministic low-rank code (the batch-centered participant mean of per-scan proposals), with the participant auxiliary losses omitted. *Direct high-dimensional effect* produces $\mathbf{b}_s \in \mathbb{R}^M$ directly, bypassing the shared low-rank basis $\mathbf{W}_b$. *Static adjacency* holds the propagation graph at full-scan FC while retaining window features and window pooling; this variant isolates the adjacency choice while preserving windowed features. *Dynamic-focused encoder* uses window-specific adjacency and removes the static-FC node feature.

D Per-paradigm reconstruction

Table 6 reports reconstruction error and edgewise correlation for each paradigm, comparing the population-only prediction with the complete P-MEGA prediction.

Table 6: Per-paradigm OOF connectome reconstruction, five folds summarized across three training seeds (mean ± SD across seeds). Within a seed, paradigm metrics are pooled over the five outer test folds, weighted by the number of test scans. MSE is scan-averaged within the paradigm; Pearson correlation is computed across edges and averaged over scans. Reduction is the within-paradigm relative decrease in MSE from the population-only prediction to the complete P-MEGA prediction.

	MSE ↓			Pearson ↑	
Paradigm	Population only	Complete	Reduction (%)	Population only	Complete
HCP					
Emotion	0.03352 ± 0.00020	0.02361 ± 0.00008	29.55 ± 0.18	0.532 ± 0.007	0.681 ± 0.001
Gambling	0.03076 ± 0.00011	0.02085 ± 0.00008	32.23 ± 0.29	0.567 ± 0.007	0.715 ± 0.001
Language	0.03072 ± 0.00006	0.02078 ± 0.00015	32.35 ± 0.59	0.489 ± 0.008	0.691 ± 0.001
Motor	0.03567 ± 0.00013	0.02526 ± 0.00012	29.18 ± 0.24	0.580 ± 0.005	0.701 ± 0.001
Relational	0.03329 ± 0.00015	0.02310 ± 0.00010	30.61 ± 0.14	0.632 ± 0.004	0.735 ± 0.001
Rest 1	0.02396 ± 0.00015	0.01447 ± 0.00004	39.61 ± 0.51	0.569 ± 0.008	0.782 ± 0.001
Rest 2	0.02163 ± 0.00017	0.01233 ± 0.00003	43.02 ± 0.55	0.495 ± 0.012	0.774 ± 0.000
Social	0.03136 ± 0.00001	0.02148 ± 0.00010	31.50 ± 0.31	0.573 ± 0.006	0.713 ± 0.001
Working memory	0.02657 ± 0.00006	0.01679 ± 0.00009	36.82 ± 0.37	0.628 ± 0.007	0.762 ± 0.001
PNC					
Emotion identification	0.04662 ± 0.00012	0.03619 ± 0.00008	22.38 ± 0.12	0.441 ± 0.007	0.619 ± 0.001
n-back	0.04351 ± 0.00011	0.03344 ± 0.00003	23.16 ± 0.16	0.381 ± 0.010	0.609 ± 0.001
Rest	0.05528 ± 0.00008	0.04545 ± 0.00003	18.83 ± 0.12	0.399 ± 0.005	0.569 ± 0.001

E SUPPORT-SWAP PROTOCOL AND PAIRED DIFFERENCES

For each eligible participant and target paradigm, all scans from the target paradigm are excluded from the support set. *True support* draws the support paradigms from the same participant; *matched support* draws the identical paradigms from one different participant matched on sex and nearest training-standardized age; *random support* draws them from one randomly selected different participant. A single donor supplies every scan in a pseudo-support set, so the three conditions share support count and paradigm composition. PNC support subsets were enumerated exhaustively; for HCP, up to four subsets were sampled at each support count with a fixed permutation seed.

Let $\mathbf{r}_{sp} = \mathbf{x}_{sp} - \mathbf{m}_{sp}^{\text{pop}}$ be the fixed target residual and $\widehat{\mathbf{b}}_{sp}^{(c)}$ the effect inferred under support condition c. We report residual correlation together with the per-edge MSE reduction

$$G_{sp}^{(c)} = \frac{1}{M}\left(\|\mathbf{r}_{sp}\|_2^2 - \|\mathbf{r}_{sp} - \widehat{\mathbf{b}}_{sp}^{(c)}\|_2^2\right). \tag{10}$$

The identity-specific gain against a control donor is

$$\Delta_{sp}^{(c)} = G_{sp}^{(\text{true})} - G_{sp}^{(c)} = \frac{1}{M}\left(\|\mathbf{r}_{sp} - \widehat{\mathbf{b}}_{sp}^{(c)}\|_2^2 - \|\mathbf{r}_{sp} - \widehat{\mathbf{b}}_{sp}^{(\text{true})}\|_2^2\right), \quad c \in \{\text{matched}, \text{random}\}. \tag{11}$$

Positive $\Delta_{sp}^{(c)}$ favors true support. Paired participant-level differences were summarized with 10,000-resample participant bootstrap CIs and FDR-corrected q values. Figure 3 shows the paired differences.

F DOWNSTREAM PREDICTION PROTOCOL

Task-specific prediction heads were attached to either the population representation $\mathbf{u}$ alone or the combined $[\mathbf{u}\|\boldsymbol{\mu}_s]$ representation, and the relevant encoder pathways were fine-tuned end to end within each outer training fold. Hyperparameters were selected on the validation partition, and outer test predictions were pooled once per training seed. Sex prediction was evaluated with AUROC; fluid reasoning in HCP (`PMAT24_A_CR`) and reading achievement in PNC (`WRAT_CR_STD`) were evaluated with RMSE and MAE after standardizing the outcomes with training-partition statistics. Baseline representations (Graph-AE, Graph-VAE, LMM-VAE) used their learned scan representation $\mathbf{u}$ under the same protocol. Both P-MEGA configurations were fine-tuned end to end, so the comparison assesses task-adapted prediction systems; the scan representation can retain individual information, and the participant representation may share predictive content with it, which offers one explanation for the outcome-dependent pattern in Table 4.

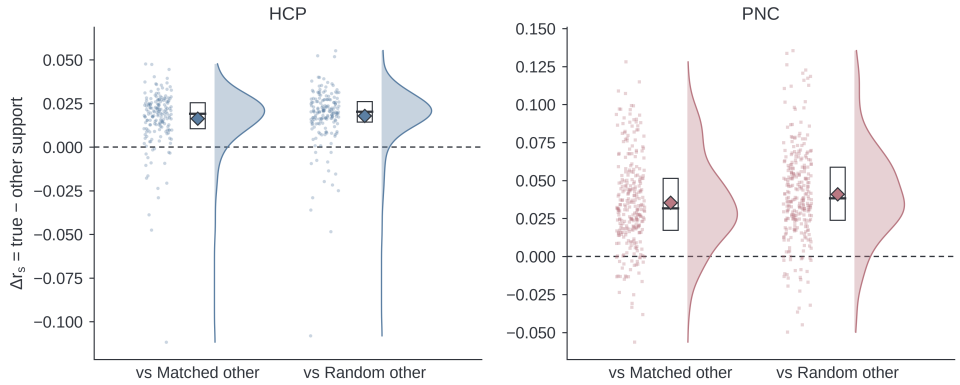

Figure 3: Paired participant-level differences in residual correlation between true same-participant support and matched- or random-other support in HCP and PNC. Diamonds mark mean estimates, boxplots summarize participant-level distributions, and the dashed line denotes zero difference.

G Network-level analyses

G.1 Methods

FC edges were grouped according to cohort-specific ROI-to-network assignments (Figure 4) after matching ROI indices and ordering to the corresponding parcellation. All analyses used OOF representations. For each training seed, the five outer test folds were concatenated to form a complete OOF aggregate in which each participant appeared exactly once; except in the stability analyses, participant-level quantities were averaged across the three training seeds before inference. Confidence intervals used 10,000 participant-level bootstrap resamples.

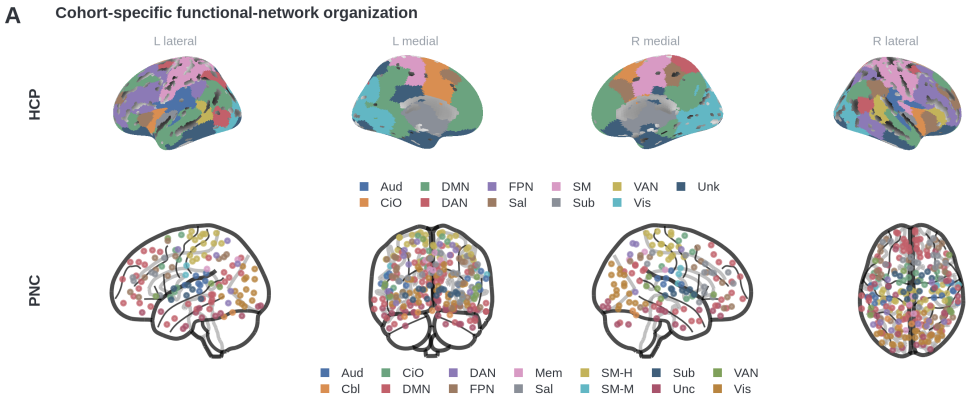

Figure 4: Cohort-specific anatomical and functional-network organization. HCP uses the Shen-268 parcellation, shown on the cortical surface and colored by the functional-network assignments used in the connectome analyses. PNC uses the Power-264 parcellation, shown at the corresponding MNI coordinates and colored by its cohort-specific network assignments. The panel provides anatomical context and displays no statistical result.

Population task contrasts. Within-participant task–rest differences were computed in the network-pair summaries of $\mathbf{m}_{sp}^{\mathrm{pop}}$. HCP contrasts included motor, language, working memory, and emotion relative to rest 1; PNC contrasts included emotion identification and working memory relative to rest. HCP age entered the adjustment models as three treatment-coded dummy variables for the four ordered bins, together with sex; PNC contrasts were adjusted for continuous age and sex. Two-sided significance used 1,000 participant-level sign permutations of the adjusted within-participant differences, with Benjamini–Hochberg FDR correction applied jointly to the task-contrast tests within each cohort.

Participant-effect variability and stability. The empirical variability of each FC edge was defined as the across-participant standard deviation of the seed-averaged $\mathbf{b}_s$, and network-pair maps aggregated these edge-wise standard deviations within and between functional networks. Across-seed stability was computed from the three complete OOF maps, and across-fold stability from five fold-specific maps after averaging the three seeds within each fold. The empirical map was also compared with the marginal standard deviation implied by the learned low-rank basis, $\sqrt{\sum_k W_{b,km}^2}$.

Brain–behavior associations. Network summaries of $\mathbf{b}_s$ were related to cognition and age with complete-case partial correlations. HCP fluid reasoning was adjusted for age-bin indicators and sex, PNC reading achievement for continuous age and sex, and age associations for sex. HCP age is ordinal and unevenly distributed across four bins, so its primary analysis used a partial Spearman correlation after rank-transforming outcome, network summary, and covariates, with the midpoint-coded partial Pearson result as a sensitivity analysis; PNC age (8–21 years) was analyzed as a continuous outcome. Two-sided p values came from the partial-correlation test with $n - k - 2$ degrees of freedom (k covariates), and FDR correction was applied separately to the network-pair associations for each phenotype. To assess whether associations reflect regional organization beyond overall fingerprint magnitude, each analysis was repeated with the centered global RMS

$$r_s^{\text{global}} = \frac{\|\mathbf{b}_s - \overline{\mathbf{b}}\|_2}{\sqrt{M}}, \qquad \overline{\mathbf{b}} = \frac{1}{S}\sum_{s=1}^{S} \mathbf{b}_s, \tag{12}$$

included as an additional covariate; these sensitivity tests formed separate phenotype-specific FDR families.

G.2 Results

Table 7: Network organization and phenotypic associations of the population and participant pathways. Task and phenotype estimates are based on seed-averaged OOF participant-level results with 10,000-resample bootstrap 95% CIs; for stability analyses the CI column reports the observed range of pairwise Spearman correlations. Numbers in parentheses give FDR-significant network pairs among all tested pairs.

Cohort	Pathway	Analysis	Leading network-pair result	Estimate	95% CI or range	q
Population task contrasts						
HCP	Population	Motor vs. rest 1	Dorsal-Attention–Salience (65/66)	$+0.127$	$[+0.120, +0.134]$	1.1×10^{-3}
HCP	Population	Language vs. rest 1	Auditory–Auditory (63/66)	-0.106	$[-0.117, -0.096]$	1.1×10^{-3}
HCP	Population	Working memory vs. rest 1	Frontal-Parietal–Frontal-Parietal (61/66)	$+0.121$	$[+0.113, +0.128]$	1.1×10^{-3}
HCP	Population	Emotion vs. rest 1	Auditory–Auditory (56/66)	-0.088	$[-0.099, -0.078]$	1.1×10^{-3}
PNC	Population	Emotion identification vs. rest	Somatomotor Hand–Somatomotor Hand (84/105)	-0.053	$[-0.063, -0.042]$	1.5×10^{-3}
PNC	Population	Working memory vs. rest	Somatomotor Hand–Somatomotor Hand (95/105)	-0.050	$[-0.060, -0.040]$	1.5×10^{-3}
Participant-effect map stability						
HCP	Participant	Across training seeds	Mean pairwise Spearman correlation	0.983	[0.979, 0.987]	–
HCP	Participant	Across outer folds	Mean pairwise Spearman correlation	0.972	[0.955, 0.987]	–
PNC	Participant	Across training seeds	Mean pairwise Spearman correlation	0.984	[0.982, 0.988]	–
PNC	Participant	Across outer folds	Mean pairwise Spearman correlation	0.957	[0.939, 0.968]	–
Brain–behavior associations						
HCP	Participant	Fluid reasoning	Default Mode–Visual (31/66)	$r = -0.199$	$[-0.262, -0.136]$	1.8×10^{-7}
PNC	Participant	Reading achievement	Somatomotor Mouth–Somatomotor Mouth (75/105)	$r = +0.252$	$[+0.207, +0.298]$	2.6×10^{-20}
HCP	Participant	Ordered age-bin trend	Frontal-Parietal–Somato-Motor (54/66)	$\rho = +0.229$	$[+0.167, +0.289]$	3.1×10^{-10}
PNC	Participant	Continuous age	Auditory–Cingulo-opercular (83/105)	$r = +0.485$	$[+0.445, +0.526]$	3.3×10^{-84}

Population pathway. The complete network-pair maps (Figure 7) reveal distributed task–rest differences in both cohorts. Between 56 and 65 of the 66 HCP network pairs survived FDR correction within each contrast (245 of 264 tests across the four contrasts); in PNC, 84 of 105 pairs were significant for emotion identification versus rest and 95 of 105 for working memory versus rest. The largest HCP effect was increased Dorsal-Attention–Salience coupling during motor relative to

rest 1 (+0.127, 95% CI [+0.120, +0.134]); the largest PNC working-memory effect was reduced within-network Somatomotor Hand coupling (−0.050, 95% CI [−0.060, −0.040]). Effect directions agreed across all three training seeds for 260 of 264 HCP tests (98.5%) and 207 of 210 PNC tests (98.6%), so the task-contrast patterns are largely insensitive to initialization. The increased within-FPN component during working memory in HCP is compatible with a role for control networks in task-dependent connectivity adjustment (Cole et al., 2014; 2013), and the motor-related Dorsal-Attention–Salience contrast shows that prominent reconstructed differences extend across network boundaries. PNC shows a different pattern, including reduced within-Somatomotor Hand components during both tasks; these results favor interpretation at the level of specific cohort–task combinations, with acquisition and preprocessing differences considered alongside task demands.

Participant pathway. The individual-variation fingerprints exhibit reproducible network organization (Table 7; Figure 5). Network-pair variability rankings were concordant across training seeds (mean pairwise Spearman 0.983 in HCP and 0.984 in PNC) and across outer folds (0.972 and 0.957), so the spatial distribution of inferred participant effects generalized across held-out participant subsets. Within-network edges showed greater across-participant variability than between-network edges in both cohorts (mean SD 0.0230 versus 0.0169 in HCP; 0.0479 versus 0.0352 in PNC), a spatial heterogeneity with broad precedent in studies of individual FC variation (Mueller et al., 2013). The empirical variability maps agreed with the marginal SD implied by the learned basis ($r = 0.799$ at the edge level and 0.916 at the network-pair level in HCP; 0.749 and 0.903 in PNC). Both quantities depend on $\mathbf{W}_b$, so this agreement demonstrates internal consistency between the inferred effects and the covariance parameterization; independent neurobiological validation of the variance map is a separate question.

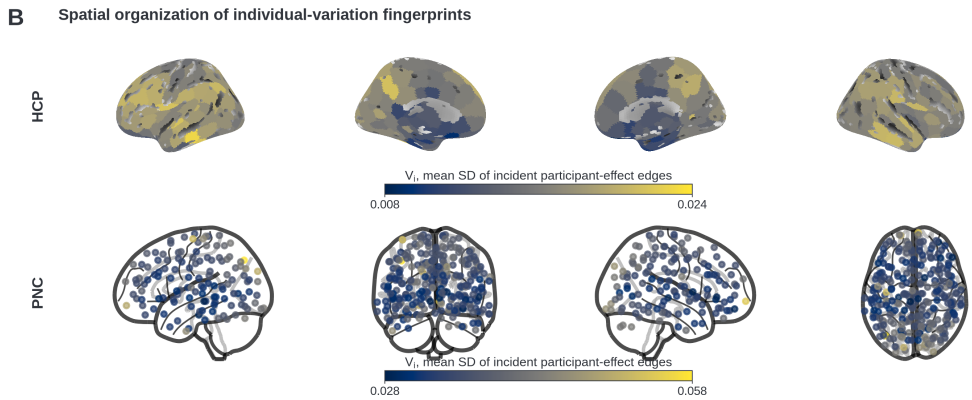

Figure 5: Spatial organization of individual-variation fingerprints. For each FC edge (i, j), participant-specific variability is the across-participant standard deviation of the seed-averaged OOF participant effect, $V_{ij} = \mathrm{SD}_s(b_{s,ij})$; each ROI is summarized as $V_i = (N-1)^{-1} \sum_{j \neq i} V_{ij}$. HCP and PNC use their respective anatomical representations and cohort-specific scales. Maps are unthresholded and descriptive; V_i summarizes variability in participant-specific connectivity, distinct from regional activation or local neural variance.

Brain–behavior associations. In HCP, 31 of 66 network-pair summaries were related to fluid reasoning after adjustment for age-bin indicators and sex (Figure 8), distributed across multiple functional systems, with the largest association for Default Mode–Visual connectivity ($r = -0.199$, 95% CI [−0.262, −0.136], $q = 1.8 \times 10^{-7}$). In PNC, 75 of 105 network-pair summaries were associated with reading achievement after adjustment for age and sex (Figure 9), with both positive and negative associations and the largest within the Somatomotor Mouth network ($r = +0.252$, 95% CI [+0.207, +0.298], $q = 2.6 \times 10^{-20}$); earlier work linked reading competence to resting-state connectivity among motor regions and between language regions (Koyama et al., 2011), which offers context at the level of the implicated systems. Figure 6 shows the FDR-significant network pairs as connectograms. Age-related organization was evident in both cohorts (Figure 10): 54 of 66 HCP network pairs showed monotonic associations across ordered age bins (largest Frontal-Parietal–Somato-Motor, $\rho = +0.229$, 95% CI [+0.167, +0.289], $q = 3.1 \times 10^{-10}$; this describes variation across ordered bins, without converting the effect to change per year), and 83 of 105 PNC pairs were associated with continuous age (largest Auditory–Cingulo-opercular, $r = +0.485$, 95%

CI $[+0.445, +0.526]$, $q = 3.3 \times 10^{-84}$). Primary and sensitivity estimators agreed closely across network pairs, so the age-related organization is robust to the age parameterization.

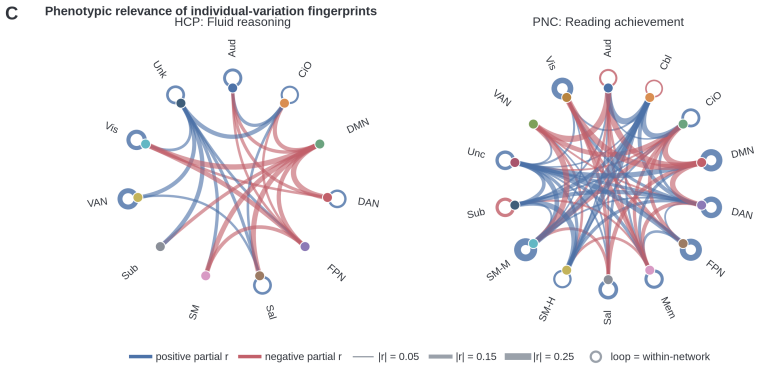

Figure 6: Phenotypic relevance of individual-variation fingerprints. Connectograms show network-pair partial correlations between the participant-specific connectivity component and the primary cognitive phenotype in each cohort (fluid reasoning in HCP; reading achievement in PNC), adjusted for age-bin indicators and sex in HCP and for continuous age and sex in PNC. Only network pairs surviving Benjamini–Hochberg FDR correction are shown; blue and red edges denote positive and negative partial correlations, edge width represents $|r|$, and loops denote within-network associations.

Global-magnitude sensitivity. Adjustment for global fingerprint magnitude (Figure 11) retained all 31 HCP fluid-reasoning associations, with the largest adjusted association for Default Mode–Dorsal-Attention connectivity ($r = -0.204$). In PNC, 70 of 105 network pairs remained significant for reading achievement, including 65 of the 75 primary pairs; ten primary associations fell below the FDR threshold, five additional pairs became significant, and the largest within-network Somatomotor Mouth association decreased from $r = +0.252$ to $r = +0.213$. Primary and globally adjusted estimates remained broadly concordant, so overall fingerprint magnitude accounts for part of the PNC reading association while substantial network-specific variation remains. Age associations were similarly robust (Figure 12): all 54 HCP age-bin associations were retained, and 85 of 105 PNC pairs were significant after adjustment, including 82 of the 83 primary pairs.

Interpretation. Task modulation and participant variability describe two dimensions of network organization: a network can show substantial task-related change and also carry individual differences across repeated observations. In contrasts formed with a common participant estimate, the additive participant term cancels algebraically, so the informative features of the population contrasts are their topology, magnitude, and reproducibility, whereas the participant-effect maps address where inferred deviations vary most across people. The signs of the phenotypic associations describe the relationship between phenotype and the inferred FC deviation; their translation into neural mechanisms requires additional evidence, and motion-sensitive and task-response-sensitive analyses (Siegel et al., 2017; Cole et al., 2019) remain important extensions, particularly for the widespread task contrasts and the developmental associations.

G.3 COMPLETE NETWORK-PAIR FIGURES

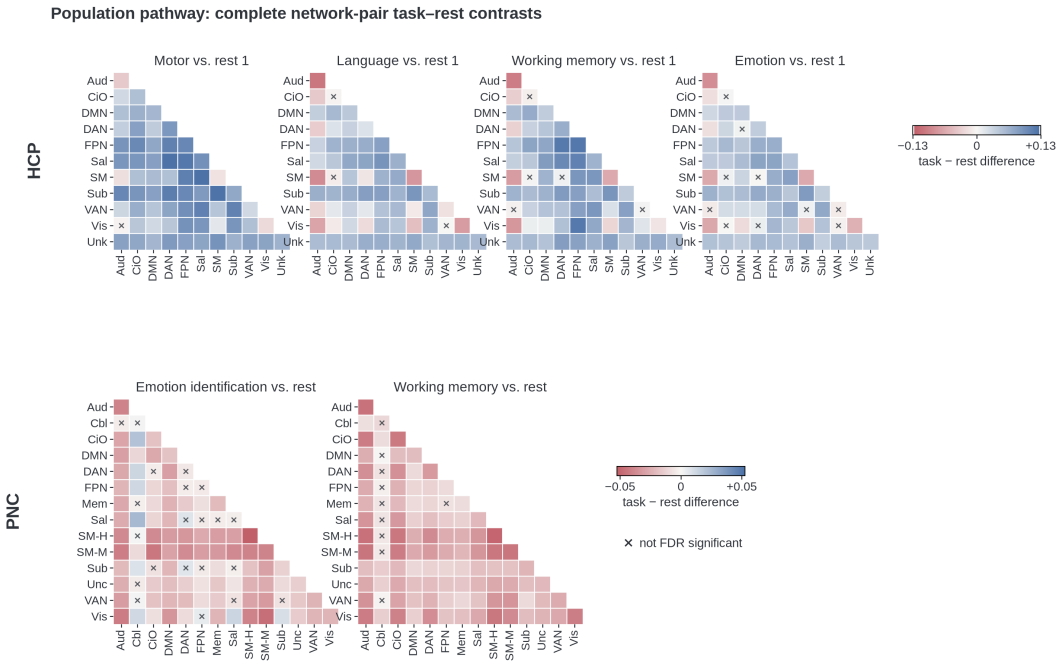

Figure 7: Complete network-pair results for the population-pathway task contrasts. Each cell is one network pair, with diagonal cells for within-network connectivity; color gives the seed-averaged OOF within-participant task–rest difference. HCP contrasts comprise motor, language, working memory, and emotion relative to rest 1; PNC contrasts comprise emotion identification and working memory relative to rest. HCP models were adjusted for age-bin indicators and sex, PNC models for continuous age and sex. Two-sided significance used 1,000 participant-level sign permutations followed by FDR correction within each cohort's task-contrast family; a gray × marks a pair that failed FDR correction. Overall, 245 of 264 HCP tests and 179 of 210 PNC tests were significant. Color scales are cohort-specific (Shen-268 with 11 networks; Power-264 with 14 networks).

H DERIVATIONS FOR THE PARTICIPANT EFFECT

This section collects analytical properties of the conditional model. The implementation uses the amortized posterior in Eq. equation 7; the closed-form expressions below serve as references for interpretation.

H.1 CONDITIONAL ELBO

Write $\mathbf{B} = \mathbf{W}_b^\top \in \mathbb{R}^{M \times d_b}$ and define the residual of the population pathway as

$$\mathbf{r}_{sp} = \mathbf{x}_{sp} - \mathcal{D}_\theta(\mathbf{u}_{sp}). \tag{13}$$

For participant s, the conditional joint density is

$$p(\{\mathbf{x}_{sp}\}_{p\in\mathcal{O}_s}, \mathbf{a}_s \mid \{\mathbf{u}_{sp}\}_{p\in\mathcal{O}_s}) = \mathcal{N}(\mathbf{a}_s; \mathbf{0}, \mathbf{I}) \prod_{p\in\mathcal{O}_s} \mathcal{N}(\mathbf{x}_{sp}; \mathcal{D}_\theta(\mathbf{u}_{sp}) + \mathbf{B}\mathbf{a}_s, \sigma_x^2 \mathbf{I}_M). \tag{14}$$

For $q_s = \mathcal{N}(\boldsymbol{\mu}_s, \mathrm{diag}(\boldsymbol{\sigma}_s^2))$, Jensen's inequality gives

$$\log p(\{\mathbf{x}_{sp}\}_p \mid \{\mathbf{u}_{sp}\}_p) \geq \sum_{p\in\mathcal{O}_s} \mathbb{E}_{q_s}[\log p(\mathbf{x}_{sp} \mid \mathbf{u}_{sp}, \mathbf{a}_s)] - \mathrm{KL}(q_s \| p), \tag{15}$$

where

$$\mathrm{KL}(q_s \| p) = \frac{1}{2} \sum_{k=1}^{d_b} \left(\mu_{sk}^2 + \sigma_{sk}^2 - 1 - \log \sigma_{sk}^2\right). \tag{16}$$

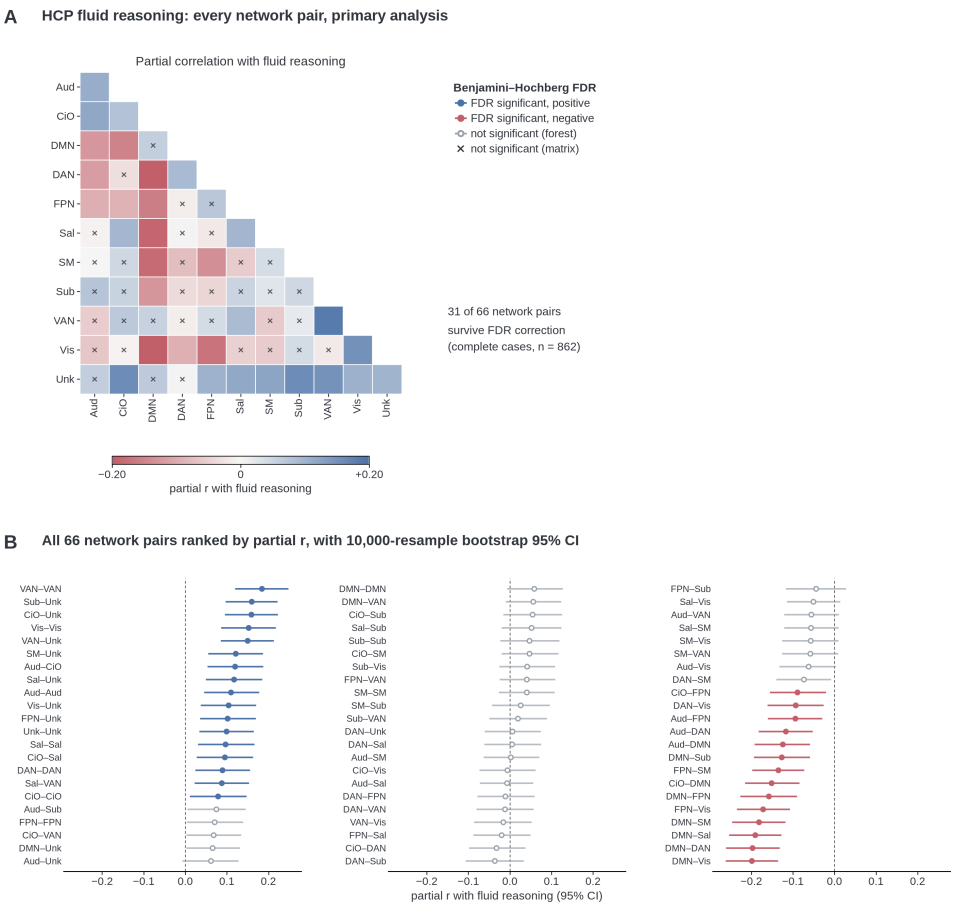

Figure 8: HCP fluid reasoning: primary network-pair analysis. (A) Network-pair partial correlations between the participant-effect summary $\mathbf{b}_s$ and fluid reasoning (`PMAT24_A_CR`), adjusted for age-bin indicators and sex; lower triangle of the 11×11 network matrix, with a gray $\times$ for pairs failing FDR correction. (B) All 66 network pairs ranked by partial r with 10,000-resample participant-bootstrap 95% CIs; filled markers denote FDR-significant associations. The analysis included 862 complete cases, and 31 of 66 pairs survived FDR correction.

The likelihood contributes once per scan, while the KL term contributes once per participant. This counting follows directly from the single latent effect shared by all scans of s.

H.2 Residual decomposition

Let $n_s = |\mathcal{O}_s|$ and $\bar{\mathbf{r}}_s = n_s^{-1} \sum_p \mathbf{r}_{sp}$. The expected reconstruction term admits the decomposition

$$\sum_p \mathbb{E}_{q_s} \|\mathbf{r}_{sp} - \mathbf{B}\mathbf{a}_s\|_2^2 = \sum_p \|\mathbf{r}_{sp} - \bar{\mathbf{r}}_s\|_2^2 + n_s \|\bar{\mathbf{r}}_s - \mathbf{B}\boldsymbol{\mu}_s\|_2^2 + n_s \operatorname{tr}\left(\mathbf{B} \operatorname{diag}(\boldsymbol{\sigma}_s^2) \mathbf{B}^\top\right). \tag{17}$$

Derivation. Insert $\mathbf{r}_{sp} - \mathbf{B}\mathbf{a}_s = (\mathbf{r}_{sp} - \bar{\mathbf{r}}_s) + (\bar{\mathbf{r}}_s - \mathbf{B}\mathbf{a}_s)$. The cross term vanishes after summing over p since $\sum_p (\mathbf{r}_{sp} - \bar{\mathbf{r}}_s) = \mathbf{0}$. Taking the Gaussian expectation yields

$$\mathbb{E}_{q_s} \|\bar{\mathbf{r}}_s - \mathbf{B}\mathbf{a}_s\|_2^2 = \|\bar{\mathbf{r}}_s - \mathbf{B}\boldsymbol{\mu}_s\|_2^2 + \operatorname{tr}\left(\mathbf{B} \operatorname{diag}(\boldsymbol{\sigma}_s^2) \mathbf{B}^\top\right), \tag{18}$$

which proves Eq. equation 17. With the population predictions fixed, the participant effect acts on the mean residual shared across scans. Joint training also updates the graph encoder and decoder, so the first term continues to shape the population pathway.

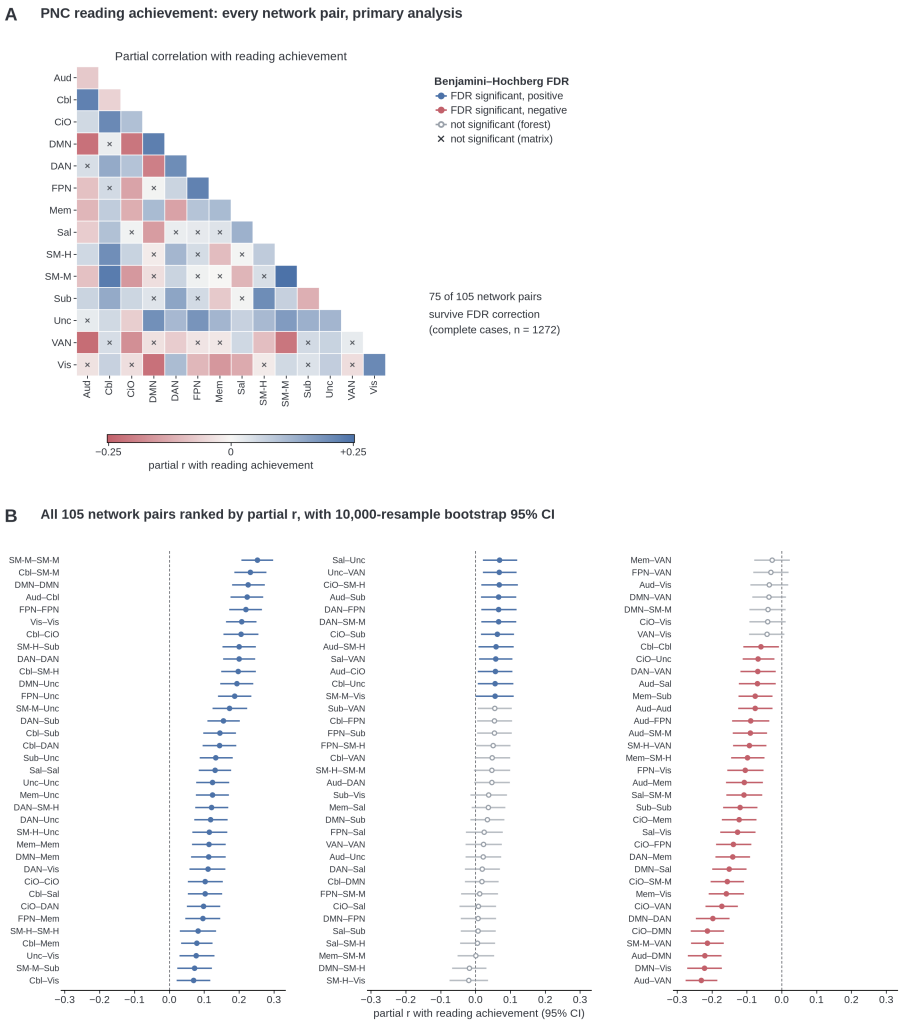

Figure 9: PNC reading achievement: primary network-pair analysis. (A) Network-pair partial correlations between $\mathbf{b}_s$ and reading achievement (`WRAT_CR_STD`), adjusted for continuous age and sex; lower triangle of the 14×14 network matrix, with a gray $\times$ for pairs failing FDR correction. (B) All 105 network pairs ranked by partial r with 10,000-resample participant-bootstrap 95% CIs; filled markers denote FDR-significant associations. The analysis included 1,272 complete cases, and 75 of 105 pairs survived FDR correction.

H.3 Exact conditional posterior and directional shrinkage

Gaussian conjugacy gives the exact posterior for a fixed global model:

$$\mathbf{V}_s^* = \left(\mathbf{I}_{d_b} + \frac{n_s}{\sigma_x^2} \mathbf{B}^\top \mathbf{B} \right)^{-1}, \tag{19}$$

$$\boldsymbol{\mu}_s^* = \frac{1}{\sigma_x^2} \mathbf{V}_s^* \mathbf{B}^\top \sum_{p \in \mathcal{O}_s} \mathbf{r}_{sp}. \tag{20}$$

Derivation. Terms in the log joint that depend on $\mathbf{a}_s$ are

$$-2 \log p(\mathbf{a}_s, \{\mathbf{r}_{sp}\}_p) = \mathbf{a}_s^\top \left(\mathbf{I}_{d_b} + \frac{n_s}{\sigma_x^2} \mathbf{B}^\top \mathbf{B} \right) \mathbf{a}_s - \frac{2}{\sigma_x^2} \mathbf{a}_s^\top \mathbf{B}^\top \sum_p \mathbf{r}_{sp} + C. \tag{21}$$

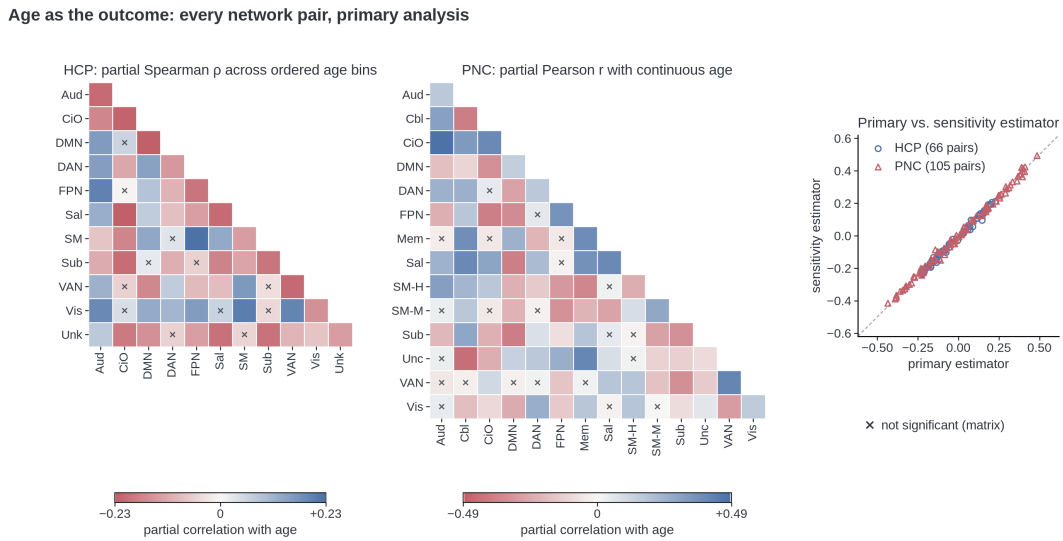

Figure 10: Age as the outcome: primary analyses and sensitivity estimators. Left and center panels show network-pair partial correlations with age, adjusted for sex; a gray × marks a pair failing FDR correction (54 of 66 HCP pairs and 83 of 105 PNC pairs were significant). HCP age is available as four ordered bins, so its primary estimator is a partial Spearman correlation across ordered bins, with a midpoint-coded partial Pearson correlation as sensitivity estimator; PNC age (8–21 years) is analyzed as a continuous outcome with partial Pearson correlation and partial Spearman correlation as sensitivity estimator. The right panel compares primary and sensitivity estimates for every network pair (dashed identity line). HCP associations represent variation across ordered age bins.

Completing the square produces Eqs. equation 19–equation 20.

For a compact singular value decomposition $\mathbf{B} = \mathbf{Q}\mathbf{\Lambda}\mathbf{R}^\top$, with singular values $\ell_j > 0$, the projected posterior mean becomes

$$\mathbf{B}\boldsymbol{\mu}_s^* = \mathbf{Q}\,\mathrm{diag}\left(\frac{n_s\ell_j^2}{\sigma_x^2 + n_s\ell_j^2}\right)\mathbf{Q}^\top\bar{\mathbf{r}}_s. \tag{22}$$

Each learned direction receives its own shrinkage factor. More scans increase the weight of the mean residual, while larger observation noise strengthens shrinkage toward zero.

Within a freely optimized diagonal variational family, the stationary coordinate variances are

$$(\sigma_{sk}^{\mathrm{MF}})^2 = \left(1 + \frac{n_s\|\mathbf{B}_{:k}\|_2^2}{\sigma_x^2}\right)^{-1}. \tag{23}$$

These values are reciprocal diagonal entries of the exact precision matrix. The learned set encoder approximates this conditional solution under amortization and a positive scale floor.

H.4 Marginal covariance across repeated scans

Integrating over the prior $\mathbf{a}_s \sim \mathcal{N}(\mathbf{0}, \mathbf{I})$ gives

$$\mathrm{Cov}(\mathbf{x}_{sp} - \mathcal{D}_\theta(\mathbf{u}_{sp}), \mathbf{x}_{sq} - \mathcal{D}_\theta(\mathbf{u}_{sq}) \mid \{\mathbf{u}\}) = \mathbf{B}\mathbf{B}^\top + \mathbb{1}\{p = q\}\sigma_x^2\mathbf{I}_M. \tag{24}$$

Indeed, the shared term contributes $\mathbf{B}\,\mathrm{Cov}(\mathbf{a}_s)\mathbf{B}^\top = \mathbf{B}\mathbf{B}^\top$ to every pair of scans from participant s. Observation noise contributes only when the two scan indices coincide. For two distinct participants, independent latent effects give zero cross-participant covariance under the conditional model.

H.5 Parameter symmetries and support size

For any orthogonal matrix $\mathbf{R}$, the transformation

$$\mathbf{B}' = \mathbf{B}\mathbf{R}^\top, \qquad \mathbf{a}_s' = \mathbf{R}\mathbf{a}_s \tag{25}$$

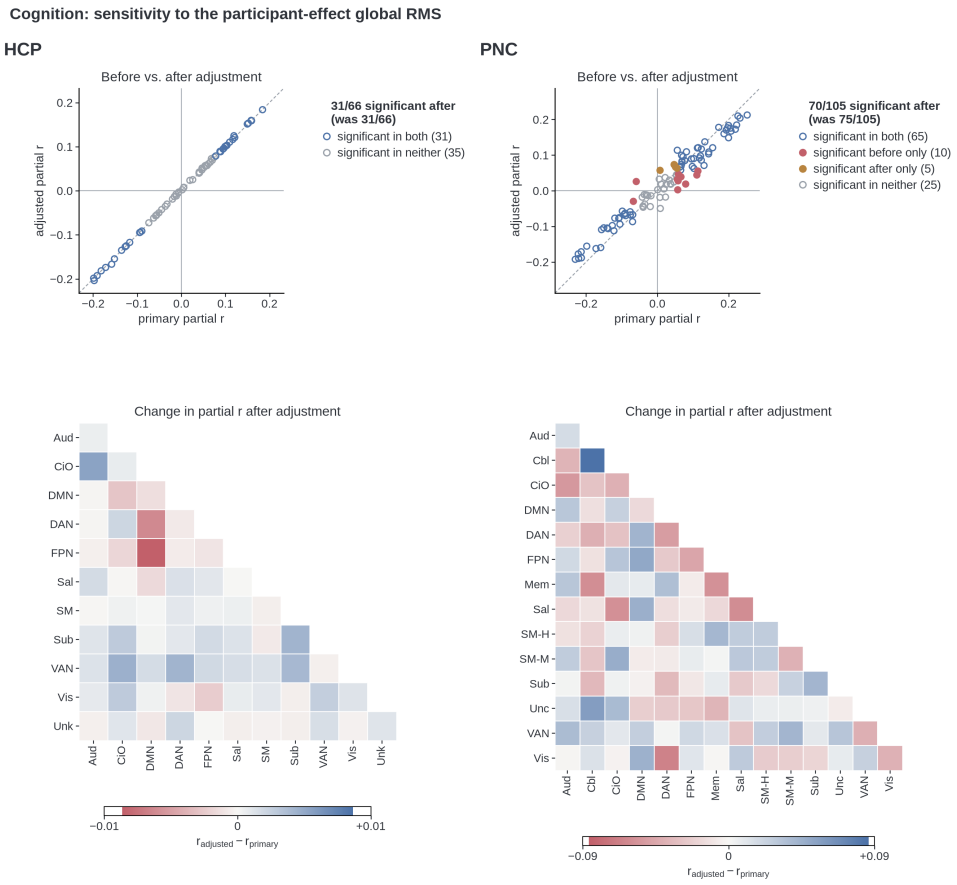

Figure 11: Cognitive associations after adjustment for participant-effect global RMS. Each brain–behavior analysis was repeated with the centered global RMS r_s^{global} as an additional covariate, forming separate phenotype-specific FDR families. Upper scatter plots compare primary and adjusted partial correlations for every network pair, with points categorized by changes in FDR status (dashed identity line); lower triangular matrices show the network-pair-specific change $r_{\text{adjusted}} - r_{\text{primary}}$ with cohort-specific color scales. In HCP, all 31 primary fluid-reasoning pairs remained significant; in PNC, 70 of 105 pairs were significant after adjustment for reading achievement, compared with 75 in the primary analysis (10 lost, 5 gained).

preserves $\mathbf{Ba}_s$ and the isotropic prior. Consequently, projected effects $\mathbf{b}_s = \mathbf{Ba}_s$ have a direct connectome-space meaning, while individual latent coordinates depend on the chosen orientation. The diagonal variational family restricts this symmetry during optimization, yet the coordinate-wise interpretation remains basis dependent.

The set encoder combines mean-pooled evidence with $\log K_s$. If every training participant has the same support count K_0, a linear count coefficient $\mathbf{w}$ and intercept $\mathbf{c}$ enter the posterior head through the single combination $\mathbf{w} \log K_0 + \mathbf{c}$. Their separate effects at new support counts are then underdetermined by the training data. Support-size behavior is reported empirically, and posterior-width trends are treated as model outputs with calibration assessed separately.

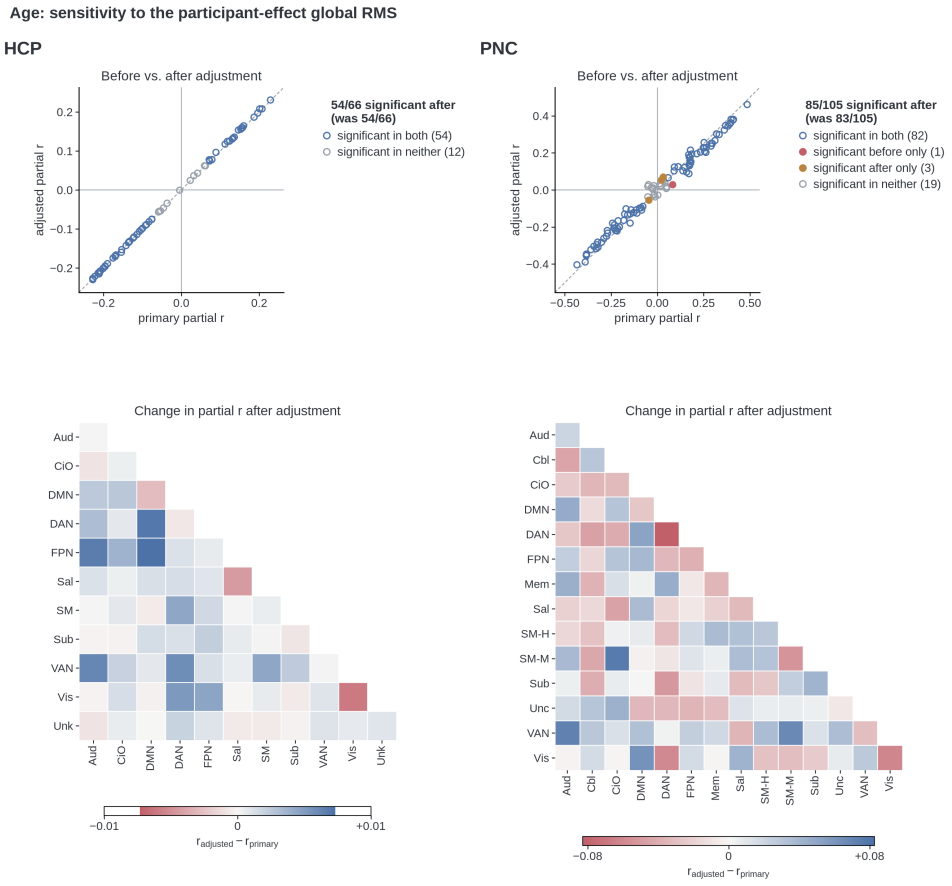

Figure 12: Age associations after adjustment for participant-effect global RMS. Each age-association analysis was repeated with r_s^{global} as an additional covariate, with adjusted tests forming separate phenotype-specific FDR families. Upper scatter plots compare primary and adjusted estimates for every network pair, categorized by changes in FDR status (dashed identity line); lower triangular matrices show the difference between adjusted and primary estimates. In HCP, all 54 primary age pairs remained significant; in PNC, the number of significant pairs increased from 83 to 85 of 105, with 82 of the 83 primary pairs retained.