

Cognitive cartography of mammalian brains using meta-analysis of AI experts

Andrea I. Luppi^{1,2,3,*}, Hana Ali², Zhen-Qi Liu¹, Filip Milisav¹, Alessandro Gozzi⁴, Danilo Bzdok^{1,5}, Bratislav Misic^{1†}

¹Montréal Neurological Institute, McGill University, Montréal, QC, Canada

²Centre for Eudaimonia and Human Flourishing, Department of Psychiatry, University of Oxford, Oxford, UK

³St John's College, University of Cambridge, Cambridge, UK

⁴Centre for Neuroscience and Cognitive Systems, Italian Institute of Technology, Rovereto, Italy

⁵MILA-Quebec Artificial Intelligence Institute, Montréal, QC, Canada

The complexity of the brain is increasingly mirrored by the complexity of the neuroscientific literature, yet no individual mind can fully grasp the diversity of scales, methodologies and model organisms. Where human experts flag, the latest AI models excel: large language models can seamlessly integrate knowledge across scientific domains. Here we show how large language models can systematically and quantitatively synthesise literature-wide neuroscientific knowledge about the cognitive operations and dysfunctions associated with each brain region. Meta-analysis of AI experts reveals structure-function mappings to which existing meta-analytic frameworks are blind, demonstrated by lesions and direct intracranial stimulation. It also unlocks the possibility of extending quantitative literature meta-analysis and decoding of brain maps to other model organisms beyond human. As proof of concept, we integrate LLM meta-analysis with species-specific transcriptomics in human, macaque, and mouse, to discover an evolutionarily conserved molecular circuit for cognition. Altogether, meta-analysis of AI experts can fundamentally catalyze neuroscientific discovery by overcoming the barrier of data aggregation from heterogeneous studies, finally bringing together a scattered literature to identify emergent patterns and latent insights across disparate subfields, modalities, and species.

INTRODUCTION

A fundamental goal of neuroscience is to understand how the architecture of the brain enables cognition. Yet no single mind can grasp the deluge of findings about causal and correlational associations between brain structure, function, and dysfunction obtained from neuroimaging, brain stimulation, and brain lesions in humans and other animals (1–10). One way to make sense of this vast and tangled literature is through meta-analysis: the de facto strategy for evidence synthesis in science, offering a structured approach to integrate neuroimaging studies and synthesise brain maps of the anatomical contributions to each cognitive operation (11–24).

Manually-curated meta-analysis (11, 25) is extremely laborious, and increasingly unfeasible given the exponential growth of the literature. However automated engines for coordinate-based meta-analysis (12, 13) are largely or entirely restricted to MRI or PET, and only available for human studies (19, 22–24, 26). These restrictions preclude translational discovery from key animal models, and from huge swaths of the literature using other methods for recording, imaging, and manipulating brain function.

Where human experts flag, the latest AI models excel: large language models (LLMs) based on the transformer architecture (27–32). Popular models such as OpenAI's Generative Pre-trained Transformer 4

(GPT-4) (33), Mistral(34), Meta's Llama (35), Google's Gemini (36) and DeepSeek (37) are trained on vast corpora of text that include the entire scientific literature, effectively making them 'AI experts' across multiple domains of knowledge. Indeed, LLMs demonstrate a remarkable capacity to synthesise related technical information and high-level concepts across diverse disciplines – from genomics to medicine to cognitive science (38–42) – and across different techniques, model organisms, and sub-fields (40, 43–46). In neuroscience, recent work revealed that popular LLMs exhibit remarkable domain-expertise, systematically outperforming human experts at predicting the outcome of neuroscience experiments (45).

Here we introduce fully automated meta-analysis of AI experts for neuroscience, combining key advantages of automation and domain-expertise. We use off-the-shelf LLMs to systematically synthesise literature-wide neuroscientific knowledge about the cognitive functions and dysfunctions associated with each brain region. This approach takes advantage of the vast knowledge corpus embedded in LLMs to automatically produce accurate mappings between brain regions and cognitive operations, without restrictions to specific techniques or modalities. We combine and compare 5 popular AI experts: OpenAI's GPT-4o-mini, Google's Gemini 2.0 Flash, Mistral AI's Mistral-7B-Instruct-v0.3, Meta's Llama-3.3-70B-Instruct-Turbo, and DeepSeek-R1-Distill-Llama-70B. We extensively benchmark brain maps generated by LLMs against functional, structural, metabolic, connectomic, transcriptomic, and chemo-architectonic databases (1, 47). We show that AI experts' summaries of the literature transcend current meta-analytic engines (12–14) to

* andrea.luppi@psych.ox.ac.uk

† bratislav.misic@mcgill.ca

shed light on brain-disorder and disorder-disorder associations that were latent in the scientific literature, as indicated by lesion- and stimulation-based maps of human cognitive circuits (6, 48–50).

Crucially, using AI experts to synthesise the neuroscientific literature unlocks the possibility of extending meta-analysis into the domain of other model organisms, which we showcase here with macaque (51–53) and mouse (54–57). We make this resource openly available for the neuroscience community at <https://github.com/Hana-Ali/neuroLLM>. Altogether, LLM experts can fundamentally catalyze neuroscientific discovery by automating data aggregation from heterogeneous studies, overcoming the barrier of terminological variations across research domains, and identifying emergent patterns and novel connections across disparate subfields, species, and modalities.

RESULTS

Our openly available framework uses AI experts (large language models) to systematically quantify relative associations between brain regions and cognitive functions. The advantage of using LLMs is that their linguistic competence and extensive neuroscientific knowledge-base enables them to recognise different names for the same brain region, much like a human expert would. This enables automated, systematic mapping of function to anatomy at scale, without requiring specific coordinates or narrowly focusing only on the human literature, as required by currently used reference solutions for coordinate-based meta-analysis in the neuroscience community (e.g., NeuroSynth, NeuroQuery (12, 13)). Thus, an LLM framework supports multiple species and brain atlases to enable comprehensive comparative analysis. Here we focus on three species: human, macaque (*Macaca mulatta*), and mouse (*Mus musculus*). We benchmark LLM-derived brain maps against multimodal, multi-species databases of brain function, structure, and molecular architecture, and we demonstrate that LLM-derived structure-function mappings provide insights to which current field-standard meta-analytic engines are blind.

Specifically, we access five state-of-the-art LLMs as automated experts, encompassing both commercial and open-source models to ensure robust and diverse analytical perspectives. Commercial models include OpenAI’s GPT-4o-mini and Google’s Gemini 2.0 Flash. Open-source models were accessed using TogetherAI’s direct API integration, and include Mistral-7B-Instruct-v0.3, Meta’s Llama-3.3-70B-Instruct-Turbo, and DeepSeek-R1-Distill-Llama-70B.

Cognitive cartography with AI experts

We begin by querying AI experts for the involvement of each region in specific functions representing broad domains of cognition. For each cognitive function, each LLM is prompted to estimate the probability that a specific brain region is involved in that specific function (see *Methods* for details). Responses are parsed to extract decimal probability values between 0 and 1. By querying each LLM about the same cognitive function systematically for each brain region (here, the 68 regions of the Desikan-Killiany atlas (58)), we assemble quantitative brain maps for specific cognitive domains. We then generate meta-analytic consensus of AI experts by z-scoring each map, and averaging the maps across all LLMs.

As proof of concept we focus on eight fundamental cognitive domains, identified by a recent computational ontology as providing a representative data-driven synthesis of human cognition: emotion, cognitive control, listening, memory, language, vision, reward, and manipulation (59). Remarkably, even though LLMs are only ever asked about individual regions, assembling LLMs’ opinions across regions does in fact produce plausible-looking brain maps of cognition, matching human intuition. To find quantitative support for this intuition, we compare the LLM-generated maps to corresponding meta-analytic maps generated by NeuroSynth (12) and NeuroQuery (13). All but one (“manipulation”) are significantly spatially correlated between the LLMs and NeuroSynth (Fig. 1). Notably, the “manipulation” LLM map has the greatest functional association in primary motor cortex, whereas this is not the case for NeuroSynth.

Fig. 1 showcases the breadth of inquiry made possible by LLMs, in the sense that their expertise extends to other model organisms. We therefore use the same procedure to generate cortical maps for the same eight cognitive domains in macaque and mouse, two widely used mammalian model organisms in neuroscience (62–64) (Fig. 1). Prima facie, the species-specific maps generated by consensus of AI experts recapitulate features of core functional neuroanatomy. For instance, in the macaque “vision” is associated with visual cortex, “listening” with superior temporal cortex and temporo-polar cortex (65), and “manipulation” with primary somatosensory and motor cortices, “reward” with orbitofrontal cortex, and “cognitive control” with dorsolateral prefrontal cortex. In mouse, the greatest associations with “vision” are likewise found in visual areas: VISal, VISam, VISl, VISp, VISpl. For “reward”, orbital, anterior cingulate (ACAd, ACAv), and limbic areas (PL, ILA) are prominent, and for “manipulation”, primary and secondary motor areas, as well as primary somatosensory areas. For “listening”, the auditory areas are most involved (AUDd, AUDp, AUDpo, AUDv). It is also noteworthy that the LLMs attribute non-zero associations with “language” to regions of the macaque and mouse brain. However, these are arguably best understood as

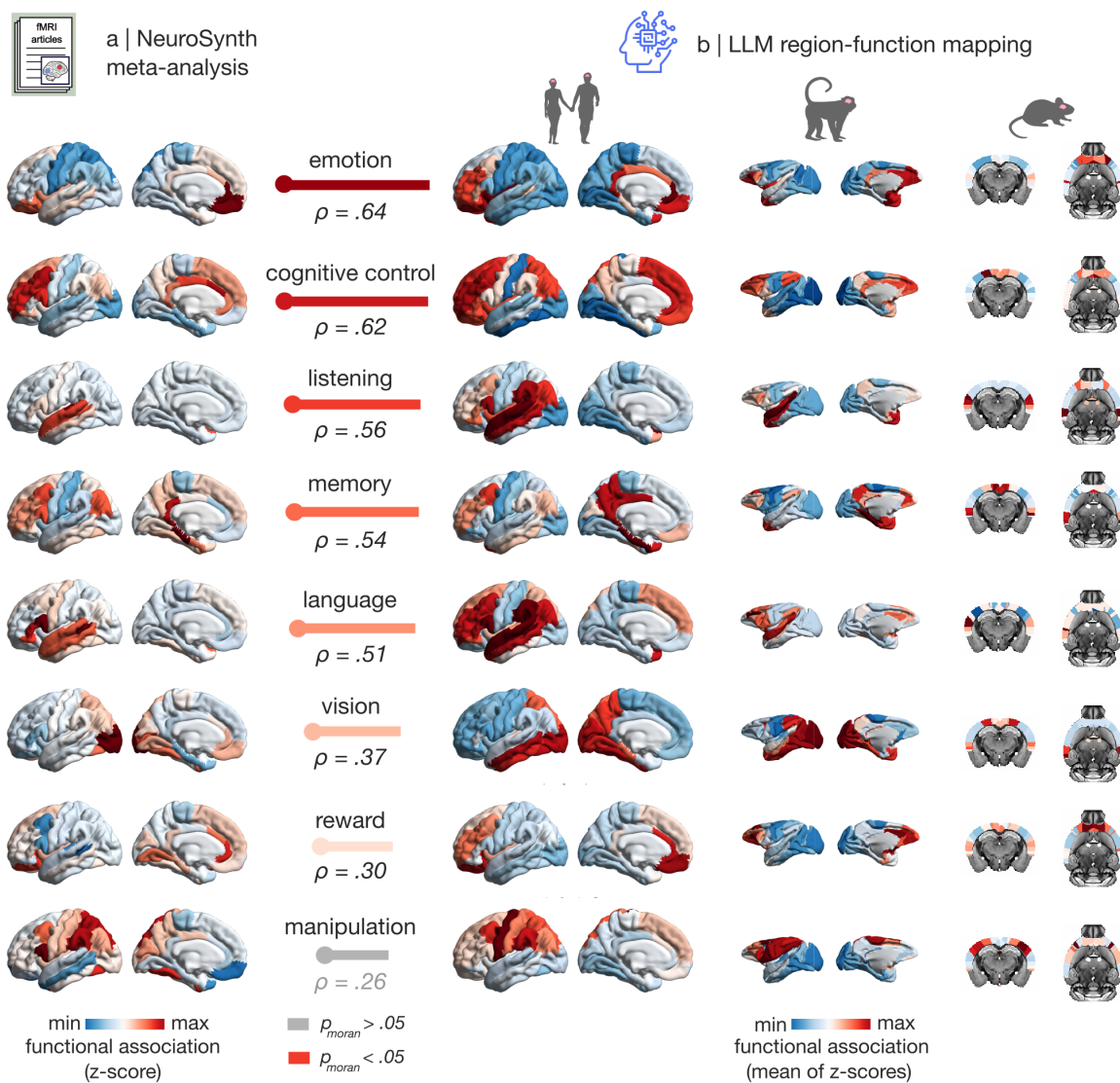


Figure 1. LLMs generate brain maps of cognition | (a) NeuroSynth meta-analytic maps for 8 representative domains of cognition identified through data-driven computational ontology (59). (b) LLMs can generate plausible brain maps of cognition for human, macaque and mouse. Human: 68 cortical regions from the Desikan-Killiany atlas (58). Macaque: 82 cortical regions from the Regional Mapping atlas (60). Mouse: 72 cortical regions from the Allen common coordinate framework atlas (61). Brain maps are z-scored prior to averaging across LLMs. Significance of correlations is assessed against a population of null maps with preserved spatial autocorrelation generated using Moran spectral randomisation.

‘informed best guesses’. Firstly, if we consider the magnitude of the LLMs’ quantitative estimate of regional association with “language”, it is significantly lower than for any other function, as we should expect (Fig. S2). Second, if we consider the spatial distribution of relative association with “language”, the resulting maps are not random or trivial: in both species, the regions that are relatively more associated with “language” display clear spatial overlap with the maps for “listening”, just as in the human (Fig. 1). In addition, these regions are in fact associated with species-specific vocalizations, including putative voice regions in superior temporal cortex and ventral-prefrontal and opercular cortex of the macaque

brain that preferentially respond to species-specific vocalisations (65–67). In other words, the AI experts appropriately convey that finding language-associated regions in these non-human organisms is unlikely, while also providing sensible extrapolations. In the next section, we assess the accuracy of the non-human maps in a more quantitative way.

Cognitive cartography in non-human species

An important use of human meta-analytic databases such as NeuroSynth is to perform principled ‘cogni-

tive decoding’: comparing a new brain map of interest against a pre-specified set of meta-analytic maps, to draw quantitative inferences about the underlying cognitive operations (68–71). Thanks to their advanced language-processing capacity, LLMs unlock the possibility of extending meta-analysis and cognitive decoding into the domain of other organisms. AI experts provide a twofold advantage: (i) they have access to a broader multimodal literature (e.g., not just human fMRI and PET); and (ii) they are not constrained by how studies report their results (e.g. with respect to a standardized coordinate system). Here we consider the capacity for AI experts to decode cortical maps for two fundamental model organisms for neuroscience: the macaque and the mouse.

For the mouse, we benchmark the generated maps against a recently-released flagship resource from the International Brain Laboratory consortium (9). Briefly, the dataset consists of cortical maps reconstructed using in vivo electrophysiological recordings of 621,733 neurons recorded during a decision-making task with 699 Neuropixel probes in 139 mice (9). Importantly, the data were annotated with task components related to sensation, choice, action and reward, as well as internal cognitive states. Specifically, we use brain maps that report for each cortical region, the success (R^2) of decoding. We then perform ‘cognitive decoding’ of each IBL map of task annotations, by spatially correlating it against the 8 representative cognitive maps for the mouse generated using consensus of AI experts. As Fig. 2a shows, we find that the meta-analytic maps display the greatest spatial correlation with the expected Neuropixel maps, such that “vision” is maximally correlated with the “stimulus” annotation, “manipulation” is maximally correlated with “choice” (behaviourally manifested by the animal using its paws to turn a wheel) as well as the “velocity” and “speed” of the wheel-turning, and “listening” is maximally correlated with the “feedback” annotation (feedback was provided as an aversive auditory tone, in case of incorrect choice)(Fig. 2a). Collectively, these results show that LLM maps can generate empirically realistic proxy maps, across domains of cognition and even across species, by synthesising findings across the literature even when explicit meta-analytic results are not available.

For the macaque, there exists no comparable comprehensive dataset. The closest available is Leavitt’s (72) quantitative summary of the evidence of involvement of every macaque region in working memory (number of studies finding evidence in favour, or null, for each region). After resampling this map to the Regional Mapping atlas used for the macaque (see *Methods*), we show that it exhibits significant spatial correlation with the map of macaque working memory obtained from AI experts (Fig. 2b). The phylogenetic proximity of the macaque to human allows us to ask an additional methodological question: is this map macaque-specific, or is it simply knowledge from the human literature projected to the macaque cortex? To address this

question, we obtain a meta-analytic map of human working memory from NeuroSynth and parcellate it into the macaque Regional Mapping atlas (53, 60, 73), comparing it against (1) the LLM-derived map of macaque working memory; and (2) the summary of macaque working memory from Leavitt (72). We find that the macaque working memory map from Leavitt (which we treat as ground truth) is best matched by the LLM-derived map ($\rho = 0.62$), outperforming the human NeuroSynth map ($\rho = 0.51$). Notably, the LLM map remains significantly correlated with the map of macaque working memory from Leavitt, even after partialling out the map of human working memory from NeuroSynth ($r = 0.46, p < 0.001$). The same results are replicated upon using NeuroQuery instead of NeuroSynth (Leavitt-NeuroQuery correlation: $\rho = 0.39$; LLM-Leavitt correlation after partialling out the NeuroQuery map: $r = 0.53, p < 0.001$). In other words, the LLM summary is species-specific, and it goes beyond meta-analyses of the human neuroimaging literature for representing the available evidence about macaque working memory. Collectively, application of AI experts to mouse and macaque literature is encouraging, demonstrating the added value of an LLM-powered approach for automated data integration that complements empirical databases.

LLMs’ neuroscientific knowledge recapitulates functional similarity and structure-function relationships

So far, we used LLMs to map functions to brain areas, but this approach does not take into account how the network structure of the brain supports coordinated and distributed interactions (8, 47, 74, 75). To quantify similarity of regions’ functional profiles in an unbiased fashion, we ask AI experts to list the top N functions of each brain region (e.g. top 5, or 10). The resulting string of text is turned into a text embedding using the OpenAI API (effectively, the reverse of the process by which LLMs generate human-readable text as output). The embedding is a high-dimensional vector representation, such that text strings with more similar meaning will have closer positions in the high-dimensional space of the linguistic embedding. Therefore we can use a measure of vector similarity in embedding space (e.g., the field standard of cosine similarity) to quantify the similarity between regions’ functional profiles, as determined by AI experts. For each pair of regions, we generate a region $\times$ region matrix of functional similarity. This approach provides a quantitative way to ask whether, according to the literature, two regions perform similar functions. Fig. 3a shows this LLM functional similarity matrix. The matrix clearly encodes non-trivial network structure, which we explore next.

Similarity of function can arise from many forms of interaction between brain regions that can be quantified across multiple spatial and temporal scales using imaging, recording and stimulation (76–78). These in-

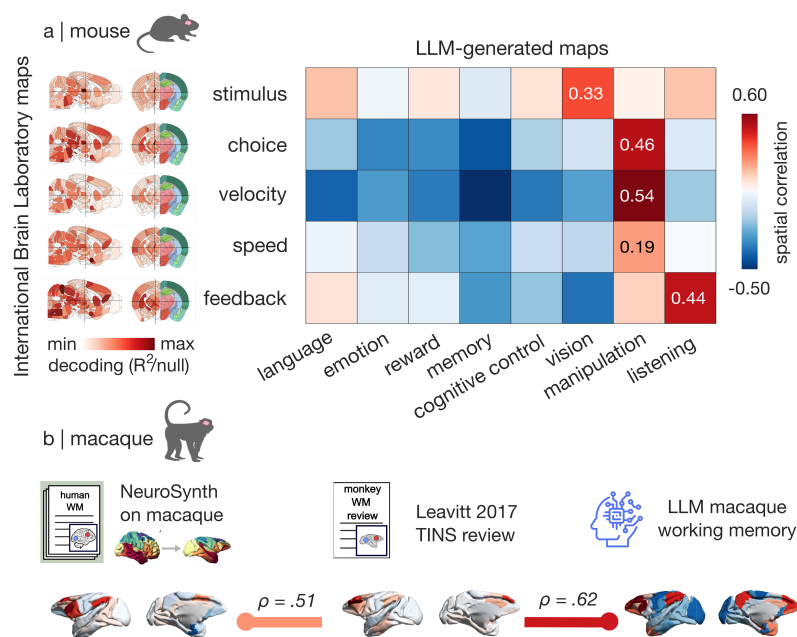


Figure 2. LLM decoding of mouse and macaque cortical maps | (a) Cortical maps of mouse cognition are obtained from recently released Neuropixels recordings from the International Brain Laboratory (9), corresponding to decoding success for five separate components of a decision-making task: stimulus encoding; behavioural choice; velocity and speed of behaviour (turning of a wheel); and feedback (reward or aversive tone). Using LLM-generated maps for the mouse brain pertaining to the same 8 representative domains of cognition as in Fig. 1, we find that the best correlation with empirical Neuropixels maps are appropriate for each task stage. **(b)** LLM meta-analysis of macaque working memory recapitulates available evidence about macaque working memory from Leavitt’s review (72), better than a NeuroSynth meta-analytic map of human working memory projected onto macaque brain.

clude (i) similarity of fMRI activity (haemodynamic co-fluctuation, also known as functional connectivity) (79), (ii) similarity of MEG activity (electrophysiological co-fluctuation) (80–83), (iii) similarity of FDG-PET activity (metabolic co-fluctuation) (84), and (iv) similarity of subjective experience upon electrical stimulation (electrical co-elicitation) (48)(Fig. 3b). The underlying principle is that if two regions activate and de-activate synchronously, it may be because they perform similar roles in the functioning of the brain (85). Ultimately, whether two brain regions perform similar or different function must arise from anatomy: their respective molecular and cytoarchitectonic make-up, and how they are physically connected with each other and with the rest of the brain (8, 76). We therefore also consider networks of (v) receptor co-expression (correlation across 19 in vivo PET maps) (47, 86); (vi) gene co-expression (correlation of transcriptomic fingerprints) (47, 87, 88); (vii) structural connectivity from diffusion MRI tractography (89); and (viii) laminar similarity (similarity of histologically derived cellular distributions across the cortical layers) (47, 90) (Fig. 3b). For completeness, we also consider (ix) the Euclidean distance between regions as an additional predictor, because the brain’s spatial embedding means that nearby regions may be more likely to perform similar functions than far-away ones (76).

To what extent does the LLMs’ quantification of functional similarity reflect other multimodal forms of interaction between brain regions? To address this question we use dominance analysis (see *Methods*; (91)), which quantifies the relative contribution of each network of inter-regional interaction for predicting the automated expert consensus of functional similarity. We repeat this analysis in macaque and mouse, using fMRI co-fluctuation (56, 63), structural connectivity (54, 92), gene co-expression (52, 53, 55), Euclidean distance, and receptor co-expression (macaque only; (51)) (Fig. S3).

Consistently across all three species, we find that the fMRI-derived networks of haemodynamic co-fluctuation display the highest relative importance for predicting the automated expert consensus of functional similarity, outperforming the anatomical and molecular ones, and in human, the electrophysiological and metabolic co-fluctuation; the only exception is the similarity of subjective effects elicited by direct intracranial stimulation, which is the best predictor of functional similarity in humans. Thus, using LLM-based automated summary of the literature allows us to extend the results of Smith et al. (85), Crossley et al. (93), Toro et al. (94) to macaque and mouse, revealing that synchronous activity predicts similarity of function. We also demonstrate—in all three species—that haemodynamic co-fluctuation out-

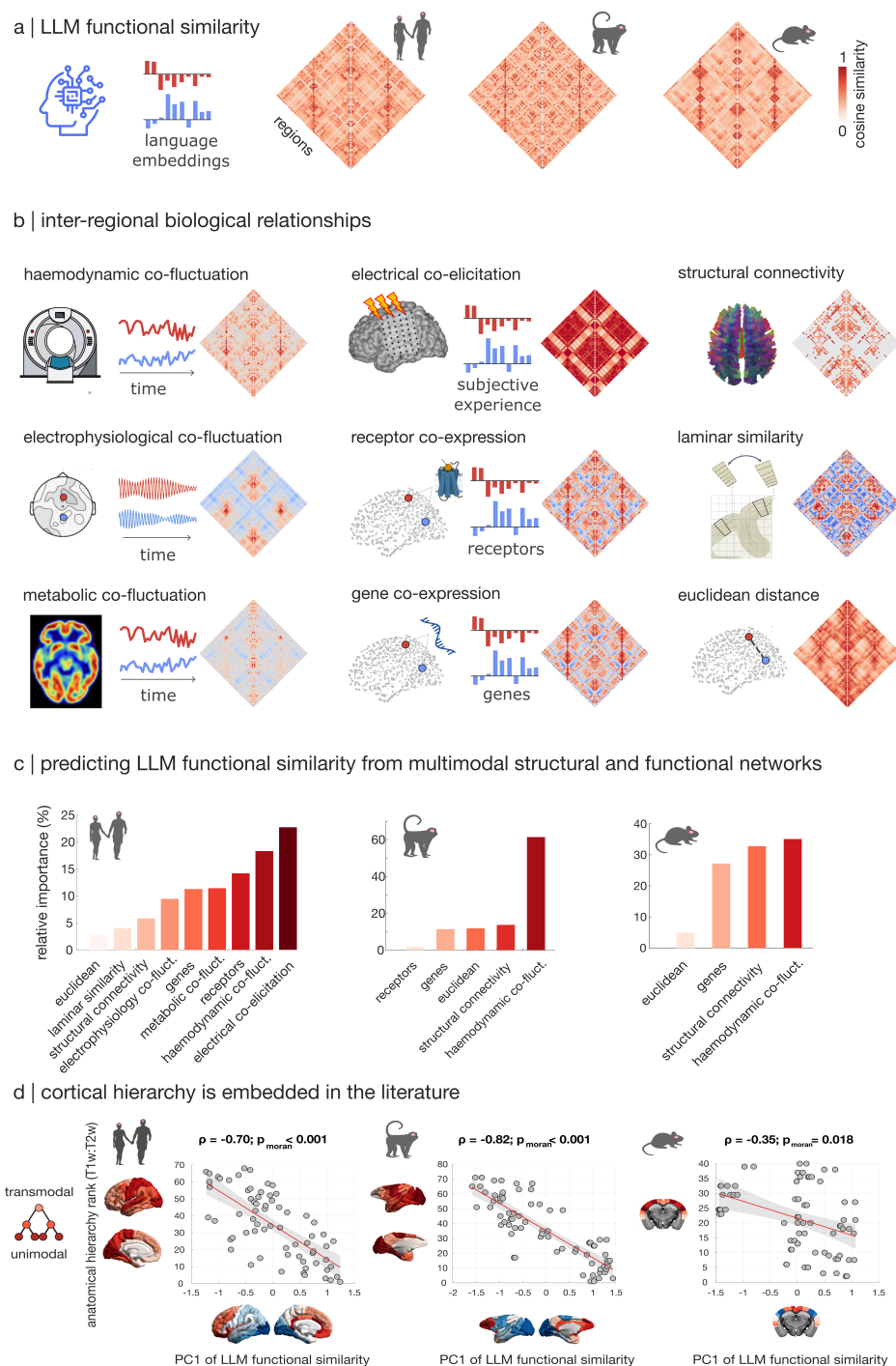


Figure 3. LLMs' quantification of functional similarity recapitulates structural and functional interactions across species
| (a) We obtain functional similarity from LLMs as the inter-regional cosine similarity of language embeddings when asked to describe the function of each region. **(b)** Multimodal quantifications of inter-regional interactions for the human brain. Functional interactions can be quantified as co-fluctuation (correlation over time) of spontaneous activity from fMRI haemodynamics ('functional connectivity'), or electrophysiology from MEG, or metabolic activity from dynamic FDG-PET; or as similarity of subjective experiences upon intracranial electrical stimulation. Anatomical relationships can be quantified as similar patterns of gene and receptor expression, similar cytoarchitecture across layers, presence of direct white matter tracts between regions, or physical proximity in space. See Fig. S3 for corresponding networks available in macaque and mouse. **(c)** Dominance analysis reveals that functional interactions best account for LLM-derived functional similarity across species. Dominance analysis distributes the fit (variance explained) of a multiple regression model across predictors, such that the contribution of each predictor can be assessed and compared to other predictors, producing a relative importance score that reflects the proportion of the variance jointly explained by all predictors, that can be attributed to each predictor. **(d)** We estimate the hierarchical position of each region using the T1w:T2w ratio, an in vivo MRI proxy of intracortical myelin (lower intracortical myelin indicates higher position on the hierarchy). In each species, we find a significant correlation with the PC1 of LLM-derived functional similarity (note that sign of the PC1 is arbitrary).

performs both anatomical and transcriptomic similarity for predicting similarity of function across the literature. Altogether, automated expert consensus of regions' functional similarity from the literature is not simply redundant with meta-analytic similarity, but rather provides a valuable complement to the growing compendium of perspectives on inter-regional interactions.

AI experts discover the latent hierarchical organisation of mammalian cortex

So far, we found that LLMs can recover biologically realistic brain topographies and networks. Does their internal representation also encode other, more complex features of mammalian brain organization that are latent in the literature? One such prominent feature is the hierarchical organization of the mammalian cortex, from unimodal cortices at one end, to transmodal association cortices at the apex (3, 53, 57, 95, 96).

To address this question, we estimate the hierarchical position of individual regions using the T1w:T2w ratio, an in vivo MRI proxy of intracortical myelin (3, 57, 95, 96). We then spatially correlate this empirical estimate of anatomical hierarchy with the principal axis of variation in the LLM functional similarity matrix. In each of the human, macaque, and mouse, we find significant spatial correlations between the empirical anatomical hierarchy and the putative LLM hierarchy (Fig. 3d), suggesting that the LLM experts can discover latent features of brain organization in each species. This is particularly noteworthy because the LLMs are never asked about “hierarchy”, and are queried about one region at a time, yet a distributed hierarchical organisation spontaneously emerges. Altogether, use of LLMs to perform automated structure-function mapping opens the path to comparative discoveries across species and across modalities.

From meta-analytic associations to direct perturbation

Do LLM-derived maps merely reflect meta-analytic statistical associations between cognitive functions and brain regions, or can they draw on a broader knowledge-base to synthesise the cognitive effects induced by direct manipulations on brain regions? To address this question, we analyze a dataset of direct intracranial electrical stimulation (iES) in surgical patients ($N = 1,537$ cortical sites in 67 participants) (48). Briefly, participants self-report their subjective experience upon direct electrical stimulation at different cortical locations. Integration of the first-person reports yields probabilistic maps of how direct stimulation elicits broad categories of cognitive processes, including ‘emotion’, ‘language’, ‘memory’, and ‘vision’, which we can use to benchmark the corresponding LLM-derived maps.

Fig. 4b shows the iES-derived maps, alongside their spatial correlations with LLM-derived maps (Fig. 4c) and

NeuroSynth meta-analytic maps (Fig. 4a). Two broad observations emerge. First, LLM maps are positively and significantly correlated with iES maps (with the exception of “vision”, which is not significant but exhibits the greatest correlation, $r = 0.51$). Second, the correlations between iES maps and LLM-generated maps are all greater in magnitude than the correlations between iES maps and NeuroSynth maps. In other words, the LLM-derived functional brain maps do not just reflect neuroimaging meta-analysis, but go beyond correlation to embody additional insights from the literature on lesions and stimulation, about how individual regions shape ongoing cognition.

Using AI experts to synthesise maps of brain disease

Up to this point, we used AI experts to generate maps of healthy brain function. But knowledge aggregation also provides an opportunity for discovery about dysfunction. In psychiatry, intense effort has been directed to understand the extent to which different brain regions are involved in specific syndromes, using numerous imaging, recording and stimulation methods. While individual techniques are increasingly powerful, psychiatric syndromes are increasingly conceptualized as multifaceted diseases that are expressed at multiple levels of brain organization (98). Here we explore whether AI experts, taking a broader view based on the entire neuroscience literature, may generate biologically insightful disease-specific topographic maps. We systematically query LLMs about each region's involvement in five psychiatric diseases: autism, bipolar disorder, attention deficit hyperactivity disorder (ADHD), major depressive disorder, and schizophrenia. As before, the approach generates disorder-specific consensus maps of regional disease involvement (Fig. 5a). At first blush, the maps recapitulate well known hot spots; for example, the involvement of subgenual anterior cingulate in depression (49, 99, 100), or the involvement of dorsolateral prefrontal cortex in schizophrenia (101).

To contextualize the potential utility of these maps, we contrast them with two other sets of maps, one approximating a “gold standard” and the other approximating the current “field standard”. First, for each disease, we derive regional density of expression of GWAS-derived risk genes (Fig. 5a, middle column; see *Methods*). Transcriptomic patterns of genes associated with risk for a disease thus provide a molecular signature of the disease (98). Second, for each disease we derive regional maps of sMRI-estimated cortical thinning from the global ENIGMA consortium (Fig. 5a, left column) (14, 17, 18). We find that 3/5 of LLM maps (autism, bipolar, and ADHD) exhibit significant spatial overlap with maps of risk-gene expression. In contrast, no significant correlation is observed between the maps of risk-gene expression density, and the corresponding ENIGMA maps of cortical morphometry. In addition, the magnitude of

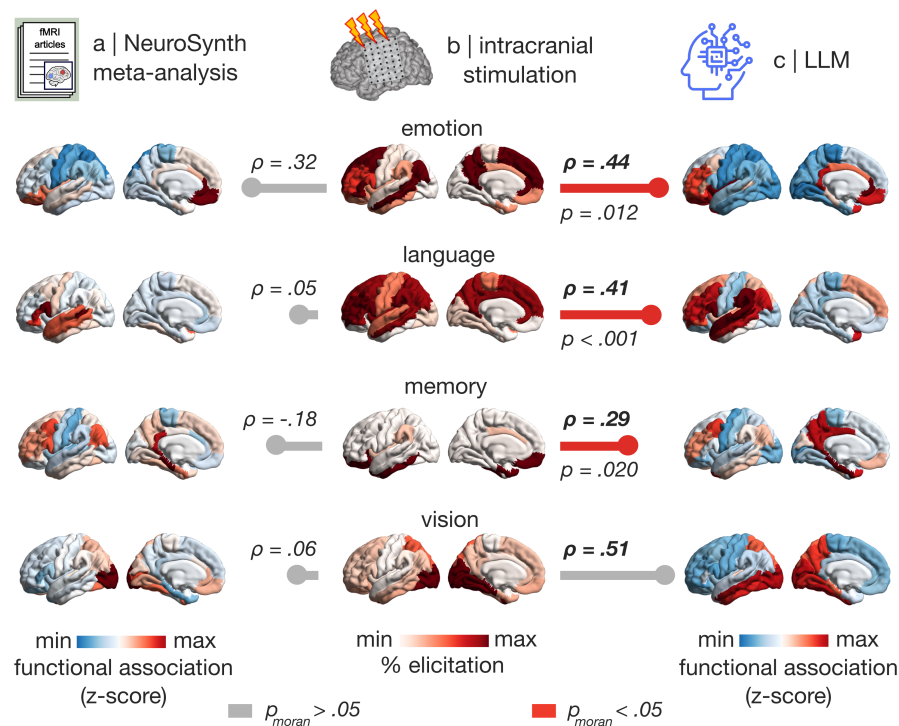


Figure 4. **LLM meta-analysis recapitulates circuits obtained from direct intracranial stimulation** | (a) NeuroSynth meta-analytic maps. (b) Prevalence of subjective experience pertaining to different domains of cognition, elicited by direct intracranial electrical stimulation (48) (c) LLM-generated maps of functional association. Significance of correlations is assessed against a population of null maps with preserved spatial autocorrelation generated using Moran spectral randomisation (97).

correlation between risk-gene expression maps is greater with all LLM generated maps compared with ENIGMA maps. In other words, automated literature summary about regional involvement with specific neuropsychiatric disorders is more sensitive to regional genetic risk than MRI-focused anatomical abnormality.

Taking this idea further, we test the capacity of LLMs to recapitulate shared relationships and comorbidities among brain diseases. We first compile risk genes across multiple neurological, psychiatric, neurodevelopmental, sleep and tumor-related syndromes, and generate brain-wide transcriptomic maps for each (98). Genes associated with risk for brain disease exhibit characteristic expression patterns. Fig. 5c shows the patterns of similarity (correlations) among these diseases based on their cortical risk gene expression. Fig. 5b instead shows the patterns of similarity among LLM-generated brain maps for the same diseases. Both heatmaps are organized according to the International Classification of Diseases (ICD-11), a global standard diagnostic manual published by the World Health Organization (102). To ask which of the two matrices (risk genes or LLM) better captures ICD classifications, we compute the modularity of each with respect to ICD categories. The LLM matrix is significantly more modular compared to the risk gene matrix (modularity $\Delta = 0.06, p < 0.001$; Fig. S4), suggesting that LLMs—even when only asked about a sin-

gle region and a single disease at a time—find that diseases from the same family display similar cortical topography. Interestingly, AI experts also identify similar functional involvement between diseases that are not part of the same category but are nevertheless related, such as for dementia and delirium with many psychiatric disorders including schizophrenia (originally termed 'dementia precox' (103)). Likewise, a relationship between sleep disturbances and psychiatric disorders is increasingly recognised (104, 105), as is a link between epilepsy and dementia (106, 107): both are correctly identified by AI experts, but not by similarity of risk-gene expression (Fig. 5b,c). Crucially, these links between disorders are an emergent property: the LLMs are never queried about more than one disorder at a time, yet similar brain maps are produced for disorders that turn out to be co-morbid. The nosological insight of AI experts may potentially stem from the fact that they are directly queried about cognitive function (rather than genetic risk), which is presumably closer to the symptoms that are used to classify each disease. In Fig. S5 we demonstrate that LLM-generated disorder maps recapitulate an alternative definition of gold standard: lesion network mapping (5, 6, 108). Namely, we use a psychosis circuit, derived from regions whose lesion induces psychosis symptoms (50) ($N = 153$ lesion cases); and a depression circuit, derived from the convergence of re-

gions whose lesion induces depressive symptoms, and whose brain stimulation (using DBS or TMS) alleviates depressive symptoms (49) (14 studies; $N = 461$ lesion cases and $N = 251$ stimulation cases). Both lesion circuits correlate with the LLM-generated map for bipolar disorder, which indeed can involve both depressive and psychotic symptoms (Fig. S5). Collectively, these results demonstrate that AI experts are sensitive to both (more distal) genetic and (more proximal) lesion circuit manifestations of cognitive dysfunction.

Opening new research directions with AI experts

What new investigations are possible with this analytic framework, that were not available before? As a final step, we showcase the capacity of LLMs to take the results of two recent multi-modal integration studies that up to this point were only possible in humans, and extrapolate them to other model organisms (109, 110). We use partial least squares analysis to integrate LLM-generated maps for 17 representative cognitive operations, and expression patterns of 81 brain-related genes (57, 111) that have available orthologs in human (from microarray (87)), macaque (from stereo-seq (52, 53)) and mouse (from in situ hybridization (55)). Across human, macaque, and mouse, we consistently find a statistically significant latent variable of gene-cognition association (Fig. S6), replicating and extending the human results obtained by Hansen et al. (110) using NeuroSynth (Fig. S7). In all three species, the main latent variable of gene-cognition association distinguishes LLM-derived emotion- and memory-related cognitive operations from sensory and motor ones, exhibiting significant associations with the term loadings from the human NeuroSynth analysis (human: Spearman $\rho = 0.66, p = 0.007$; macaque: $\rho = 0.77, p < 0.001$; mouse: $\rho = 0.51, p = 0.045$; Fig. 6). Likewise, all three species exhibit a significant association with the gene loadings obtained in the original analysis of Hansen et al. (110), even though the macaque gene expression is from stereo-seq ($\rho = 0.56, p < 0.001$) and the mouse gene expression is from in situ hybridization ($\rho = 0.52, p < 0.001$) (Fig. 6). Altogether, three separate data-driven analyses in three separate species, each with a different technique for spatial transcriptomics, converge to identify a conserved molecular architecture for cognition in the mammalian brain.

In the Supplement we also extend across species the recent results of Luppi et al. (109). We combine principles of network control theory (112, 113) and *ex vivo* tractography with our automated AI consensus, to reveal that the architecture of macaque and mouse brain networks is wired to favour energetically efficient transitions between cognitive states - just like the human brain (Fig. S8). Altogether, use of LLMs to perform automated structure-function mapping opens the path to consistent discoveries across species and across modalities, beyond

what is possible with the current human-specific meta-analytic tools.

Robustness and sensitivity

To evaluate robustness of our findings, we assess their sensitivity with respect to important methodological choices: choice of meta-analytic framework for comparison; inclusion of subcortex; and variability among LLMs' responses. We replicate our comparison between LLMs and formal neuroimaging meta-analysis using NeuroQuery (13) instead of NeuroSynth. NeuroQuery employs techniques from (pre-LLM) natural language processing to infer semantic relationships among terms used in the literature, thereby combining the results of similar terms (13). Whereas NeuroSynth generates brain maps of the consistency of observations reported in the literature, the brain maps generated by NeuroQuery predict the probability that a study on the cognitive operation of interest will report observations at each location (13). We find similarly good agreement between LLM-generated maps and meta-analytic maps generated by NeuroQuery (Fig. S1).

We further replicate meta-analysis results upon using the more fine-grained Glasser parcellation (114) (Fig. S15 and Fig. S16), and after including subcortical structures: thalamus, caudate, putamen, globus pallidus, accumbens, hippocampus, and amygdala (both left and right), corresponding to the Cammoun atlas (115) (Fig. S17). Notably, the LLM-derived clustering remains more faithful to ICD-11 classification than the clustering based on molecular fingerprints of risk-gene expression, even when including subcortical structures, which are implicated in many neurological, psychiatric, and neurodegenerative disorders (Fig. S18).

Finally, we consider the extent to which AI experts agree and disagree about structure-function mappings, and what that teaches us about the current state of knowledge for specific regions and functions. Our main results are obtained by combining (averaging) the outputs of 5 widely used large language models: OpenAI's GPT-4o-mini, Google's Gemini 2.0 Flash, Mistral-7B-Instruct-v0.3, Meta's Llama-3.3-70B-Instruct-Turbo, and DeepSeek-R1-Distill-Llama-70B. However, the workflow could also be implemented using a different combination of LLMs, or even a single one. Different models are trained on different corpora of data, and also differ in terms of specific architecture (e.g., number of parameters). To what extent do they agree in their summaries of the brain mapping literature? We systematically assess how the different LLMs used here differ in terms of their confidence about individual regions, individual cognitive operations, and global patterns of structure-function mapping. Cases where different LLMs diverge in their assessment of a region's involvement with a particular cognitive function (Fig. S13) may

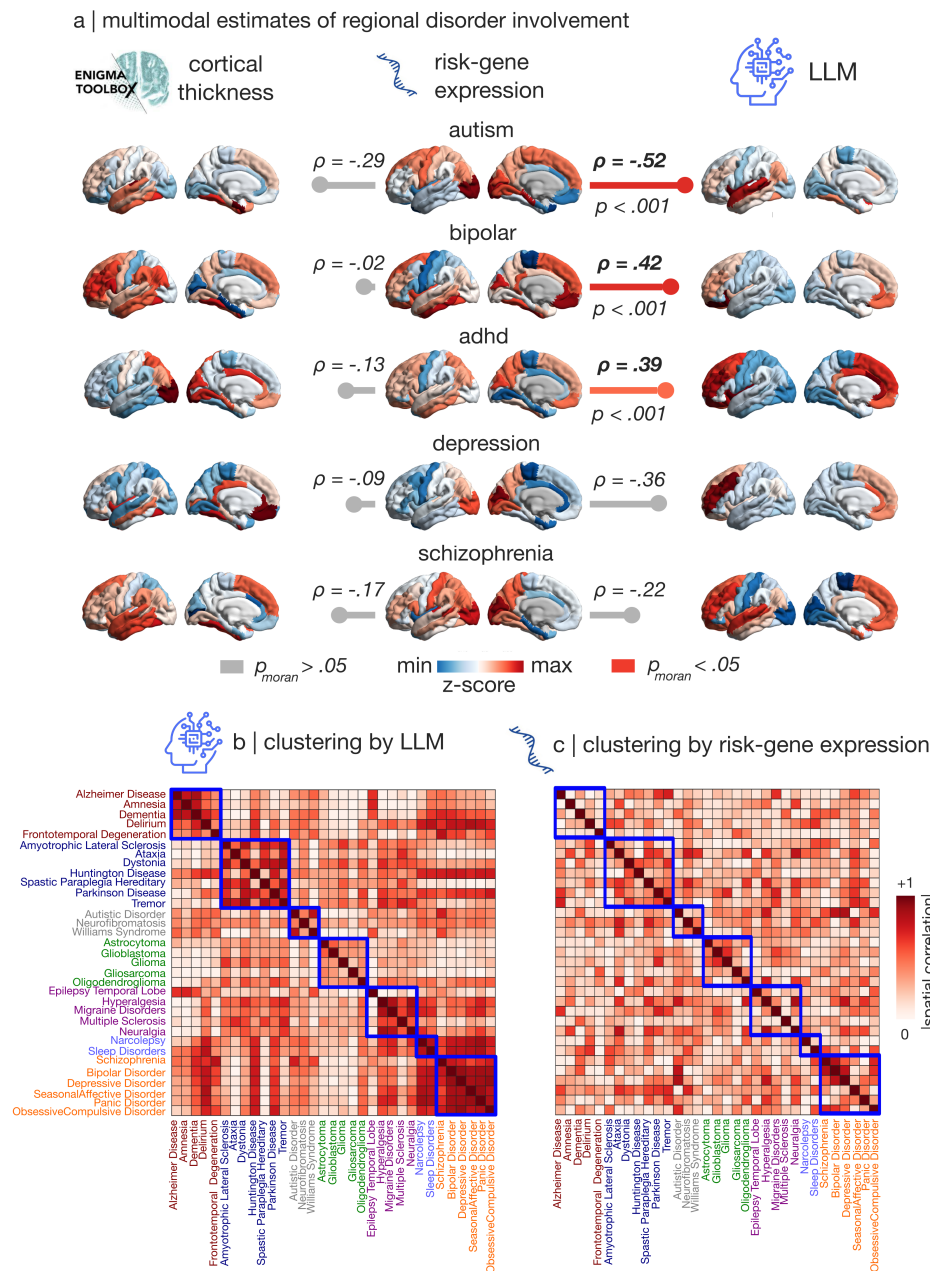


Figure 5. LLMs recapitulate brain circuits of cognitive dysfunction from risk genes and atrophy | (a) Multi-modal regional disorder involvement from ENIGMA meta-analysis of cortical thickness (left); from transcriptomic expression of risk-genes identified by genome-wide association studies (GWAS) (middle); and from consensus of AI experts (right). (b,c) Similarity (magnitude of spatial correlation) between disease maps from LLM-quantified regional involvement (b) aligns with symptom dimensions significantly better than similarity from transcriptomically derived maps of molecular risk (c). Modular structure is based on assignment into 7 disease categories from the International Classification of Diseases manual: Neurocognitive disorders and Dementias; Neurodegenerative and movement disorders; Neurodevelopmental disorders; Brain tumors; Neurological disorders; Sleep and Wake disorders; and Mood and Behaviour disorders. The difference in modular structure between LLM similarity and molecular similarity is significantly greater than chance (Fig. S4).

indicate deficiencies in the literature. Broadly speaking, Mistral tends to be the most confident LLM (on average assigning significantly higher probability of association between regions and functions)(Fig. S10 and Fig. S11a).

The same pattern is observed for macaque and mouse (Fig. S12 and Fig. S11b,c). However, in general we find excellent agreement between the cognitive brain maps generated by different LLMs (range: 0.59 to 0.80,

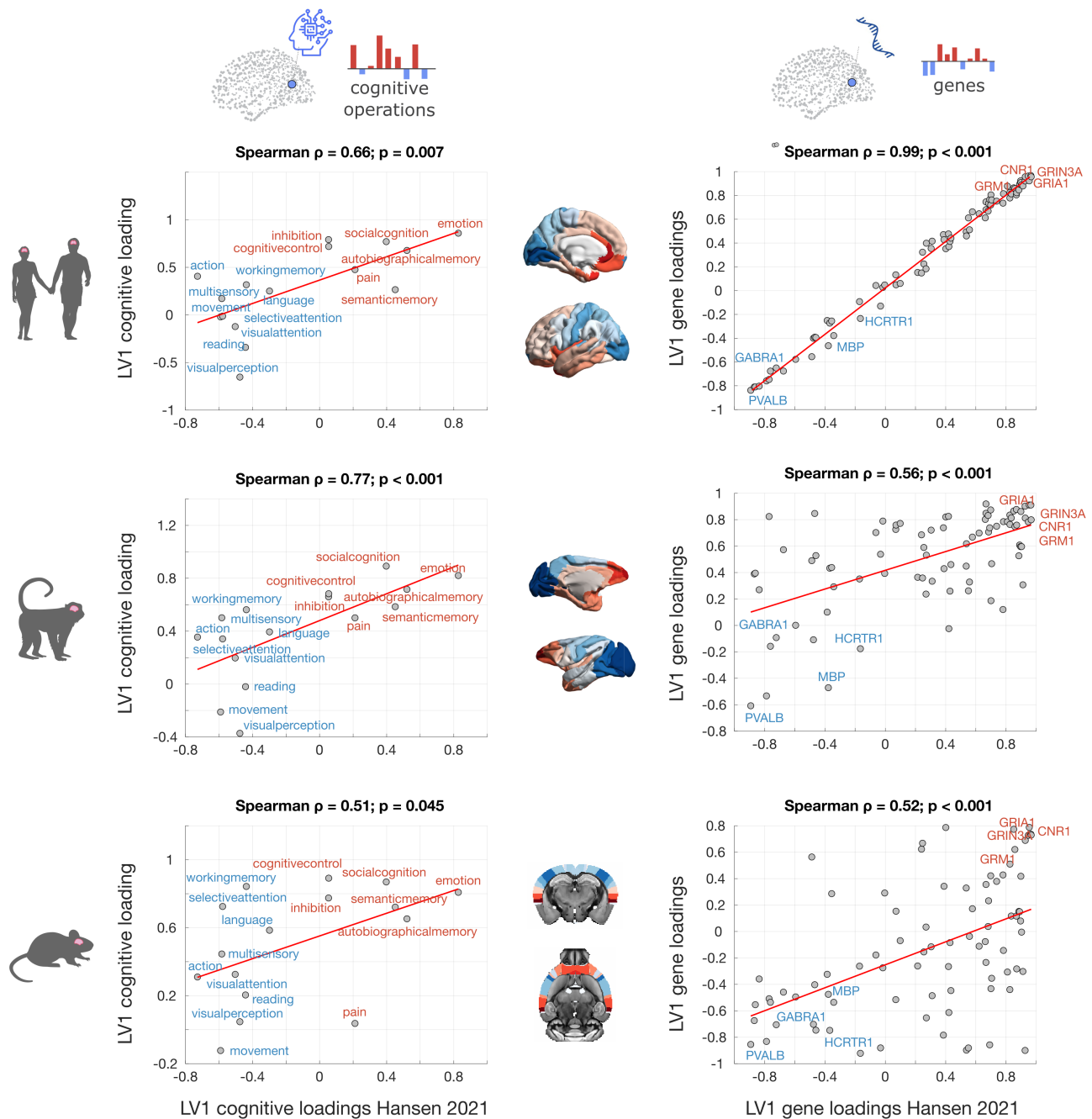


Figure 6. AI-driven discovery of consistent gene-cognition associations across mammalian brains | The LLM-derived maps of brain function provide essentially identical insights as NeuroSynth, about the relationship between human cognition and the human brain's molecular architecture. Despite using only a subset of genes and a subset of terms, and despite using cognitive maps from LLMs rather than from NeuroSynth, we successfully replicate the results of Hansen et al. (110) to a high degree. The main latent variable of gene-cognition association from Partial Least Squares analysis displays the same anterior-posterior brain pattern, with both gene and term loadings onto this latent variable being highly correlated with those in the original study (genes: Spearman $\rho = 0.99$, $p < 0.001$; terms: $\rho = 0.66$, $p = 0.007$). Significance is assessed against a population of null maps with preserved spatial autocorrelation generated using Moran spectral randomisation (97).

median = 0.74) (Fig. S10, Fig. S9). We also find very high spatial correlations (almost all > 0.90) between the maps obtained from re-running the same workflow a second time, indicating that results are highly consistent both within and across LLMs (Fig. S14). Altogether, although here we combined the outputs of multiple models as a well documented strategy to boost LLMs' factual accuracy (116–118), our results indicate that using a subset of LLMs would produce comparable insights.

DISCUSSION

The goal of cognitive neuroscience is to understand how brain architecture gives rise to cognition, behaviour, and dysfunction. The literature is vast and variegated, encompassing both healthy humans and patients, but also more experimentally accessible animal models such as non-human primates and rodents; and spanning a wide range of methodologies, from different neuroimaging modalities, to lesions and brain stimulation. We developed and made openly available a fully automated computational workflow to synthesise the knowledge of the scientific literature embedded in large language models' training corpora.

Automated meta-analysis of AI experts goes beyond traditional coordinate-based meta-analysis, extending its reach from a single imaging modality to the broader literature on direct interventions, lesions, and disease. In parallel, use of AI experts to summarise the literature extends to macaque and mouse analytic approaches that were previously the sole domain of human neuroscience. This approach sheds light on brain-disorder and disorder-disorder associations that were latent in the scientific literature, and revealed common molecular signatures of cognition and affect across multiple species.

The evaluations performed here suggest that widely used LLMs have great value as automated experts for synthesising the functions of specific regions, and the neural circuits underlying specific cognitive functions. Brain maps of specific cognitive functions assembled through consensus of AI experts bear significant spatial overlap with conventional meta-analytic maps, while also going beyond them, capturing maps of direct electrical stimulation or lesion network mapping that are not available from current meta-analytic engines (5, 48, 49). This overwhelming utility of AI experts is consistent with the fact that they can draw on the entire scientific literature and they interpret findings similarly to human experts—though it is worth bearing in mind that the literature itself is not free from bias (13, 119).

Our workflow is publicly available and easy to use, supporting many popular LLMs (both paid and free) to maximise accessibility. We chose to use popular general-purpose LLMs without need for any fine-tuning, to maximise the ease and immediacy of use even for those with minimal knowledge of LLMs. Indeed, LLMs con-

sistently outperform human experts by a wide margin (83% vs 66% correct) across domains of neuroscience even without fine-tuning—providing confidence in our own use of the LLMs as 'automated experts' (45). Although fine-tuning may further improve performance (41, 45), we expect that as LLMs' rapid evolution continues, performance will improve even for general-purpose models (45, 120). Ultimately, changes in cost, speed, and even environmental impact may also need to be factored into one's choice of model(s) to use. More broadly, to extract quantitative data in a systematic fashion we made use of strictly regimented prompts. Our approach therefore does not currently include the opportunity for users to engage in dialogue with the LLMs, challenge them, or pursue clarifying questions. Thus, one way to expand our workflow could be to produce a supporting explanation for the provided output, including references (43, 44)—although such strategies may introduce additional scope for confabulation of plausible-sounding but fictitious evidence. Future work may expand our approach in several directions. Beyond fine-tuning, accuracy and interpretability may be boosted via retrieval-augmented generation (121). Although we already prompt the LLMs to behave as neuroscience experts and only rely on peer-reviewed articles, textbooks, and reputable sources in their simulated review of the literature, alternative prompt engineering strategies may also be pursued (44, 122–124).

The approach that we adopted combines the outputs of several popular large language models, in part to overcome differences in training corpora and algorithms, and in part to overcome the inherent stochasticity of LLM outputs (125). Output aggregation is a known performance-boosting strategy, both for human experts and for models (34, 116–118, 126, 127). Although we found broad agreement between the cognitive maps generated by different AI experts, jointly considering multiple LLMs' outputs enables us to also obtain a quantitative estimate of the literature's uncertainty about each region-function association (Fig. S13). Resolving such uncertainties can then become the target of future experimental work, highlighting the potential of AI experts to drive hypothesis-generation.

This workflow differs from a classic statistical meta-analysis in the traditional sense of directly aggregating effect sizes from specific individual studies to obtain an overall effect size (21). Rather, the automated consensus of AI experts introduced here offers a quantitative overview of the literature, akin to what may be obtained from aggregating human experts' opinions (e.g., (128)). Our results suggest that this approach is useful in the sense that AI experts—even though only prompted about specific functions in specific brain regions—can recover additional organizational features, such as inter-regional covariance patterns and hierarchical differentiation. Like other forms of literature aggregation, this approach may be susceptible to the literature's well documented biases, such as confirmation bias and publica-

tion bias (12, 13, 119, 129). Unlike coordinate-based meta-analyses, but more akin to human experts, LLMs supplement evidence with (educated) guesswork and extrapolation. A clear example is language: for macaque and mouse, LLMs consistently provided maps that are analogous to the respective species' maps for 'listening'. Crucially, cases when LLMs' answers consistently diverge from human expectation—or from modality-specific meta-analyses—may offer valuable clues about latent patterns and missing links in the literature, providing the opportunity for automated hypothesis-generation and highlighting follow-up experiments and analyses (119, 130).

At a time of exciting technological, analytical and data sharing advances (1, 63, 131, 132), and an exponentially expanding literature, there is an innate sense of loss and even sadness as one realizes that it is simply impossible to keep abreast of all new knowledge about the brain. The present work showcases how the use of widely accessible AI technologies can synthesize knowledge about the brain in a way that is readily useful for generating quantitative insight and asking fundamentally new questions about brain function.

METHODS

Automated consensus of AI experts using Large Language Models

We introduce a computational framework for automated meta-analysis of AI experts, synthesising knowledge about brain region functions across species using Large Language Models (LLMs). This framework implements two primary analysis modes: (1) hypothesis-driven functional association, to quantify the likelihood of a specific cognitive function (or dysfunction) being associated with a particular brain region; and (2) data-driven function listing, to identify the top functional associations for different brain regions.

Our framework supports multiple species and brain atlases to enable comprehensive comparative analysis. Our analysis is primarily conducted across three species: human, macaque (*Macaca mulatta*), and mouse (*Mus musculus*), using the Desikan-Killiany (58), Regional Mapping (60), and Allen Mouse Brain Atlas (61), respectively. Brain regions are analysed separately per hemisphere (though the option to perform joint analysis across hemispheres is available).

Large Language Models

Five state-of-the-art LLMs are employed in this study, encompassing both commercial and open-source models to ensure robust and diverse analytical perspectives. The commercial models include OpenAI's GPT-4o-mini, and Google's Gemini 2.0 Flash.

Open-source models accessed through TogetherAI include Mistral AI's Mistral-7B-Instruct-v0.3, Meta's Llama-3.3-70B-Instruct-Turbo, and DeepSeek's DeepSeek-R1-Distill-Llama-70B. All models are accessed via direct API integration with OpenAI, Google AI, and TogetherAI platforms. Default model parameters are used for all models. All models implement exponential backoff retry logic to handle rate limiting and temporary API failures.

Analysis is performed on a per-region basis: the LLM is queried about a specific region using its anatomical name as per the corresponding atlas. LLMs' extensive neuroscientific knowledge-base enables them to recognise different names for the same brain region, much like a human expert would—thereby enabling systematic mapping of function to anatomy like coordinate-based meta-analytic engines, while overcoming the need for exact coordinates that restricts them to human MRI.

Functional association

For predetermined cognitive functions, LLMs are prompted to estimate the probability that each brain region is involved in that specific function:

Guidelines 1. Consider only **species-specific** neuroscience literature. Do **not** use data from other species. 2. The probability should be a single decimal number between 0 and 1. 3. This number should approximate the relative frequency with which this function is linked to the given brain region in literature. 4. **DO NOT** provide explanations, citations, or any extra text— **return only the probability value.**

Expected Output Format 0.XX” “You are an expert in neuroscience literature analysis. Your task is to estimate the probability that the brain region **region** hemisphere part **species** brain is involved in **function**. These functions should be based on a simulated review of neuroscience literature, reflecting how frequently these functions are associated with this brain region across **peer-reviewed** studies, textbooks, and reputable sources.

Guidelines 1. Consider only **species-specific** neuroscience literature. Do **not** use data from other species. 2. The probability should be a single decimal number between 0 and 1. 3. This number should approximate the relative frequency with which this function is linked to the given brain region in literature. 4. **DO NOT** provide explanations, citations, or any extra text— **return only the probability value.**

Expected Output Format 0.XX”

Responses are parsed to extract decimal probability values between 0 and 1, with text parsing algorithms designed to handle various response formats and convert non-numeric responses to standardised probability scores.

Embedding of top functions

For each brain region, LLMs are queried to identify the top N primary functions based on neuroscience literature (here, $N = 5$). The standardised prompt template instructs models to:

"List out the top 5 functions that region [X] in the [hemisphere] hemisphere of the [species] brain is involved in. These functions should be based on a simulated review of neuroscience literature, reflecting how frequently these functions are associated with this brain region across peer-reviewed studies, textbooks, and reputable sources."

Raw LLM responses undergo systematic cleaning using regular expressions to extract function lists, remove extraneous text, and standardise formatting. Cleaned responses of the top functions are then concatenated into a single text string, which is subsequently converted into semantic embeddings using OpenAI’s text-embedding-3-large model (3,073 dimensions). These embeddings serve as high-dimensional representations of each region’s functional profile according to each LLM.

After this, inter-regional similarity is computed using cosine similarity on the semantic embeddings derived from the above queries. For each species and LLM, pairwise cosine similarity matrices are generated to quantify functional similarity between all brain regions. The cosine similarity measure is chosen for its effectiveness in high-dimensional semantic spaces and its interpretability as a measure of functional overlap between regions. Similarity matrices preserve the original atlas ordering of brain regions to maintain neuroanatomical consistency and enable systematic comparison across models. Missing or failed embeddings are handled through exclusion from similarity calculations, with appropriate documentation of data completeness.

Meta-analytic cognitive brain maps from NeuroSynth and NeuroQuery

Continuous measures of the association between voxels in MNI-152 standard space and cognitive categories were obtained from NeuroSynth, an automated term-based meta-analytic tool that synthesizes results from more than 14000 published human fMRI studies by searching for high-frequency key words (such as “pain” and “attention” terms) that are systematically mentioned in the papers alongside fMRI voxel coordinates ([https://](https://github.com/neurosynth/neurosynth)

github.com/neurosynth/neurosynth, using the volumetric association test maps (12)). This measure of association strength is the tendency that a given term is reported in the functional neuroimaging study if there is activation observed at a given voxel. Note that NeuroSynth does not distinguish between areas that are activated or deactivated in relation to the term of interest, nor the degree of activation, only that certain brain areas are frequently reported in conjunction with certain words. The probabilistic measure reported by NeuroSynth can be interpreted as a quantitative representation of how regional fluctuations in activity are related to psychological processes.

For our main analyses, we used NeuroSynth maps pertaining to eight fundamental domains of cognition from a recent data-driven ontology: emotion, cognitive control, listening, memory, language, vision, reward, and manipulation (59). Voxelwise maps in MNI-152 standard space were parcellated into 68 cortical regions from the Desikan-Killiany atlas (58). An additional 14 subcortical regions were also included for the sensitivity analyses, corresponding to the Cammoun atlas (115). We also obtained the NeuroSynth map for ‘working memory’ and parcellated it according to the Regional Mapping atlas (60), adapted from the macaque to the human brain by Luppi et al. (53), Bezgin et al. (73). For the multivariate analysis linking cognition and genes, we used a set of 17 cognitive terms, obtained as the intersection of two term lists: the original list of 123 NeuroSynth terms from (110), and a list of 24 NeuroSynth terms from (68).

For replication, we also obtained brain maps for the same eight domains of cognition from NeuroQuery (<https://neuroquery.org/>) (13). NeuroQuery contains all the terms from the neuroscientific vocabularies that occur in more than 5 out of 10000 articles, and uses natural language processing and semantic smoothing to combine semantically related terms used in the literature, which may be treated as distinct (leading to loss of statistical power) by earlier automated approaches such as NeuroSynth (13). Additionally, whereas NeuroSynth is designed for in-sample statistical inference, NeuroQuery is intended to provide out-of sample prediction: given a query, it uses its semantic model to list related studies and their reported activations (13). Therefore, NeuroQuery provides a valuable complement to NeuroSynth as a means of obtaining quantitative summaries of the human cognitive mapping literature.

Cognitive circuits from intracranial electrical stimulation

We analyze a dataset of direct intracranial electrical stimulation (iES) in surgical patients ($N = 1,537$ cortical sites in 67 participants) (48). Briefly, participants self-reported their subjective experience upon direct electrical stimulation at different cortical locations, which cover all human intrinsic brain networks. Integration of the first-person reports yields probabilistic maps of

how direct stimulation elicits broad categories of cognitive processes. As reported in the original publication (48), “eight categories were employed: (1) somato-motor effects, including somatosensations and any observable motor effects, such as muscle twitches or limb movements; (2) visual effects, from simple phosphenes to more complex perturbations of visual perception such as distortion of faces; (3) olfactory effects; (4) vestibular effects, such as feelings of rotation, flotation and acceleration; (5) emotional effects of either positive or negative valence; (6) language effects, including arrest or alteration (for example, slurring) of speech; (7) memory recall; and (8) physiological and interoceptive effects, such as perceived changes in body temperature or heart rate”. Among these, four overlap with the cognitive domains from (59) and were used for the present work: ‘emotion’, ‘language’, ‘memory’, and ‘vision’. Classified results were then pooled to yield an estimate of the prevalence of each effect type for each of the brain’s intrinsic connectivity networks.

Cognitive circuits from Lesion Network Mapping

Classical neuropsychology identifies the causal roles of brain regions by identifying spatial overlap among lesions that cause the same symptoms. Lesion network mapping (5, 6, 108) extends this principle from overlap in 3D anatomical space, to overlap in ‘network space’: it identifies regions that overlap in terms of their functional connectivity with the sites of lesions that cause the same symptoms. This approach can also be extended to sites of stimulation that ameliorate the same symptoms (49). Here we include publicly available maps from two such studies: (1) a psychosis circuit, derived from the functional connectivity of regions whose lesion induces psychosis symptoms (50) ($N = 153$ lesion cases), available in MNI-152 space from NeuroVault at <https://neurovault.org/collections/20510>; and (2) a depression circuit, derived from the functional convergence of regions whose lesion induces depressive symptoms, and whose brain stimulation (using DBS or TMS) alleviates depressive symptoms (49) (14 studies; $N = 461$ lesion cases and $N = 251$ stimulation cases), available from NeuroVault at <https://neurovault.org/images/787859>.

Regional morphometric disease association from ENIGMA consortium

Spatial maps of case-versus-control cortical thickness were obtained by including all the neurological, neurodevelopmental, and psychiatric diagnostic categories available from the ENIGMA (Enhancing Neuroimaging Genetics through Meta-Analysis) consortium (17, 18) and the Enigma toolbox (<https://github.com/MICA-MNI/ENIGMA>) (14) and recent related publications ([\[crossdisorder_vulnerability\]\(#\)\) \(133\), except for obesity and schizotypy. This resulted in a total of 11 maps, pertaining to 22q11.2 deletion syndrome \(134\), attention-deficit/hyperactivity disorder \(135\), autism spectrum disorder \(136\), idiopathic generalized epilepsy \(137\), right temporal lobe epilepsy \(137\), left temporal lobe epilepsy \(137\), depression \(138\), obsessive-compulsive disorder \(139\), schizophrenia \(140\), bipolar disorder \(141\), and Parkinson’s disease \(142\). The ENIGMA consortium is a data-sharing initiative that relies on standardized image acquisition and processing pipelines, such that cortical thickness maps are comparable \(18\). Altogether, over 17 000 patients were scanned across the eleven diagnostic categories, against almost 22 000 controls. The values for each map are z-scored effect sizes \(Cohen’s \$d\$ \) of cortical thickness in patient populations versus healthy controls. Imaging and processing protocols can be found at <http://enigma.ini.usc.edu/protocols/>. For every brain region, we constructed an 11-element vector of cortical thickness changes, where each element represents a diagnostic category’s change in cortical thickness at the region.](https://github.com/netneurolab/hansen_</p>
</div>
<div data-bbox=)

Molecular signatures of regional disease risk

Genome-wide association studies

The ENIGMA toolbox (<https://github.com/MICA-MNI/ENIGMA>) (14) provides lists of disease-associated genes, obtained from genome-wide association studies: ADHD (143), autistic spectrum disorder (144), bipolar disorder (145), major depressive disorder (100) and schizophrenia (146).

DisGeNET database

Zeighami et al. (98) obtained lists of genes pertaining to 40 common human brain diseases, identified through the DisGeNET database (www.disgenet.org) (147, 148), a platform aggregated from multiple sources including curated repositories, GWAS catalogs, animal models, and the scientific literature. After removing substance-related disorders, those with overlapping names (e.g., ‘post-partum depression’ vs ‘depression’) and those with very broad names (e.g., ‘learning disorders’), we retained 33 disorders, categorised into 7 groups: Neurocognitive disorders and Dementias (Alzheimer Disease; Amnesia; Dementia; Delirium; Frontotemporal Lobar Degeneration); Neurodegenerative and movement disorders (Amyotrophic Lateral Sclerosis; Ataxia; Dystonia; Huntington Disease; Hereditary Spastic Paraplegia; Parkinson Disease; Tremor); Neurodevelopmental disorders (Autistic Disorder; Neurofibromatosis 1; Williams Syndrome); Brain tumors (Astrocytoma; Glioblastoma; Glioma; Gliosarcoma; Oligodendroglioma); Neurological disorders (Temporal Lobe Epilepsy; Hyperalgesia; Mi-

graine Disorders; Multiple Sclerosis; Neuralgia); Sleep and Wake disorders (Narcolepsy; Sleep Disorders); and Mood and Behaviour disorders (Schizophrenia; Bipolar Disorder; Depressive Disorder; Seasonal Affective Disorder; Panic Disorder; Obsessive Compulsive Disorder). As reported in (98), the gene set sizes range widely from frontotemporal lobar degeneration (11 genes) to schizophrenia (733 genes).

For both GWAS-derived and DisGeNET-derived lists of risk genes, one map of regional molecular risk per disease was obtained by averaging the normalised maps of the corresponding genes.

Multimodal networks of inter-regional interactions

Haemodynamic connectivity from functional MRI

Haemodynamic connectivity, commonly simply referred to as “functional connectivity”, captures how similarly pairs of cortical regions exhibit fMRI BOLD activity at rest (79). The fMRI BOLD time-series picks up on magnetic differences between oxygenated and deoxygenated haemoglobin to measure the haemodynamic response: the oversupply of oxygen to active brain regions (149).

The dataset of functional and structural neuroimaging data used in this work came from the Human Connectome Project (HCP, <http://www.humanconnectome.org/>), Release Q3 (150). Per HCP protocol, all subjects gave written informed consent to the HCP consortium. These data contained fMRI and diffusion weighted imaging (DWI) acquisitions from 100 unrelated subjects of the HCP 900 data release (150). All HCP scanning protocols were approved by the local Institutional Review Board at Washington University in St. Louis. Detailed information about the acquisition and imaging is provided in the dedicated HCP publications. Briefly, data were acquired using the following parameters. Structural MRI: 3D MPRAGE T1-weighted, TR = 2,400 ms, TE = 2.14 ms, TI = 1,000 ms, flip angle = 8 deg, FOV = 224 × 224, voxel size = 0.7 mm isotropic. Two sessions of 15-min resting-state fMRI: gradient-echo EPI, TR = 720 ms, TE = 33.1 ms, flip angle = 52 deg, FOV = 208 × 180, voxel size = 2 mm isotropic. Here, we used functional data from only the first scanning session, in LR direction. HCP-minimally preprocessed data (151) were used for all acquisitions. The minimal preprocessing pipeline includes bias field correction, functional realignment, motion correction, and spatial normalisation to Montreal Neurological Institute (MNI-152) standard space with 2mm isotropic resampling resolution (151). We also removed the first 10 volumes, to allow magnetisation to reach steady state. Additional denoising steps were performed using the SPM12-based toolbox CONN (<http://www.nitrc.org/projects/conn>), version 17f (152). To reduce noise due to cardiac and motion artifacts, we applied the anatomical CompCor method of denoising the functional data. The anatomical Comp-

Cor method (also implemented within the CONN toolbox) involves regressing out of the functional data the following confounding effects: the first five principal components attributable to each individual’s white matter signal, and the first five components attributable to individual cerebrospinal fluid (CSF) signal; six subject-specific realignment parameters (three translations and three rotations) as well as their first-order temporal derivatives (153). Linear detrending was also applied, and the subject-specific denoised BOLD signal timeseries were band-pass filtered to eliminate both low-frequency drift effects and high-frequency noise, thus retaining frequencies between 0.008 and 0.09 Hz. The parcellated time-series were used to construct functional connectivity matrices as a Pearson correlation coefficient between pairs of regional time-series. A group-average functional connectivity matrix was constructed as the mean functional connectivity across all individuals.

Metabolic connectivity from dynamic FDG-PET

Metabolic connectivity indexes how similarly two cortical regions metabolize glucose over time and therefore how similarly two cortical regions consume energy (47). Here we followed the same procedures as in (47). Briefly, volumetric 4D PET images of [^{18}F]-fluorodeoxyglucose (FDG, a glucose analogue) tracer uptake over time were obtained from Jamadar et al. (84). Specifically, 26 healthy participants (77% female, 18–23 years old) were recruited from the general population and underwent a 95 minute simultaneous MR-PET scan in a Siemens (Erlangen) Biograph 3-Tesla molecular MR scanner. Participants were positioned supine in the scanner bore with their head in a 16-channel radiofrequency head coil and were instructed to lie as still as possible with eyes open and think of nothing in particular. FDG (average dose 233 MBq) was infused over the course of the scan at a rate of 36 mL/h using a BodyGuard 323 MR-compatible infusion pump (Caesarea Medical Electronics, Caesarea, Israel). Infusion onset was locked to the onset of the PET scan. This data has been validated and analyzed previously in (154, 155).

PET images were reconstructed and preprocessed according to (155). Specifically, the 5700-second PET time-series for each subject was binned into 356 3D sinogram frames each of 16-second intervals. The attenuation for all required data was corrected via the pseudo-CT method (156). Ordinary Poisson-Ordered Subset Expectation Maximization algorithm (3 iterations, 21 subsets) with point spread function correction was used to reconstruct 3D volumes from the sinogram frames. The reconstructed DICOM slices were converted to NIFTI format with size 344×344×127 (voxel size: 2.09×2.09×2.03 mm³) for each volume. A 5 mm FWHM Gaussian postfilter was applied to each 3D volume. All 3D volumes were temporally concatenated to form a 4D (344×344×127×356) NIFTI volume. A guided motion

correction method using simultaneously acquired MRI was applied to correct the motion during the PET scan. 225 16-second volumes were retained commencing for further analyses.

Next, the 225 PET volumes were motion corrected (FSL MCFLIRT (157)) and the mean PET image was brain extracted and used to mask the 4D data. The fPET data were further processed using a spatiotemporal gradient filter to remove the accumulating effect of the radiotracer and other low-frequency components of the signal (158). Finally, each time point of the PET volumetric time-series were registered to MNI152 template space using Advanced Normalization Tools in Python (ANTSpy, <https://github.com/ANTsX/ANTsPy>), parcellated to 68 cortical regions according to the Desikan-Killiany atlas, and time-series at pairs of cortical regions were correlated (Pearson's r) to construct a metabolic connectivity matrix for each subject. A group-averaged metabolic connectome was obtained by averaging connectivity across subjects.

Electrophysiological connectivity from magnetoencephalography

Electrophysiological connectivity was measured using magnetoencephalography (MEG) recordings, which tracks the magnetic field produced by neural currents (47). Resting state MEG data was acquired for $n = 33$ unrelated healthy young adults (age range 22–35 years) from the Human Connectome Project (S900 release (150)). The data includes resting state scans of approximately 6 minutes long and noise recording for all participants. MEG anatomical data and 3T structural MRI of all participants were also obtained for MEG preprocessing.

The present MEG data was first processed and used by Hansen et al. (47), Shafiei et al. (80). Resting state MEG data was preprocessed using the open-source software, Brainstorm (<https://neuroimage.usc.edu/brainstorm/> (159)), following the online tutorial for the HCP dataset (<https://neuroimage.usc.edu/brainstorm/Tutorials/HCP-MEG>). MEG recordings were registered to individual structural MRI images before applying the following preprocessing steps. First, notch filters were applied at 60, 120, 180, 240, and 300 Hz, followed by a high-pass filter at 0.3 Hz to remove slow-wave and DC-offset artifacts. Next, bad channels from artifacts (including heartbeats, eye blinks, saccades, muscle movements, and noisy segments) were removed using Signal-Space Projections (SSP).

Pre-processed sensor-level data was used to construct a source estimation on HCP's fsLR4k cortex surface for each participant. Head models were computed using overlapping spheres and data and noise covariance matrices were estimated from resting state MEG and noise recordings. Linearly constrained minimum variance (LCMV) beamformers was used to obtain the source activity for each participant. Data covariance regulariza-

tion was performed and the estimated source variance was normalized by the noise covariance matrix to reduce the effect of variable source depth. All eigenvalues smaller than the median eigenvalue of the data covariance matrix were replaced by the median. This helps avoid instability of data covariance inversion caused by the smallest eigenvalues and regularizes the data covariance matrix. Source orientations were constrained to be normal to the cortical surface at each of the 8 000 vertex locations on the cortical surface, then parcellated according to the Desikan-Killiany atlas (58).

After preprocessing and parcellating the data, amplitude envelope correlations were performed between time-series at each pair of brain regions, for six canonical frequency bands separately (delta (2–4 Hz), theta (5–7 Hz), alpha (8–12 Hz), beta (15–29 Hz), low gamma (30–59 Hz), and high gamma (60–90 Hz)). Amplitude envelope correlation is applied instead of directly correlating the time-series because of the high sampling rate (2034.5 Hz) of the MEG recordings. An orthogonalization process was applied to correct for the spatial leakage effect by removing all shared zero-lag signals (160). The composite electrophysiological connectivity matrix is the first principal component of all six connectivity matrices (47).

Gene co-expression from microarray transcriptomics

Correlated gene expression represents the transcriptional similarity between pairs of cortical regions (47, 88). Regional microarray expression data were obtained from 6 post-mortem brains (1 female, ages 24.0–57.0, 42.50 ± 13.38) provided by the Allen Human Brain Atlas (AHBA, <https://human.brain-map.org> (87)). We followed the same preprocessing as recently described (53). Briefly, the Allen Human Brain Atlas (AHBA) is a publicly available transcriptional atlas containing gene expression data measured with DNA microarrays and sampled from hundreds of histologically validated neuroanatomical structures across six (five male and one female) normal postmortem human brains. We extracted and mapped gene expression data to the 68 cortical ROIs of the Desikan-Killiany atlas using the abagen toolbox (version 0.1.1; <https://github.com/rmarkello/abagen> (161)). Data was pooled between homologous cortical regions to ensure adequate coverage of both left (data from six donors) and right hemisphere (data from two donors). Distances between samples were evaluated on the cortical surface with a 2mm distance threshold. Only probes where expression measures were above a background threshold in more than 50% of samples were selected. A representative probe for a gene was selected based on highest intensity. 15,633 genes survived these preprocessing and quality assurance steps. Finally, the region $\times$ region correlated gene expression matrix was constructed by correlating (Pearson's r) the normalised gene expression profile at every pair of brain regions.

Receptor co-expression from Positron Emission Tomography

Receptor similarity indexes the degree to which the receptor density profiles at two cortical regions are correlated (47, 86). Here we followed the same procedures as in (47). Briefly, receptor densities were estimated using PET tracer studies for a total of 18 receptors and transporters, across 9 neurotransmitter systems, recently made available by Hansen and colleagues at https://github.com/netneurolab/hansen_receptors (86). These include dopamine (D_1 (162), D_2 (163–166), DAT (167), noradrenaline (NAT (168–171), serotonin (5-HT_{1A} (172), 5-HT_{1B} (172–177), 5-HT_{2A} (178), 5-HT₄ (178), 5-HT₆ (179, 180) 5-HTT (178)), acetylcholine ($\alpha_4\beta_2$ (181, 182), M_1 (183), VACHT (184, 185), glutamate (mGluR₅ (186, 187) NMDA (188, 189), GABA (GABA_A (190))), histamine (H_3 (191)), cannabinoid (CB₁ (192–195)) and opioid (MOR (196)). Volumetric PET images were registered to the MNI-ICBM 152 nonlinear 2009 (version c, asymmetric) template, averaged across participants within each study, then parcellated and receptors/transporters with more than one mean image of the same tracer (5-HT_{1B}, D_2 , VACHT) were combined using a weighted average (86). A region-by-region receptor similarity matrix was constructed by correlating (Pearson's r) receptor profiles at every pair of cortical regions.

Laminar similarity from BigBrain histology

Laminar similarity is estimated from histological data and aims to uncover how similar pairs of cortical regions are in terms of cellular distributions across the cortical laminae. Specifically, we use data from the BigBrain, a high-resolution (20 μ m) histological reconstruction of a post-mortem brain from a 65 year old male (90, 197). Cell-staining intensity profiles were sampled across 50 equivolumetric surfaces from the pial surface to the white matter surface to estimate laminar variation in neuronal density and soma size. Intensity profiles at various cortical depths can be used to approximately identify boundaries of cortical layers that separate supragranular (cortical layers I–III) granular (cortical layer IV), and infragranular (cortical layers V–VI) layers.

The data were obtained on *fsaverage* surface (164k vertices) from the BigBrainWarp toolbox (198) and were parcellated into 68 cortical regions according to the Desikan-Killiany atlas (58). The region $\times$ region laminar similarity matrix was calculated as the partial correlation (Pearson's r) of cell intensities between pairs of cortical regions, after correcting for the mean intensity across cortical regions. Laminar similarity was first introduced in Paquola et al. (90) and has also been referred to as “microstructure profile covariance”.

Similarity of subjective response to intracranial electrical stimulation

We obtained a matrix of inter-regional similarity of subjective experiences elicited by direct intracranial electrical stimulation, using the data from (48) as described above. Based on the intrinsic connectivity network that it belongs to, each cortical region has a ‘fingerprint’ given by its probability of eliciting each of the 8 classes of subjective experiences from (48) upon electrical stimulation. Correlation of these probability fingerprints produces a matrix of similarity of subjective elicitation between each pair of cortical regions.

Structural connectivity

We used diffusion MRI (dMRI) data from the same 100 unrelated participants (54 females and 46 males, mean age = 29.1 ± 3.7 years) of the HCP 900 participants data release (199), as described above. The diffusion weighted imaging (DWI) acquisition protocol is covered in detail elsewhere (151). The diffusion MRI scan was conducted on a Siemens 3T Skyra scanner using a 2D spin-echo single-shot multiband EPI sequence with a multi-band factor of 3 and monopolar gradient pulse. The spatial resolution was 1.25 mm isotropic. TR=5500 ms, TE=89.50ms. The b-values were 1000, 2000, and 3000 s/mm². The total number of diffusion sampling directions was 90, 90, and 90 for each of the shells in addition to 6 b0 images. We used the version of the data made available in DSI Studio-compatible format at <http://brain.labsolver.org/diffusion-mri-templates/hcp-842-hcp-1021> (200).

We adopted previously reported procedures to reconstruct the human connectome from DWI data. The minimally-preprocessed DWI HCP data (151) were corrected for eddy current and susceptibility artifact. DWI data were then reconstructed using q-space diffeomorphic reconstruction (QSDR (201)), as implemented in DSI Studio (www.dsi-studio.labsolver.org). QSDR is a model-free method that calculates the orientational distribution of the density of diffusing water in a standard space, to conserve the diffusible spins and preserve the continuity of fiber geometry for fiber tracking. QSDR first reconstructs diffusion-weighted images in native space and computes the quantitative anisotropy (QA) in each voxel. These QA values are used to warp the brain to a template QA volume in Montreal Neurological Institute (MNI) space using a nonlinear registration algorithm implemented in the statistical parametric mapping (SPM) software. A diffusion sampling length ratio of 2.5 was used, and the output resolution was 1 mm. A modified FACT algorithm (202) was then used to perform deterministic fiber tracking on the reconstructed data, with the following parameters (203): angular cutoff of 55°, step size of 1.0 mm, minimum length of 10 mm, maximum length of 400 mm, spin density function smoothing

of 0.0, and a QA threshold determined by DWI signal in the cerebrospinal fluid. Each of the streamlines generated was automatically screened for its termination location. A white matter mask was created by applying DSI Studio's default anisotropy threshold (0.6 Otsu's threshold) to the spin distribution function's anisotropy values. The mask was used to eliminate streamlines with premature termination in the white matter region. Deterministic fiber tracking was performed until 1,000,000 streamlines were reconstructed for each individual.

For each individual, their structural connectome was reconstructed by drawing an edge between each pair of regions i and j from the Desikan-Killiany cortical atlas (204) if there were white matter tracts connecting the corresponding brain regions end-to-end; edge weights were quantified as the number of streamlines connecting each pair of regions, normalised by ROI distance and size.

A group-consensus matrix A across participants was then obtained using the distance-dependent procedure of Betzel and colleagues, to mitigate concerns about inconsistencies in reconstruction of individual participants' structural connectomes (205). This approach seeks to preserve both the edge density and the prevalence and length distribution of inter- and intra-hemispheric edge length distribution of individual participants' connectomes, and it is designed to produce a representative connectome (205, 206). This procedure produces a binary consensus network indicating which edges to preserve. The weight of each non-zero edge is then computed as the mean of the corresponding non-zero edges across participants.

Macaque multimodal data

Macaque functional MRI

The non-human primate MRI data was made available as part of the Primate neuroimaging Data-Exchange (PRIME-DE) monkey MRI data sharing initiative, a recently introduced open resource for non-human primate imaging (63).

The data preprocessing and denoising followed the same procedures as in a previous publication (69). We used fMRI data from rhesus macaques (*Macaca mulatta*) scanned at Newcastle University. This sample includes 14 exemplars (12 male, 2 female); Age distribution: 3.9-13.14 years; Weight distribution: 7.2-18 kg (full sample description available online: http://fcon_1000.projects.nitrc.org/indi/PRIME/files/newcastle.csv and http://fcon_1000.projects.nitrc.org/indi/PRIME/newcastle.html).

Ethics approval: All of the animal procedures performed were approved by the UK Home Office and comply with the Animal Scientific Procedures Act (1986) on the care and use of animals in research and with the European Directive on the protection of animals used in re-

search (2010/63/EU). We support the Animal Research Reporting of In Vivo Experiments (ARRIVE) principles on reporting animal research. All persons involved in this project were Home Office certified and the work was strictly regulated by the U.K. Home Office. Local Animal Welfare Review Body (AWERB) approval was obtained. The 3Rs principles compliance and assessment was conducted by National Centre for 3Rs (NC3Rs). Animal in Sciences Committee (UK) approval was obtained as part of the Home Office Project License approval. Animal care and housing: All animals were housed and cared for in a group-housed colony, and animals performed behavioural training on various tasks for auditory and visual neuroscience. No training took place prior to MRI scanning. Macaque MRI acquisition

Animals were scanned in a vertical Bruker 4.7T primate dedicated scanner, with single channel or 4-8 channel parallel imaging coils used. No contrast agent was used. Optimization of the magnetic field prior to data acquisition was performed by means of 2nd order shim, Bruker and custom scanning sequence optimization. Animals were scanned upright, with MRI compatible headpost or non-invasive head immobilisation, and working on tasks or at rest (here, only resting-state scans were included). Eye tracking, video and audio monitoring were employed during scanning. Resting-state scanning was performed for 21.6 minutes, with a TR of 2600ms, 17ms TE, Effective Echo Spacing of 0.63ms, voxels size $1.22 \times 1.22 \times 1.24$. Phase Encoding Direction: Encoded in columns. Structural scans comprised a T1 structural, MDEFT sequence with the following parameters: TE: 6ms; TR: 750 ms; Inversion delay: 700ms; Number of slices: 22; In-plane field of view: $12.8 \times 9.6 \text{ cm}^2$ on a grid of 256×192 voxels; Voxel resolution: $0.5 \times 0.5 \times 2\text{mm}$; Number of segments: 8.

The macaque MRI data were preprocessed using the recently developed pipeline for non-human primate MRI analysis, *Pypreclin*, which addresses several specificities of monkey research. The pipeline is described in detail in the associated publication (207). Briefly, it includes the following steps: (i) Slice-timing correction. (ii) Correction for the motion-induced, time-dependent B0 inhomogeneities. (iii) Reorientation from acquisition position to template; here, we used the recently developed National Institute of Mental Health Macaque Template (NMT): a high-resolution template of the average macaque brain generated from in vivo MRI of 31 rhesus macaques (*Macaca mulatta*). (iv) Realignment to the middle volume using FSL MCFLIRT function. (v) Normalisation and masking using Joe's Image Program (JIP-align) routine (<http://www.nmr.mgh.harvard.edu/~jbm/jip/>, Joe Mandeville, Massachusetts General Hospital, Harvard University, MA, USA), which is specifically designed for preclinical studies: the normalization step aligns (affine) and warps (non-linear alignment using distortion field) the anatomical data into a generic template space. (vi) B1 field correction for low-frequency intensity non-uniformities present in the data. (vii) Coreg-

istration of functional and anatomical images, using JIP-align to register the mean functional image (moving image) to the anatomical image (fixed image) by applying a rigid transformation. The anatomical brain mask was obtained by warping the template brain mask using the deformation field previously computed during the normalization step. Then, the functional images were aligned with the template space by composing the normalization and coregistration spatial transformations.

Denoising: The aCompCor denoising method implemented in the CONN toolbox (152) was used to denoise the macaque functional MRI data, to ensure consistency with the human data analysis pipeline. White matter and CSF masks were obtained from the corresponding probabilistic tissue maps of the high-resolution NMT template (eroded by 1 voxel); their first five principal components were regressed out of the functional data, as well as linear trends and 6 motion parameters (3 translations and 3 rotations) and their first derivatives. To make human and macaque data comparable, the macaque data were also bandpass filtered in the same 0.008-0.09 Hz range used for human data. Out of the 14 total animals present in the Newcastle sample, 10 had available awake resting-state fMRI data; of these 10, all except the first animal had two scanning sessions available. Thus, the total was 19 distinct sessions across 10 individual macaques. Data were parcellated into the 82-ROI cortical parcellation of Kötter and Wanke (60). The parcellated time-series were used to construct functional connectivity matrices as a Pearson correlation coefficient between pairs of regional time-series. A group-average functional connectivity matrix was constructed as the mean functional connectivity across all individuals.

Macaque structural connectivity

Anatomical connectivity for the macaque brain was obtained from the fully weighted, whole-cortex macaque connectome recently developed by Shen and colleagues (208). This connectome was generated by combining information from two different axonal tract-tracing studies from the CoCoMac database (<http://cocomac.g-node.org/main/index.php>) (209) with diffusion-based tractography obtained from nine adult macaques (*Macaca mulatta* and *Macaca fascicularis*) (208). The resulting connectome provides a matrix of weighted, directed anatomical connectivity between each of the cortical ROIs of the Regional Mapping atlas of Kötter and Wanke (208).

Macaque cortical gene expression from stereo-seq

We used cortex-wide macaque gene expression data recently made available by (52), who combined single-nucleus RNA sequencing (“snRNA-seq”) with high-resolution, large-field-of view spatial transcrip-

tomics from spatiotemporal enhanced resolution omics-sequencing (“stereo-seq”) (210). Specifically, the authors made available (<https://macaque.digital-brain.cn/spatial-omics>) post-mortem gene expression data covering 143 regions of the left cortical hemisphere of one 6yo male cynomolgus macaque (*Macaca fascicularis*). We refer the reader to (52) for details.

Briefly, Chen and colleagues obtained 119 coronal sections at 500- μ m spacing, covering the entire cortex of the left hemisphere, which were used for stereo-seq transcriptomics (52). Adjacent 50- μ m thick sections were also acquired for regional microdissection and single-nucleus RNA sequencing (snRNA-seq) analysis, as well as 10- μ m sections adjacent to each stereo-seq section, which were used for the anatomical parcellation of brain regions via immunostaining (52). As reported in (52), for each coronal section, the cortical region and layer parcellation were manually delineated on stereo-seq data background, based on cytoarchitectural pattern (e.g. cell density, cell size) revealed by total mRNA expression, nucleic acid staining, and NeuN staining of adjacent sections. Gene expression data were made available at <https://macaque.digital-brain.cn/spatial-omics> for 143 cortical regions of the left hemisphere, including prefrontal, frontal, cingulate, somatosensory, insular, auditory, temporal, parietal, occipital and piriform areas. On the data-sharing portal, separate normalised gene expression data are made available for each region and for each of its cortical layers. To make the gene expression comparable across our datasets, we used the version of the macaque gene expression data that was mirrored between hemispheres and mapped onto the “regional mapping” macaque atlas of Kötter and Wanke (60), as made available by (53). To obtain inter-regional covariance of gene expression, each gene expression pattern was z-scored and then genes were correlated across regions, thereby obtaining a regions-by-regions matrix.

Macaque receptor density from in vitro receptor autoradiography

In vitro autoradiography data for 14 neurotransmitter receptors were obtained from (51): AMPA, kainate, NMDA, GABA_A, GABA_B, GABA_{A/BZ}, M₁, M₂, M₃, α_1 , α_2 , 5HT_{1A}, 5HT_{2A}, D₁. The authors applied quantitative *in vitro* receptor autoradiography to label 14 neurotransmitter receptors in three male *Macaca fascicularis* brains (7.3 $\pm$ 0.6 years old; body weight 6 $\pm$ 0.8 kg) obtained from Covance Preclinical Services, where they were housed and used as control animals for pharmaceutical studies performed in compliance with legal requirements. Animal experimental procedures and husbandry had the approval of the respective Institutional Animal Care and Use Committee and were carried out in accordance with the European Council Directive of 2010 (51). We refer the reader to (51) for details.

Briefly, the data of density of receptors per neuron were made available for 109 cortical areas of the

macaque brain, which were identified based on their cytoarchitecture and receptor-architecture characteristics (51). To enable comparison across datasets, we used the version of the macaque receptor density data that was mapped onto the Regional Mapping macaque cortical parcellation of Köster and Wanke (60) as made available by (53). The matrix of similarity of regional receptor expression was obtained by correlating the z-scored receptor density patterns.

Mouse multimodal data

Mouse functional MRI

The mouse fMRI data used here have been published before (56). In vivo experiments were conducted in accordance with the Italian law (DL 26/214, EU 63/2010, Ministero della Sanita, Roma) and with the National Institute of Health recommendations for the care and use of laboratory animals (56). The animal research protocols for this study were reviewed and approved by the Italian Ministry of Health and the animal care committee of Istituto Italiano di Tecnologia (IIT). All surgeries were performed under anesthesia.

Adult (< 6 months old) male C57BL/6J mice were used throughout the study. Mice were group housed in a 12:12 hours light-dark cycle in individually ventilated cages with access to food and water ad libitum and with temperature maintained at 21 ± 1 degrees centigrade and humidity at $60 \pm 10\%$. All the imaged mice were bred in the same vivarium and scanned with the same MRI scanner and imaging protocol employed for the awake scans (see below).

A first group of mice ($n = 10$, awake dataset) underwent head-post surgery, scanner habituation and fMRI image acquisitions as described below. See (56) for the full surgical, habituation, and scanner protocol. The scans so obtained constitute the awake rsfMRI mouse dataset we used throughout our study. To prevent motion, each mouse was secured using an implanted head-post in the custom-made MRI-compatible animal cradle and the body of the mouse was gently restrained (for details of the headpost implantation and habituation protocol, see the original publication (56)). All scans were acquired at the IIT laboratory in Rovereto (Italy) on a 7.0 Tesla MRI scanner (Bruker Biospin, Ettlingen) with a BGA-9 gradient set, a 72 mm birdcage transmit coil, and a four-channel (awake, halothane) or three-channel (medetomidine-isoflurane) solenoid receive coil. Awake scans were acquired using a single-shot echo planar imaging (EPI) sequence with the following parameters: TR/TE=1000/15 ms, flip angle=60 degrees, matrix=100 x 100, FOV=2.3 x 2.3 cm, 18 coronal slices (voxel-size 230 x 230 x 600 μm), slice thickness=600 μm and 1920 time points, for a total time of 32 minutes (56). Preprocessing of fMRI images was carried out as described in previous work (56). Briefly, the first 2 min-

utes of the time series were removed to account for thermal gradient equilibration. Functional MRI time-series were then time despiked (3dDespike, AFNI), motion corrected (MCFLIRT, FSL), skull stripped (FAST, FSL) and spatially registered (ANTs registration suite) to an in-house mouse brain template with a spatial resolution of $0.23 \times 0.23 \times 0.6 \text{ mm}^3$. Denoising involved the regression of 25 nuisance parameters. These were: average cerebral spinal fluid signal plus 24 motion parameters determined from the 3 translation and rotation parameters estimated during motion correction, their temporal derivatives and corresponding squared regressors. No global signal regression was employed. In-scanner head motion was quantified via calculations of frame-wise displacement (FD). To rule out a contribution of residual head-motion, we further introduced frame-wise fMRI scrubbing ($\text{FD} > 0.075 \text{ mm}$). The resulting time series were band-pass filtered (0.01-0.1 Hz band) and then spatially smoothed with a Gaussian kernel of 0.5 mm full width at half maximum. Finally, the time-series were trimmed to ensure that the same number of timepoints were included for all animals, resulting in 1414 volumes per animal. Finally data were parcellated into 72 cortical symmetric regions from the Allen Mouse Brain Atlas (CCFv3) (61).

Mouse structural connectivity

For the mouse structural connectome, we used a parcellated version of the high-resolution mouse connectome of Coletta et al. (54). Below, we summarise how Coletta and colleagues obtained the high-resolution mouse structural connectome.

The present mouse structural connectome is based on “high-resolution models of the mouse brain connectome (100 μm^3) previously released by Knox and colleagues (211). The Knox connectome is based on 428 viral microinjection experiments in C57BL/6J male mice obtained from the Allen Mouse Brain Connectivity Atlas (<http://connectivity.brain-map.org/>). The connectome data were derived from imaging enhanced green fluorescent protein (eGFP)-labeled axonal projections that were then registered to the Allen Mouse Brain Atlas and aggregated according to a voxel-wise interpolation model” (54).

Before constructing the SC matrix, Coletta et al. ensured symmetry along the right-left axis for all the major macrostructures of the mouse brain. To this purpose, they “flipped each macrostructure (isocortex, hippocampal formation, subcortical plate, pallidum, striatum, pons, medulla, midbrain, thalamus, hypothalamus, cerebellum, and olfactory bulb) along the sagittal midline (once for the right hemisphere and once for the left hemisphere) and took the intersection with the respective nonflipped macrostructure. This procedure resulted in the removal of a set of nonsymmetric voxel (total fraction, 8.6%), the vast majority of which reside in fringe

white/gray matter or cerebrospinal fluid/gray matter interfaces. The removal of these nonsymmetric voxels did not substantially affect the network structure of the resampled connectome, as assessed with a spatial correlation analysis between the symmetrized and nonsymmetrized right ipsilateral (i.e., squared) connectome". Coletta et al. then filtered out fiber tracts and ventricular spaces, and estimated SC using a resampled version of the recently published voxel scale model of the mouse structural connectome (54), to make the original matrix computationally tractable. Resampling of the Knox et al. connectome was carried out by aggregating neighboring voxels according to a Voronoi diagram based on Euclidean distance between neighboring voxels, preserving the intrinsic architectural foundation of the connectome while minimizing spatial blurring and boundary effects between ontogenically distinct neuroanatomical divisions of the mouse brain, or white/gray matter, and parenchymal/ventricular interfaces (see Coletta et al. for details of the Voronoi aggregation scheme (54)). By averaging the connectivity profile of neighboring voxels based on their relative spatial arrangement, this strategy has also the advantage of mitigating limitations related to the enforced smoothness of source space used by the original kernel interpolation used by (54). A whole-brain connectome was then built under the assumption of brain symmetry. The obtained Voronoi diagram made it possible to map the results back into the original $100 \mu^3$ three-dimensional coordinate system of the Allen Institute mouse brain connectome [CCFv3] (61). For each pair of the 72 Allen Atlas cortical regions that we employed, their structural connectivity was obtained by averaging the connectivity of the respective constituent voxels.

Mouse brain gene expression from in situ hybridization

Mouse gene expression profiles were obtained using in situ hybridization data from the Allen Mouse Brain Atlas (55). Briefly, the Allen Mouse Brain Atlas consists of data acquired from a pipeline that includes semi-automated riboprobe generation, tissue preparation and sectioning, in-situ hybridization (ISH), imaging, and data post-processing. These data were acquired both sagittally and (for a smaller set of genes) coronally, and were further processed, aligned by the Allen Institute to their Common Coordinate Framework version 3 (CCFv3) reference atlas (61) and summarized voxelwise through a measure termed gene expression energy (defined as the sum of expressing pixel intensity divided by the sum of all pixels in a division), resulting in 3D gene expression images at a $200 \mu\text{m}$ isotropic resolution. This gene expression energy increases in regions of high expression, and is bounded by zero in regions of no expression. For our study, we used gene expression energy data from the coronal dataset (4345 gene expression images corresponding to 4082 unique genes), because of its whole-

brain coverage and data quality. Voxelwise gene expression data were further summarized as normalized mean expression within regions of interest as defined by the CCFv3 reference atlas. For each ROI and each gene, voxelwise expression energy data were averaged over voxels containing valid expression signal. Finally, independently for each region, gene expression data were normalized by regressing out mean gene expression across the brain.

Anatomical hierarchy from in vivo T1w:T2w ratio

For each species, we used maps of intracortical myelination (from T1w:T2w ratio) as an in vivo marker of the anatomical cortical hierarchy (53, 57, 95, 212). For the human, we used the map available in the neuromaps toolbox (1). For the macaque, we used the T1w:T2w ratio map originally from (213) and resampled to the Regional Mapping macaque cortical parcellation of Kötter and Wanke (60) by (53). For the mouse, we used the T1w:T2w ratio data from (57).

Statistical analyses

The use of non-parametric tests alleviated the need to assume normality of data distributions (which was not formally tested). All tests were two-sided, with an α value of 0.05. The effect sizes were estimated using Cohen's measure of the standardized mean difference, d . To ensure robustness to possible outliers, correlations were quantified using Spearman's rank-based non-parametric correlation coefficient.

Spatial autocorrelation-preserving surrogates with Moran spectral randomisation

Significance of spatial correlations is assessed against a population of 5,000 null maps with preserved spatial autocorrelation generated using Moran spectral randomisation. For each species, we generate null maps based on the inverse Euclidean distances between parcel centroids for that species, as implemented in the BrainSpace toolbox (<https://brainspace.readthedocs.io/en/latest/>) (97). Moran spectral randomisation quantifies the spatial autocorrelation in the data in terms of Moran's I coefficient, by computing spatial eigenvectors known as Moran eigenvector maps. The Moran eigenvectors are then used to generate null maps by imposing the spatial structure of the empirical data on randomised surrogate data (97, 214).

Multivariate gene-cognition association with Partial Least Squares

Partial Least Squares (PLS) analysis was used to relate regional gene expression to functional associations. PLS analysis is an unsupervised multivariate statistical technique that decomposes relationships between two datasets $X_{n \times g}$ and $Y_{n \times f}$ into orthogonal sets of latent variables with maximum covariance, which are linear combinations of the original data (215, 216). In the present case, $X_{n \times g}$ is regional gene expression across n regions (68 for the human; 82 for the macaque; and 72 for the mouse) and g genes (the same set of 81 for each species). $Y_{n \times f}$ is the matrix of regional functional associations, across 17 cognitive operations.

PLS finds components from the predictor variables (regional gene expression) that have maximum covariance with the response variables (regional functional associations). The PLS components (i.e., linear combinations of the weighted variables) are ranked by the covariance between predictor and response variables so that the first few PLS components provide a low-dimensional representation of the covariance between the higher-dimensional data matrices. Concretely, this is achieved by performing singular value decomposition (SVD) on the matrix $Y'X$, such that:

$$Y'X = U_{g \times t} S_{t \times t} V_{t \times f}' \quad (1)$$

where $U_{g \times t}$ and $V_{t \times f}$ are orthonormal matrices consisting of left and right singular vectors, and $S_{t \times t}$ is a diagonal matrix of singular values. The i th columns of U and V constitute a latent variable, and the i th singular value in S represents the covariance between singular vectors. The i th singular value is proportional to the amount of covariance between gene expression and cognitive relevance captured by the i th latent variable, where the effect size can be estimated as the ratio of the squared singular value to the sum of all squared singular values.

The statistical significance of the covariance explained by each PLS model is tested by permuting the response variables 10 000 times while considering the spatial dependency of the data by using spatial autocorrelation-preserving permutation tests (214, 217). The p-value is computed as the proportion of null singular values that are greater in magnitude than the empirical singular values. Thus, these p-values represent the probability that the observed spatial correspondence between genes and cognitive operations could occur by randomly correlating maps with comparable spatial autocorrelation.

Network null models

We used two different network null models to disambiguate the role of connectome topology and geometric

embedding in shaping control energy (217). The first null model is the well-known Maslov-Sneppen degree-preserving rewired network, whereby edges are swapped so as to randomise the topology while preserving the exact binary degree of each node (degree sequence), and the overall distribution of edge weights (218). As a second, more stringent null model, we adopted a null model that in addition to preserving exactly the same degree sequence and exactly the same edge weight distribution as the original network, also approximately preserves the original network's edge length distribution (based on Euclidean distance between regions), and the weight-length relationship (219).

For each null model, we generated a population of 500 null networks starting from the empirical connectome, and computed the control energy between each pair of cognitive brain states from NeuroSynth, as done for the empirical connectome. We compared the overall control energy between all possible states obtained from the empirical connectome and from the distribution of null instances.

Dominance analysis

To consider all inter-regional interactions together and evaluate their respective contributions to LLM-derived functional similarity, we performed a dominance analysis with all multimodal empirical networks as predictors, and the network of functional similarity from LLMs as target. Dominance analysis seeks to determine the relative contribution ("dominance" of each independent variable to the overall fit (adjusted R^2)) of the multiple linear regression model (<https://github.com/dominance-analysis/dominance-analysis>) (91). This is done by fitting the same regression model on every combination of predictors ($2^p - 1$ submodels for a model with p predictors). Total dominance is defined as the average of the relative increase in R^2 when adding a single predictor of interest to a submodel, across all $2^p - 1$ submodels. The sum of the dominance of all input variables is equal to the total adjusted R^2 of the complete model, making the percentage of relative importance an intuitive method that partitions the total effect size across predictors. Therefore, unlike other methods of assessing predictor importance, such as methods based on regression coefficients or univariate correlations, dominance analysis accounts for predictor-predictor interactions and is interpretable.

Data and code availability

We have made available code to perform cognitive cartography of brain function and dysfunction using large language models at <https://github.com/Hana-Ali/neuroLLM>. NeuroSynth meta-analytic maps are freely available from the NeuroSynth

database at <https://github.com/neurosynth/neurosynth>. NeuroQuery meta-analytic maps are freely available from <https://neuroquery.org/>. ENIGMA cortical thickness data are provided as part of the ENIGMA Toolbox (v1.1.3), available at <https://github.com/MICA-MNI/ENIGMA>. Maps of subjective experience probabilities elicited by intracranial electrical stimulation are available from the supplementary materials of (48). Lesion network mapping circuits for depression and psychosis are available from NeuroVault at <https://neurovault.org/images/787859> and <https://neurovault.org/collections/20510>. Mouse Neuropixels maps are available from the International Brain Laboratory at <https://www.internationalbrainlab.com/data>. The neuromaps toolbox for fetching human brain maps (version 0.0.1) is freely available at <https://netneurolab.github.io/neuromaps/>. Human Connectome Project Young Adult fMRI and diffusion MRI data are available from <https://www.humanconnectome.org/study/hcp-young-adult>. Diffusion MRI data for the Human Connectome Project in DSI Studio-compatible format are available at <http://brain.labsolver.org/diffusion-mri-templates/hcp-842-hcp-1021>. The Human PET receptor and transporter maps are available at https://github.com/netneurolab/hansen_receptors. The macaque fMRI data are available from the PRIME-DE database (http://fcon_1000.projects.nitrc.org/indi/indiPRIME.html). The macaque connectome is available on Zenodo at <https://doi.org/10.5281/zenodo.1471588>. The CoCo-Mac database on which it is based, is also available online at <http://cocomac.g-node.org/main/index.php?>. The mouse fMRI data and mouse structural connectome are available from author A.G. Human gene expression data from the Allen Human Brain Atlas are available at <https://human.brain-map.org/>. Macaque cortical gene expression data from (52) are available at <https://macaque.digital-brain.cn/spatial-omics>. The dataset is provided by Brain Science Data Center, Chinese Academy of Sciences (<https://braindatacenter.cn/>). Mouse gene expression data from in situ hybridization are available from the Allen Mouse Brain Atlas (55) at <https://mouse.brain-map.org/>. The original macaque receptor density data from autoradiography are available from <https://balsa.wustl.edu/study/P2Nql> and <https://search.kg.ebrains.eu/instances/de62abc1-7252-4774-9965-5040f5e8fb6b> (51). The original map of macaque intracortical myelination from T1w:T2w ratio from (51) is available at <https://balsa.wustl.edu/study/P2Nql>. Human networks of metabolic connectivity, electrophysiological connectivity, and laminar similarity are available on GitHub at https://github.com/netneurolab/hansen_many_networks/tree/v1.0.0. The abagen toolbox for processing of the AHBA human transcriptomic dataset is available at <https://abagen.readthedocs.io/>. The BrainSpace toolbox for Moran spectral randomisation is available at <https://brainspace.readthedocs.io/en/latest/>. MATLAB code used to generate geometry-

preserving null networks is freely available at <https://www.brainnetworkslab.com/coderesources>. Third-party Python software (version 3.8 was used) for dominance analysis is freely available at <https://github.com/dominance-analysis/dominance-analysis>. The CONN toolbox (v17f) for fMRI preprocessing is freely available at <http://www.nitrc.org/projects/conn>.

Acknowledgments

The authors are grateful to Dr Michael D. Fox and Dr Shan Siddiqi for making available the lesion circuit maps of depression and psychosis; and to members of the Network Neuroscience Lab for valuable feedback. AIL acknowledges the support of St John's College, Cambridge; and a Wellcome Early Career Award (grant number 226924/Z/23/Z). We also acknowledge support of the Brain Canada Foundation, through the Canada Brain Research Fund with the support of Health Canada [to D.B.]; the National Institutes of Health (grants no. NIH R01 AG068563A and NIH R01 R01DA053301-01A1) [to D.B.]; the Canadian Institute of Health Research (grants no. CIHR 438531 and CIHR 470425) [to D.B.]; the Healthy Brains Healthy Lives initiative (Canada First Research Excellence fund) [to D.B.]; Google (Research Award, Teaching Award) [to D.B.]; the CIFAR Artificial Intelligence Chairs programme (Canada Institute for Advanced Research) [to D.B.]; the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program [DISCONN; no. 802371 to A.G.; and no. 101125054 - BRAINAMICS to A.G.]; BM acknowledges support from the Natural Sciences and Engineering Research Council of Canada (RGPIN-2017-04265), Canadian Institutes of Health Research (PJT-180439), and Canada Research Chairs Program (CRC-2022-00169). ZQL acknowledges support from the Fonds de Recherche du Québec – Nature et Technologies (FRQNT). Human connectome data were provided by the Human Connectome Project, WU-Minn Consortium (1U54MH091657; Principal Investigators David Van Essen and Kamil Ugurbil) funded by the 16 National Institutes of Health (NIH) institutes and centers that support the NIH Blueprint for Neuroscience Research, and by the McDonnell Center for Systems Neuroscience at Washington University. For the purpose of open access, the authors have applied a Creative Commons Attribution (CC BY) licence to any Author Accepted Manuscript version arising from this submission. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not reflect the views of the funders.

Conflicts of interest

None.

- [1] Ross D Markello, Justine Y Hansen, Zhen-Qi Liu, Vincent Bazinet, Golia Shafiei, Laura E Suárez, Nadia Blostein, Jakob Seidlitz, Sylvain Baillet, Theodore D Satterthwaite, et al. Neuromaps: structural and functional interpretation of brain maps. *Nature Methods*, 19(11):1472–1479, 2022.
- [2] Markus Axer and Katrin Amunts. Scale matters: The nested human connectome. *Science*, 378(6619):500–504, 2022.
- [3] Valerie J Sydnor, Bart Larsen, Danielle S Bassett, Aaron Alexander-Bloch, Damien A Fair, Conor Liston, Allyson P Mackey, Michael P Milham, Adam Pines, David R Roalf, Jakob Seidlitz, Ting Xu, Armin Raznahan, and Theodore D Satterthwaite. Neurodevelopment of the association cortices: Patterns, mechanisms, and implications for psychopathology. *Neuron*, 2021. ISSN 08966273. doi:10.1016/j.neuron.2021.06.016. URL <https://doi.org/10.1016/j.neuron.2021.06.016>. [Online; accessed 2021-07-26].
- [4] Anton Arkhipov, Nuno da Costa, Saskia de Vries, Trygve Bakken, Corbett Bennett, Amy Bernard, Jim Berg, Michael Buice, Forrest Collman, Tanya Daigle, et al. Integrating multimodal data to understand cortical circuit architecture and function. *Nature Neuroscience*, pages 1–14, 2025.
- [5] Shan H. Siddiqi, Konrad P. Kording, Josef Parvizi, and Michael D. Fox. Causal mapping of human brain function. *Nature Reviews Neuroscience*, 23(6):361–375, 2022. doi:10.1038/s41583-022-00583-8.
- [6] Michael D Fox. Mapping symptoms to brain networks with the human connectome. *New England Journal of Medicine*, 379(23):2237–2245, 2018.
- [7] Justine Y Hansen and Bratislav Misic. Integrating and interpreting brain maps. *Trends in Neurosciences*, 2025.
- [8] Laura E. Suárez, Ross D. Markello, Richard F. Betzel, and Bratislav Misic. Linking structure and function in macroscale brain networks. *Trends in Cognitive Sciences*, 24(4):302–315, 4 2020. ISSN 1879307X. doi:10.1016/j.tics.2020.01.008. PMID: 32160567 publisher: Elsevier Ltd.
- [9] Dora Angelaki, Brandon Benson, Julius Benson, Daniel Birman, Niccolò Bonacchi, Kcénia Bougrova, Sebastian A Bruijns, Matteo Carandini, Joana A Catarino, et al. A brain-wide map of neural activity during complex behaviour. *Nature*, 645(8079):177–191, 2025.
- [10] Huda Akil, Maryann E Martone, and David C Van Essen. Challenges and opportunities in mining neuroscience data. *science*, 331(6018):708–712, 2011.
- [11] Peter T Fox and Jack L Lancaster. Mapping context and content: the brainmap model. *Nature Reviews Neuroscience*, 3(4):319–321, 2002.
- [12] Tal Yarkoni, Russell A Poldrack, Thomas E Nichols, David C Van Essen, and Tor D Wager. Large-scale automated synthesis of human functional neuroimaging data. *Nature methods*, 8(8):665–670, 2011.
- [13] Jérôme Dockès, Russell A Poldrack, Romain Primet, Hande Gözükan, Tal Yarkoni, Fabian Suchanek, Bertrand Thirion, and Gaël Varoquaux. Neuroquery, comprehensive meta-analysis of human brain mapping. *eLife*, 9: e53385, 2020.
- [14] Sara Larivière, Casey Paquola, Bo-yong Park, Jessica Royer, Yezhou Wang, Oualid Benkarim, de Reinder Vos Wael, Sofie L. Valk, Sophia I. Thomopoulos, Matthias Kirschner, Lindsay B. Lewis, Alan C. Evans, Sanjay M. Sisodiya, Carrie R. McDonald, Paul M. Thompson, and Boris C. Bernhardt. The enigma toolbox: multiscale neural contextualization of multisite neuroimaging datasets. *Nature Methods* 2021 18:7, 18(7):698–700, 6 2021. ISSN 1548-7105. doi:10.1038/s41592-021-01186-4. publisher: Nature Publishing Group.
- [15] Thomas Donoghue and Bradley Voytek. Automated meta-analysis of the event-related potential (erp) literature. *Scientific Reports*, 12(1):1867, 2022.
- [16] Taylor Salo, Tal Yarkoni, Thomas E. Nichols, Jean-Baptiste Poline, Murat Bilgel, Katherine L. Bottenhorn, Dorota Jarecka, James D. Kent, Adam Kimbler, Dylan M. Nielson, Kendra M. Oudyk, Julio A. Peraza, Alexandre Pérez, Puck C. Reeders, Julio A. Yanes, and Angela R. Laird. Nimare: Neuroimaging meta-analysis research environment. *Aperture Neuro*, 3:1 – 32, 2023. doi: 10.52294/001c.87681. URL <https://doi.org/10.52294/001c.87681>.
- [17] Paul M Thompson, Jason L Stein, Sarah E Medland, Derrek P Hibar, Alejandro Arias Vasquez, Miguel E Renteria, Roberto Toro, Neda Jahanshad, Gunter Schumann, Barbara Franke, et al. The enigma consortium: large-scale collaborative analyses of neuroimaging and genetic data. *Brain imaging and behavior*, 8:153–182, 2014.
- [18] Paul M Thompson, Neda Jahanshad, Christopher RK Ching, Lauren E Salminen, Sophia I Thomopoulos, Joanna Bright, Bernhard T Baune, Sara Bertolin, Janita Bralten, Willem B Bruin, et al. Enigma and global neuroscience: A decade of large-scale studies of the brain in health and disease across more than 40 countries. *Translational psychiatry*, 10(1):100, 2020.
- [19] Simon B Eickhoff, Angela R Laird, Christian Grefkes, Ling E Wang, Karl Zilles, and Peter T Fox. Coordinate-based activation likelihood estimation meta-analysis of neuroimaging data: A random-effects approach based on empirical estimates of spatial uncertainty. *Human brain mapping*, 30(9):2907–2926, 2009.
- [20] Tor D Wager, Martin Lindquist, and Lauren Kaplan. Meta-analysis of functional neuroimaging data: current and future directions. *Social cognitive and affective neuroscience*, 2(2):150–158, 2007.
- [21] Peter T Fox, Lawrence M Parsons, and Jack L Lancaster. Beyond the single study: function/location metanalysis in cognitive neuroimaging. *Current opinion in neurobiology*, 8(2):178–187, 1998.
- [22] Angela R Laird, Jack J Lancaster, and Peter T Fox. Brainmap: the social evolution of a human brain mapping database. *Neuroinformatics*, 3:65–77, 2005.
- [23] Simon B Eickhoff, Danilo Bzdok, Angela R Laird, Florian Kurth, and Peter T Fox. Activation likelihood estimation meta-analysis revisited. *Neuroimage*, 59(3):2349–2361, 2012.
- [24] Angela R Laird, P Mickle Fox, Cathy J Price, David C Glahn, Angela M Uecker, Jack L Lancaster, Peter E Turkeltaub, Peter Kochunov, and Peter T Fox. Ale meta-analysis: Controlling the false discovery rate and performing statistical contrasts. *Human brain mapping*, 25(1):155–164, 2005.

- [25] AmanPreet Badhwar, David Kennedy, Jean-Baptiste Poline, and Roberto Toro. Distributed collaboration: the case for the enhancement of brainspell's interface. *Giga-Science*, 5(suppl_1):s13742–016, 2016.
- [26] Simon B Eickhoff, Thomas E Nichols, Angela R Laird, Felix Hoffstaedter, Katrin Amunts, Peter T Fox, Danilo Bzdok, and Claudia R Eickhoff. Behavior, sensitivity, and power of activation likelihood estimation characterized by massive empirical simulation. *Neuroimage*, 137:70–85, 2016.
- [27] Sébastien Bubeck, Varun Chandrasekaran, Ronen Eldan, Johannes Gehrke, Eric Horvitz, Ece Kamar, Peter Lee, Yin Tat Lee, Yuanzhi Li, Scott Lundberg, Harsha Nori, Hamid Palangi, Marco Tulio Ribeiro, and Yi Zhang. Sparks of artificial general intelligence: Early experiments with gpt-4. *arXiv preprint arXiv:2303.12712*, 2023. Preprint (March 2023).
- [28] Michael R. King. Administration of the text-based portions of a general iq test to five different large language models. *TechRxiv preprint*, 2023. preprint.
- [29] Jacob Devlin, Ming-Wei Chang, Kenton Lee, and Kristina Toutanova. Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: human language technologies, volume 1 (long and short papers)*, pages 4171–4186, 2019.
- [30] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser, and Illia Polosukhin. Attention is all you need. *Advances in neural information processing systems*, 30, 2017.
- [31] Tom Brown, Benjamin Mann, Nick Ryder, Melanie Subbiah, Jared D Kaplan, Prafulla Dhariwal, Arvind Neelakantan, Pranav Shyam, Girish Sastry, Amanda Askell, et al. Language models are few-shot learners. *Advances in neural information processing systems*, 33:1877–1901, 2020.
- [32] Thomas Serre and Ellie Pavlick. From prediction to understanding: Will ai foundation models transform brain science? *Neuron*, 2025.
- [33] Josh Achiam, Steven Adler, Sandhini Agarwal, Lama Ahmad, Ilge Akkaya, Florencia Leoni Aleman, Diogo Almeida, Janko Altschmidt, Sam Altman, Shyamal Anadkat, et al. Gpt-4 technical report. *arXiv preprint arXiv:2303.08774*, 2023.
- [34] Albert Q Jiang, Alexandre Sablayrolles, Antoine Roux, Arthur Mensch, Blanche Savary, Chris Bamford, Devendra Singh Chaplot, Diego de las Casas, Emma Bou Hanna, Florian Bressand, et al. Mixtral of experts. *arXiv preprint arXiv:2401.04088*, 2024.
- [35] Hugo Touvron, Louis Martin, Kevin Stone, Peter Albert, Amjad Almahairi, Yasmine Babaei, Nikolay Bashlykov, Soumya Batra, Prajwal Bhargava, Shruti Bhosale, et al. Llama 2: Open foundation and fine-tuned chat models. *arXiv preprint arXiv:2307.09288*, 2023.
- [36] Gemini Team, Rohan Anil, Sebastian Borgeaud, Jean-Baptiste Alayrac, Jiahui Yu, Radu Soricut, Johan Schalkwyk, Andrew M Dai, Anja Hauth, Katie Millican, et al. Gemini: a family of highly capable multimodal models. *arXiv preprint arXiv:2312.11805*, 2023.
- [37] Daya Guo, Dejian Yang, Haowei Zhang, Junxiao Song, Ruoyu Zhang, Runxin Xu, Qihao Zhu, Shirong Ma, Peiyi Wang, Xiao Bi, et al. Deepseek-r1: Incentivizing reasoning capability in llms via reinforcement learning. *arXiv preprint arXiv:2501.12948*, 2025.
- [38] Daniel Martin Katz, Michael James Bommarito, Shang Gao, and Pablo Arredondo. Gpt-4 passes the bar exam. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 382(2270): 20230254, 2024. doi:10.1098/rsta.2023.0254. Epub 2024 Feb 26.
- [39] Tiffany H. Kung, Morgan Cheatham, Arielle Medenilla, Czarina Sillos, Lorie De Leon, Camille Elepaño, Maria Madriaga, Rimel Aggabao, Giezel Diaz-Candido, James Maningo, and Victor Tseng. Performance of chatgpt on usmle: Potential for ai-assisted medical education using large language models. *PLOS Digital Health*, 2(2):e0000198, February 2023. doi: 10.1371/journal.pdig.0000198.
- [40] Jawad Ibn Ahad, Rafeed Mohammad Sultan, Abraham Kaikobad, Fuad Rahman, Mohammad Ruhul Amin, Nabeel Mohammed, and Shafin Rahman. Empowering meta-analysis: Leveraging large language models for scientific synthesis. *2024 IEEE International Conference on Big Data (BigData)*, pages 1615–1624, 2024.
- [41] Marcel Binz, Elif Akata, Matthias Bethge, Franziska Brändle, Fred Callaway, Julian Coda-Forno, Peter Dayan, Can Demircan, Maria K Eckstein, Noémi Éltető, et al. A foundation model to predict and capture human cognition. *Nature*, pages 1–8, 2025.
- [42] Jack Stanley, Emmett Rabot, Siva Reddy, Eugene Belilovsky, Laurent Mottron, and Danilo Bzdok. Large language models deconstruct the clinical intuition behind diagnosing autism. *Cell*, 188(8):2235–2248, 2025.
- [43] Marcin P Joachimiak, J Harry Caufield, Nomi L Harris, Hyeongsik Kim, and Christopher J Mungall. Gene set summarization using large language models. *ArXiv*, pages arXiv–2305, 2024.
- [44] Mengzhou Hu, Sahar Alkhairy, Ingoo Lee, Rudolf T Pillich, Dylan Fong, Kevin Smith, Robin Bachelder, Trey Ideker, and Dexter Pratt. Evaluation of large language models for discovery of gene set function. *Nature methods*, 22(1):82–91, 2025.
- [45] Xiaoliang Luo, Akilles Rechartdt, Guangzhi Sun, Kevin K Nejad, Felipe Yáñez, Bati Yilmaz, Kangjoo Lee, Alexandra O Cohen, Valentina Borghesani, Anton Pashkov, et al. Large language models surpass human experts in predicting neuroscience results. *Nature human behaviour*, 9(2):305–315, 2025.
- [46] Danilo Bzdok, Andrew Thieme, Oleksiy Levkovskyy, Paul Wren, Thomas Ray, and Siva Reddy. Data science opportunities of large language models for neuroscience and biomedicine. *Neuron*, 112(5):698–717, 2024.
- [47] Justine Y Hansen, Golia Shafiei, Katharina Voigt, Emma X Liang, Sylvia ML Cox, Marco Leyton, Sharna D Jamadar, and Bratislav Misic. Integrating multimodal and multiscale connectivity blueprints of the human cerebral cortex in health and disease. *PLoS biology*, 21(9):e3002314, 2023.
- [48] Kieran CR Fox, Lin Shi, Sori Baek, Omri Raccach, Brett L Foster, Srijani Saha, Daniel S Margulies, Aaron Kucyi, and Josef Parvizi. Intrinsic network architecture predicts the effects elicited by intracranial electrical stimulation of the human brain. *Nature human behaviour*, 4(10): 1039–1052, 2020.
- [49] Shan H Siddiqi, Frederic LWVJ Schaper, Andreas Horn, Joey Hsu, Jaya L Padmanabhan, Amy Brodtmann, Robin FH Cash, Maurizio Corbetta, Ki Sueng Choi,

- Darin D Dougherty, et al. Brain stimulation and brain lesions converge on common causal circuits in neuropsychiatric disease. *Nature human behaviour*, 5(12):1707–1716, 2021.
- [50] Andrew R Pines, Summer B Frandsen, William Drew, Garance M Meyer, Calvin Howard, Stephan T Palm, Frederic LWVJ Schaper, Christopher Lin, Konstantin Butenko, Michael A Ferguson, et al. Mapping lesions that cause psychosis to a human brain circuit and proposed stimulation target. *JAMA psychiatry*, 82(4):368–378, 2025.
- [51] Sean Froudish-Walsh, Ting Xu, Meiqi Niu, Lucija Rapan, Ling Zhao, Daniel S Margulies, Karl Zilles, Xiao-Jing Wang, and Nicola Palomero-Gallagher. Gradients of neurotransmitter receptor expression in the macaque cortex. *Nature neuroscience*, 26(7):1281–1294, 2023.
- [52] Ao Chen, Yidi Sun, Ying Lei, Chao Li, Sha Liao, Juan Meng, Yiqin Bai, Zhen Liu, Zhifeng Liang, Zhiyong Zhu, et al. Single-cell spatial transcriptome reveals cell-type organization in the macaque cortex. *Cell*, 186(17):3726–3743, 2023.
- [53] Andrea I Luppi, Zhen-Qi Liu, Justine Y Hansen, Rodrigo Cofre, Meiqi Niu, Elena Kuzmin, Seán Froudish-Walsh, Nicola Palomero-Gallagher, and Bratislav Misic. Benchmarking macaque brain gene expression for horizontal and vertical translation. *Science Advances*, 11(9):eads6967, 2025.
- [54] Ludovico Coletta, Marco Pagani, Jennifer D. Whitesell, Julie A. Harris, Boris Bernhardt, and Alessandro Gozzi. Network structure of the mouse brain connectome with voxel resolution. *Science Advances*, 6(51):7187–7205, dec 1 2020. ISSN 23752548. doi: 10.1126/SCIADV.ABB7187/SUPPL_FILE/ABB7187_SM.PDF. URL <https://www.science.org/doi/full/10.1126/sciadv.abb7187>. PMID: 33355124 publisher: American Association for the Advancement of Science.
- [55] Ed S Lein, Michael J Hawrylycz, Nancy Ao, Mikael Ayres, Amy Bensinger, Amy Bernard, Andrew F Boe, Mark S Boguski, Kevin S Brockway, Emi J Byrnes, et al. Genome-wide atlas of gene expression in the adult mouse brain. *Nature*, 445(7124):168–176, 2007.
- [56] Daniel Gutierrez-Barragan, Neha Atulkumar Singh, Filomena Grazia Alvino, Ludovico Coletta, Federico Rocchi, Elizabeth De Guzman, Alberto Galbusera, Mauro Ubaldi, Stefano Panzeri, and Alessandro Gozzi. Unique spatiotemporal fMRI dynamics in the awake mouse brain. *Current biology*, 32(2):1–14, 1 2021. ISSN 1879-0445. doi:10.1016/J.CUB.2021.12.015. URL <https://pubmed.ncbi.nlm.nih.gov/34998465/>. PMID: 34998465 publisher: Curr Biol.
- [57] Ben D Fulcher, John D Murray, Valerio Zerbi, and Xiao-Jing Wang. Multimodal gradients across mouse cortex. *Proc Natl Acad Sci USA*, 116(10):4689–4695, 2019.
- [58] Rahul S Desikan, Florent Ségonne, Bruce Fischl, Brian T Quinn, Bradford C Dickerson, Deborah Blacker, Randy L Buckner, Anders M Dale, R Paul Maguire, Bradley T Hyman, et al. An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest. *NeuroImage*, 31(3):968–980, 2006.
- [59] Elizabeth Beam, Christopher Potts, Russell A Poldrack, and Amit Etkin. A data-driven framework for mapping domains of human neurobiology. *Nature neuroscience*, 24(12):1733–1744, 2021.
- [60] Rolf Kötter and Egon Wanke. Mapping brains without coordinates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1456):751–766, 2005.
- [61] Quanxin Wang, Song-Lin Ding, Yang Li, Josh Royall, David Feng, Phil Lesnar, Nile Graddis, Maitham Naeemi, Benjamin Facer, Anh Ho, et al. The allen mouse brain common coordinate framework: a 3d reference atlas. *Cell*, 181(4):936–953, 2020.
- [62] Pieter R Roelfsema and Stefan Treue. Basic neuroscience research with nonhuman primates: a small but indispensable component of biomedical research. *Neuron*, 82(6):1200–1204, 2014.
- [63] Michael P Milham, Lei Ai, Bonhwang Koo, Ting Xu, Céline Amiez, Fabien Balezeau, Mark G Baxter, Erwin LA Blezer, Thomas Brochier, Aihua Chen, et al. An open resource for non-human primate imaging. *Neuron*, 100(1):61–74, 2018.
- [64] Helen C Barron, Rogier B Mars, David Dupret, Jason P Lerch, and Cassandra Sampaio-Baptista. Cross-species neuroscience: closing the explanatory gap. *Philosophical Transactions of the Royal Society B*, 376(1815):20190633, 2021.
- [65] Michael Ortiz-Rios, Paweł Kuśmieriek, Iain DeWitt, Dennis Archakov, Frederico AC Azevedo, Mikko Sams, Iiro P Jääskeläinen, Georgios A Keliris, and Josef P Rauschecker. Functional mri of the vocalization-processing network in the macaque brain. *Frontiers in neuroscience*, 9:113, 2015.
- [66] Christopher I Petkov, Christoph Kayser, Thomas Steudel, Kevin Whittingstall, Mark Augath, and Nikos K Logothetis. A voice region in the monkey brain. *Nature neuroscience*, 11(3):367–374, 2008.
- [67] Benjamin Wilson, Yukiko Kikuchi, Li Sun, David Hunter, Frederic Dick, Kenny Smith, Alexander Thiele, Timothy D Griffiths, William D Marslen-Wilson, and Christopher I Petkov. Auditory sequence processing reveals evolutionarily conserved regions of frontal cortex in macaques and humans. *Nature communications*, 6(1):8901, 2015.
- [68] Daniel S Margulies, Satrajit S Ghosh, Alexandros Goulas, Marcel Falkiewicz, Julia M Huntenburg, Georg Langs, Gleb Bezgin, Simon B Eickhoff, F Xavier Castellanos, Michael Petrides, Elizabeth Jefferies, and Jonathan Smallwood. Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proceedings of the National Academy of Sciences of the United States of America*, 113(44):12574–12579, 2016. ISSN 1091-6490. doi:10.1073/pnas.1608282113. PMID: 27791099 arXiv: 1408.1149 Citation Key: Margulies2016 ISBN: 0027-8424.
- [69] Andrea I. Luppi, Pedro A. M. Mediano, Fernando E. Rosas, Negin Holland, Tim D. Fryer, John T. O'Brien, James B. Rowe, David K. Menon, Daniel Bor, and Emmanuel A. Stamatakis. A synergistic core for human brain evolution and cognition. *Nature Neuroscience*, 25(6):771–782, 6 2022. ISSN 1546-1726. doi: 10.1038/s41593-022-01070-0. number: 6 publisher: Nature Publishing Group.
- [70] Russell A Poldrack. Inferring mental states from neuroimaging data: from reverse inference to large-scale decoding. *Neuron*, 72(5):692–697, 2011.
- [71] Timothy N Rubin, Oluwasanmi Koyejo, Krzysztof J Gorgolewski, Michael N Jones, Russell A Poldrack, and Tal Yarkoni. Decoding brain activity using a large-scale

- probabilistic functional-anatomical atlas of human cognition. *PLoS computational biology*, 13(10):e1005649, 2017.
- [72] Matthew L Leavitt, Diego Mendoza-Halliday, and Julio C Martinez-Trujillo. Sustained activity encoding working memories: not fully distributed. *Trends in neurosciences*, 40(6):328–346, 2017.
- [73] Gleb Bezgin, Ana Solodkin, Rembrandt Bakker, Petra Ritter, and Anthony R. McIntosh. Mapping complementary features of cross-species structural connectivity to construct realistic “virtual brains”. *Human Brain Mapping*, 38(4):2080–2093, 2017. ISSN 1097-0193. doi: 10.1002/hbm.23506.
- [74] Panagiotis Fotiadis, Linden Parkes, Kathryn A Davis, Theodore D Satterthwaite, Russell T Shinohara, and Dani S Bassett. Structure–function coupling in macroscale human brain networks. *Nature Reviews Neuroscience*, 25(10):688–704, 2024.
- [75] Steven E. Petersen and Olaf Sporns. Brain networks and cognitive architectures. *Neuron*, 88(1):207–219, 10 2015. ISSN 10974199. doi: 10.1016/j.neuron.2015.09.027. PMID: 26447582 publisher: Cell Press.
- [76] Vincent Bazinet, Justine Y Hansen, and Bratislav Misic. Towards a biologically annotated brain connectome. *Nature reviews neuroscience*, 24(12):747–760, 2023.
- [77] Andrea Avena-Koenigsberger, Bratislav Misic, and Olaf Sporns. Communication dynamics in complex brain networks. *Nature Reviews Neuroscience*, 19(1):17–33, 2017. ISSN 1471-003X. doi:10.1038/nrn.2017.149. PMID: 29238085 publisher: Nature Publishing Group Citation Key: Avena-Koenigsberger2017.
- [78] Caio Seguin, Olaf Sporns, and Andrew Zalesky. Brain network communication: concepts, models and applications. *Nature Reviews Neuroscience*, pages 1–18, jul 12 2023. ISSN 1471-0048. doi:10.1038/s41583-023-00718-5. URL <https://www.nature.com/articles/s41583-023-00718-5>. publisher: Nature Publishing Group.
- [79] Michael D Fox and Marcus E Raichle. Spontaneous fluctuations in brain activity observed with functional magnetic resonance imaging. *Nature reviews neuroscience*, 8(9):700–711, 2007.
- [80] Golia Shafiei, Sylvain Baillet, and Bratislav Misic. Human electromagnetic and haemodynamic networks systematically converge in unimodal cortex and diverge in transmodal cortex. *PLoS biology*, 20(8):e3001735, 2022.
- [81] Catherine Jean Chu, Naoaki Tanaka, J Diaz, Brian L Edlow, Ona Wu, M Hämäläinen, S Stufflebeam, Sydney S Cash, and Mark A Kramer. Eeg functional connectivity is partially predicted by underlying white matter connectivity. *Neuroimage*, 108:23–33, 2015.
- [82] Matthew J Brookes, Joanne R Hale, Johanna M Zumer, Claire M Stevenson, Susan T Francis, Gareth R Barnes, Julia P Owen, Peter G Morris, and Srikanth N Nagarajan. Measuring functional connectivity using meg: methodology and comparison with fmri. *Neuroimage*, 56(3):1082–1104, 2011.
- [83] Pilar Garces, Ernesto Pereda, Juan A Hernández-Tamames, Francisco Del-Pozo, Fernando Maestu, and José Ángel Pineda-Pardo. Multimodal description of whole brain connectivity: A comparison of resting state meg, fmri, and dwi. *Human brain mapping*, 37(1):20–34, 2016.
- [84] Sharna D Jamadar, Phillip GD Ward, Emma X Liang, Edwina R Orchard, Zhaolin Chen, and Gary F Egan. Metabolic and hemodynamic resting-state connectivity of the human brain: a high-temporal resolution simultaneous bold-fmri and fdg-fpet multimodality study. *Cerebral Cortex*, 31(6):2855–2867, 2021.
- [85] Stephen M Smith, Peter T Fox, Karla L Miller, David C Glahn, Peter Mickle Fox, Clare E Mackay, Nicola Filippini, Kate E Watkins, Roberto Toro, Angela R Laird, and Christian F Beckmann. Correspondence of the brain’s functional architecture during activation and rest. *Proceedings of the National Academy of Sciences*, 106(31):13040–13045, 2009. ISSN 0027-8424. doi: 10.1073/pnas.0905267106. PMID: 19620724 arXiv: 0905267106 Citation Key: Smith2009b ISBN: 1091-6490 (Electronic)\r1091-6490 (Linking).
- [86] Justine Y Hansen, Golia Shafiei, Ross D Markello, Sylvia Cox, Kelly Smart, Etienne Aumont, Stijn Servaes, Stephanie Scala, Gabriel Wainstein, Gleb Bezgin, Thomas Funck, W Schmitz, Marc-andre Bedard, R Nathan Spreng, Jean-paul Soucy, and Synthia Guimond. Mapping neurotransmitter systems to the structural and functional organization of the human neocortex. *Nature Neuroscience*, (accepted):1–26, 2022.
- [87] Michael J Hawrylycz, Ed S Lein, Angela L Guillozet-Bongaarts, Elaine H Shen, Lydia Ng, Jeremy A Miller, Louie N Van De Lagemaat, Kimberly A Smith, Amanda Ebbert, Zackery L Riley, et al. An anatomically comprehensive atlas of the adult human brain transcriptome. *Nature*, 489(7416):391, 2012.
- [88] Jonas Richiardi, Andre Altmann, Anna-Clare Milazzo, Catie Chang, M Mallar Chakravarty, Tobias Banaschewski, Gareth J Barker, Arun LW Bokde, Uli Bromberg, Christian Büchel, et al. Correlated gene expression supports synchronous activity in brain networks. *Science*, 348(6240):1241–1244, 2015.
- [89] Olaf Sporns, Giulio Tononi, and Rolf Kötter. The human connectome: a structural description of the human brain. *PLoS computational biology*, 1(4):e42, 2005.
- [90] Casey Paquola, Reinder Vos De Wael, Konrad Wagstyl, Richard AI Bethlehem, Seok-Jun Hong, Jakob Seidlitz, Edward T Bullmore, Alan C Evans, Bratislav Misic, Daniel S Margulies, et al. Microstructural and functional gradients are increasingly dissociated in transmodal cortices. *PLoS biology*, 17(5):e3000284, 2019.
- [91] Razia Azen and David V Budesu. The dominance analysis approach for comparing predictors in multiple regression. *Psychological methods*, 8(2):129, 2003.
- [92] Rembrandt Bakker, Thomas Wachtler, and Markus Diesmann. Cocomac 2.0 and the future of tract-tracing databases. *Frontiers in Neuroinformatics*, 6, dec 27 2012. ISSN 1662-5196. doi:10.3389/fninf.2012.00030. URL <https://www.frontiersin.org/articles/10.3389/fninf.2012.00030>. publisher: Frontiers.
- [93] Nicolas A. Crossley, Andrea Mechelli, Petra E. Vértes, Toby T. Winton-Brown, Ameera X. Patel, Cedric E. Ginestet, Philip McGuire, and Edward T. Bullmore. Cognitive relevance of the community structure of the human brain functional coactivation network. *Proceedings of the National Academy of Sciences*, 110(28):11583–11588, 2013. doi:10.1073/pnas.1220826110.
- [94] Roberto Toro, Peter T Fox, and Tomáš Paus. Functional coactivation map of the human brain. *Cerebral cortex*,

- 18(11):2553–2559, 2008.
- [95] Julia M Huntenburg, Pierre-Louis Bazin, and Daniel S Margulies. Large-scale gradients in human cortical organization. *Trends Cogn Sci*, 22(1):21–31, 2018.
- [96] Joshua B Burt, Murat Demirtaş, William J Eckner, Natasha M Navejar, Jie Lisa Ji, William J Martin, Alberto Bernacchia, Alan Anticevic, and John D Murray. Hierarchy of transcriptomic specialization across human cortex captured by structural neuroimaging topography. *Nature neuroscience*, 21(9):1251–1259, 2018.
- [97] Reinder Vos de Wael, Oualid Benkarim, Casey Paquola, Sara Lariviere, Jessica Royer, Shahin Tavakol, Ting Xu, Seok-Jun Hong, Georg Langs, Sofie Valk, et al. Brainspace: a toolbox for the analysis of macroscale gradients in neuroimaging and connectomics datasets. *Communications biology*, 3(1):103, 2020.
- [98] Yashar Zeighami, Trygve E Bakken, Thomas Nickl-Jockschat, Zeru Peterson, Anil G Jegga, Jeremy A Miller, Jay Schulkin, Alan C Evans, Ed S Lein, and Michael Hawrylycz. A comparison of anatomic and cellular transcriptome structures across 40 human brain diseases. *Plos Biology*, 21(4):e3002058, 2023.
- [99] Helen S. Mayberg, Andres M. Lozano, Valerie Voon, Heather E. McNeely, David Seminowicz, Clement Hamani, Jill M. Schwab, and Sidney H. Kennedy. Deep brain stimulation for treatment-resistant depression. *Neuron*, 45(5):651–660, 2005. doi: 10.1016/j.neuron.2005.02.014.
- [100] David M Howard, Mark J Adams, Toni-Kim Clarke, Jonathan D Hafferty, Jude Gibson, Masoud Shirali, Jonathan RI Coleman, Saskia P Hagenaars, Joey Ward, Eleanor M Wigmore, et al. Genome-wide meta-analysis of depression identifies 102 independent variants and highlights the importance of the prefrontal brain regions. *Nature neuroscience*, 22(3):343–352, 2019.
- [101] David C. Glahn, J. Daniel Ragland, Alon Abramoff, Jeffrey Barrett, Angela R. Laird, Carrie E. Bearden, and Dawn I. Velligan. Beyond hypofrontality: a quantitative meta-analysis of functional neuroimaging studies of working memory in schizophrenia. *Human Brain Mapping*, 25(1):60–69, 2005. doi:10.1002/hbm.20138.
- [102] World Health Organization. *International Classification of Diseases for Mortality and Morbidity Statistics (11th Revision)*. World Health Organization, Geneva, 2019. URL <https://icd.who.int/>. Accessed: 2025-11-06.
- [103] Kenneth S Kendler. The development of kraepelin’s concept of dementia praecox: A close reading of relevant texts. *JAMA psychiatry*, 77(11):1181–1187, 2020.
- [104] Daniel Freeman, Bryony Sheaves, Felicity Waite, Allison G Harvey, and Paul J Harrison. Sleep disturbance and psychiatric disorders. *The Lancet Psychiatry*, 7(7): 628–637, 2020.
- [105] Philip R Jansen, Kyoko Watanabe, Sven Stringer, Nathan Skene, Julien Bryois, Anke R Hammerschlag, Christiaan A de Leeuw, Jeroen S Benjamins, Ana B Muñoz-Manchado, Mats Nagel, et al. Genome-wide analysis of insomnia in 1,331,010 individuals identifies new risk loci and functional pathways. *Nature genetics*, 51(3): 394–403, 2019.
- [106] Yi Fang, Xiaoli Si, Jiali Wang, Zhiyun Wang, Ying Chen, Yi Liu, Yaping Yan, Jun Tian, Baorong Zhang, and Jiali Pu. Alzheimer disease and epilepsy: A mendelian randomization study. *Neurology*, 101(4):e399–e409, 2023.
- [107] Anita Kamondi, Madeleine Grigg-Damberger, Wolfgang Löscher, Heikki Tanila, and Andras Attila Horvath. Epilepsy and epileptiform activity in late-onset alzheimer disease: clinical and pathophysiological advances, gaps and conundrums. *Nature Reviews Neurology*, 20(3):162–182, 2024.
- [108] Aaron D Boes, Sashank Prasad, Hesheng Liu, Qi Liu, Alvaro Pascual-Leone, Verne S Caviness Jr, and Michael D Fox. Network localization of neurological symptoms from focal brain lesions. *Brain*, 138(10):3061–3075, 2015.
- [109] Andrea I. Luppi, S. Parker Singleton, Justine Y. Hansen, Danilo Bzdok, Amy Kuceyeski, Richard F. Betzel, and Bratislav Misic. Contributions of network structure, chemoarchitecture and diagnostic categories to transitions between cognitive topographies. *Nature Biomedical Engineering*, 8:1142–1161, 2024. doi: 10.1101/2023.03.16.532981. PMID: 36993597 PMCID: PMC10055141.
- [110] Justine Y. Hansen, Ross D. Markello, Jacob W. Vogel, Jakob Seidlitz, Danilo Bzdok, and Bratislav Misic. Mapping gene transcription and neurocognition across human neocortex. *Nature Human Behaviour*, pages 1–11, 3 2021. ISSN 2397-3374. doi:10.1038/s41562-021-01082-z. publisher: Nature Publishing Group.
- [111] Andrea I Luppi, Lynn Uhrig, Jordy Tasserie, Golia Shafiei, Kanako Muta, Junichi Hata, Hideyuki Okano, Daniel Golkowski, Andreas Ranft, Rudiger Ilg, et al. Comprehensive profiling of anaesthetised brain dynamics across phylogeny. *bioRxiv*, pages 2025–03, 2025.
- [112] Shi Gu, Fabio Pasqualetti, Matthew Cieslak, Qawi K Telesford, Alfred B Yu, Ari E Kahn, John D Medaglia, Jean M Vettel, Michael B Miller, Scott T Grafton, and Danielle S Bassett. Controllability of structural brain networks. *Nature Communications*, 6:8414, 2015. ISSN 20411723. doi:10.1038/ncomms9414. PMID: 26423222 Citation Key: Gu2015.
- [113] Jason Z Kim, Jonathan M. Soffer, Ari E Kahn, Jean M Vettel, Fabio Pasqualetti, and Danielle S Bassett. Role of graph architecture in controlling dynamical networks with applications to neural systems. *Nature Physics*, 14(1):91–98, 2018. ISSN 17452481. doi: 10.1038/NPHYS4268.
- [114] Matthew F Glasser, Timothy S Coalson, Emma C Robinson, Carl D Hacker, John Harwell, Essa Yacoub, Kamil Ugurbil, Jesper Andersson, Christian F Beckmann, Mark Jenkinson, et al. A multi-modal parcellation of human cerebral cortex. *Nature*, 536(7615):171–178, 2016.
- [115] Leila Cammoun, Xavier Gigandet, Djalel Meskaldji, Jean Philippe Thiran, Olaf Sporns, Kim Q Do, Philippe Maeder, Reto Meuli, and Patric Hagmann. Mapping the human connectome at multiple scales with diffusion spectrum mri. *J Neurosci Meth*, 203(2):386–397, 2012.
- [116] Ninad Naik. Probabilistic consensus through ensemble validation: A framework for llm reliability. *arXiv preprint arXiv:2411.06535*, 2024.
- [117] Dongfu Jiang, Xiang Ren, and Bill Yuchen Lin. Llm-blender: Ensembling large language models with pairwise ranking and generative fusion. *arXiv preprint arXiv:2306.02561*, 2023.
- [118] Seyed Pouyan Mousavi Davoudi, Alireza Shafiee Fard, and Alireza Amiri-Margavi. Collective reasoning among llms a framework for answer validation without ground truth. *arXiv preprint arXiv:2502.20758*, 2025.
- [119] Jessica B Voytek and Bradley Voytek. Automated cog-

- nome construction and semi-automated hypothesis generation. *Journal of neuroscience methods*, 208(1):92–100, 2012.
- [120] Juraj Gottweis, Wei-Hung Weng, Alexander Daryin, Tao Tu, Anil Palepu, Petar Sirkovic, Artiom Myaskovsky, Felix Weissenberger, Keran Rong, Ryutaro Tanno, et al. Towards an ai co-scientist. *arXiv preprint arXiv:2502.18864*, 2025.
- [121] Hao Yu, Aoran Gan, Kai Zhang, Shiwei Tong, Qi Liu, and Zhao Feng Liu. Evaluation of retrieval-augmented generation: A survey. In *CCF Conference on Big Data*, pages 102–120. Springer, 2024.
- [122] Guohao Li, Hasan Hammoud, Hani Itani, Dmitrii Khizbullin, and Bernard Ghanem. Camel: Communicative agents for "mind" exploration of large language model society. *Advances in Neural Information Processing Systems*, 36:51991–52008, 2023.
- [123] Noah Shinn, Federico Cassano, Ashwin Gopinath, Karthik Narasimhan, and Shunyu Yao. Reflexion: Language agents with verbal reinforcement learning. *Advances in Neural Information Processing Systems*, 36: 8634–8652, 2023.
- [124] V Nair, E Schumacher, G Tso, and A Kannan. Dera: enhancing large language model completions with dialog-enabled resolving agents. *arxiv. Preprint*, 2023.
- [125] Jiayi Yuan, Hao Li, Xinheng Ding, Wenya Xie, Yu-Jhe Li, Wentian Zhao, Kun Wan, Jing Shi, Xia Hu, and Zirui Liu. Understanding and mitigating numerical sources of non-determinism in llm inference. In *The Thirty-ninth Annual Conference on Neural Information Processing Systems*.
- [126] Thomas G. Dietterich. Ensemble methods in machine learning. *Multiple Classifier Systems*, pages 1–15, 2000. doi:10.1007/3-540-45014-9_1.
- [127] David V. Budescu and Eva Chen. Identifying expertise to extract the wisdom of crowds. *Management Science*, 61 (2):267–280, 2015. doi:10.1287/mnsc.2014.1909.
- [128] Petilla terminology: nomenclature of features of gabaergic interneurons of the cerebral cortex. *Nature Reviews Neuroscience*, 9(7):557–568, 2008.
- [129] Timothy EJ Behrens, Peter Fox, Angie Laird, and Stephen M Smith. What is the most interesting part of the brain? *Trends in cognitive sciences*, 17(1):2–4, 2013.
- [130] Valentina Pacella, Victor Nozais, Lia Talozzi, Majd Abdallah, Demian Wassermann, Stephanie J Forkel, and Michel Thiebaut de Schotten. The morphospace of the brain-cognition organisation. *Nature Communications*, 15(1):8452, 2024.
- [131] Krzysztof J Gorgolewski, Gael Varoquaux, Gabriel Rivera, Yannick Schwarz, Satrajit S Ghosh, Camille Maumet, Vanessa V Sochat, Thomas E Nichols, Russell A Poldrack, Jean-Baptiste Poline, et al. Neurovault. org: a web-based repository for collecting and sharing unthresholded statistical maps of the human brain. *Front Neuroinform*, 9:8, 2015.
- [132] 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.
- [133] Justine Y Hansen, Golia Shafiei, Jacob W Vogel, Kelly Smart, Carrie E Bearden, Martine Hoogman, Barbara Franke, Daan Van Rooij, Jan Buitelaar, Carrie R McDonald, et al. Local molecular and global connectomic contributions to cross-disorder cortical abnormalities. *Nature communications*, 13(1):4682, 2022.
- [134] Daqiang Sun, Christopher R. K. Ching, Amy Lin, Jennifer K. Forsyth, Leila Kushan, Ariana Vajdi, Maria Jalbrzikowski, Laura Hansen, Julio E. Villalon-Reina, Xiaoping Qu, Rachel K. Jonas, Therese van Amelsvoort, Geor Bakker, Wendy R. Kates, Kevin M. Antshel, Wanda Fremont, Linda E. Campbell, Kathryn L. McCabe, Eileen Daly, Maria Gudbrandsen, Clodagh M. Murphy, Declan Murphy, Michael Craig, Jacob Vorstman, Ania Fiksinski, Sanne Koops, Kosha Ruparel, David R. Roalf, Raquel E. Gur, J. Eric Schmitt, Tony J. Simon, Naomi J. Goodrich-Hunsaker, Courtney A. Durdle, Anne S. Bassett, Eva W. C. Chow, Nancy J. Butcher, Fidel Vila-Rodriguez, Joanne Doherty, Adam Cunningham, Marianne B. M. van den Bree, David E. J. Linden, Hayley Moss, Michael J. Owen, Kieran C. Murphy, Donna M. McDonald-McGinn, Beverly Emanuel, Theo G. M. van Erp, Jessica A. Turner, Paul M. Thompson, and Carrie E. Bearden. Large-scale mapping of cortical alterations in 22q11.2 deletion syndrome: Convergence with idiopathic psychosis and effects of deletion size. *Molecular Psychiatry*, 25(8):1822–1834, 8 2020. ISSN 1476-5578. doi:10.1038/s41380-018-0078-5. PMID: 29895892 PMCID: PMC6292748.
- [135] Martine Hoogman, Ryan Muetzel, Joao P. Guimaraes, Elena Shumskaya, Maarten Mennes, Marcel P. Zwiers, Neda Jahanshad, Gustavo Sudre, Thomas Wolfers, Eric A. Earl, Juan Carlos Soliva Vila, Yolanda Vives-Gilbert, Sabin Khadka, Stephanie E. Novotny, Catharina A. Hartman, Dirk J. Heslenfeld, Lizanne J. S. Schwere, Sara Ambrosino, Bob Oranje, Patrick de Zeeuw, Tiffany M. Chaim-Avancini, Pedro G. P. Rosa, Marcus V. Zanetti, Charles B. Malpas, Gregor Kohls, Georg G. von Polier, Jochen Seitz, Joseph Biederman, Alysa E. Doyle, Anders M. Dale, Theo G. M. van Erp, Jeffery N. Epstein, Terry L. Jernigan, Ramona Baur-Streubel, Georg C. Ziegler, Kathrin C. Zierhut, Anouk Schranter, Marie F. Hovik, Astri J. Lundervold, Clare Kelly, Hazel McCarthy, Norbert Skokauskas, Ruth L. O’Gorman Tuura, Anna Calvo, Sara Lera-Miguel, Rosa Nicolau, Kaylita C. Chantiluke, Anastasia Christakou, Alasdair Vance, Mara Cercignani, Matt C. Gabel, Philip Asherson, Sarah Baumeister, Daniel Brandeis, Sarah Hohmann, Ivanei E. Bramati, Fernanda Tovar-Moll, Andreas J. Fallgatter, Bernd Kardatzki, Lena Schwarz, Anatoly Anikin, Alexandr Baranov, Tinatin Gogberashvili, Dmitry Kapilushniy, Anastasia Solovieva, Hanan El Marroun, Tonya White, Georgii Karkashadze, Leyla Namazova-Baranova, Thomas Ethofer, Paulo Mattos, Tobias Banaschewski, David Coghill, Kerstin J. Plessen, Jonna Kuntsi, Mitul A. Mehta, Yannis Paloyelis, Neil A. Harrison, Mark A. Bellgrove, Tim J. Silk, Ana I. Cubillo, Katya Rubia, Luisa Lazaro, Silvia Brem, Susanne Walitza, Thomas Frodl, Mariam Zentis, Francisco X. Castellanos, Yuliya N. Yoncheva, Jan Haavik, Liesbeth Reneman, Annette Conzelmann, Klaus-Peter Lesch, Paul Pauli, Andreas Reif, Leanne Tamm, Kerstin Konrad, Eileen Oberwelland Weiss, Geraldo F. Busatto, Mario R. Louza, Sarah Durston, Pieter J. Hoekstra, Jaap Oosterlaan, Michael C. Stevens, J. Antoni Ramos-Quiroga, Oscar Vilarroya, Damien A. Fair, Joel T. Nigg, Paul M. Thompson, Jan K. Buitelaar, Stephen V. Faraone, Philip Shaw, Henning Tiemeier, Janita Bralten, and Barbara Franke. Brain imaging of the cortex

- in adhd: A coordinated analysis of large-scale clinical and population-based samples. *The American Journal of Psychiatry*, 176(7):531–542, 7 2019. ISSN 1535-7228. doi:10.1176/appi.ajp.2019.18091033. PMID: 31014101 PMCID: PMC6879185.
- [136] Daan van Rooij, Evdokia Anagnostou, Celso Arango, Guillaume Auzias, Marlene Behrmann, Geraldo F. Busatto, Sara Calderoni, Eileen Daly, Christine Deruelle, Adriana Di Martino, Ilan Dinstein, Fabio Luis Souza Duran, Sarah Durston, Christine Ecker, Damien Fair, Jennifer Fedor, Jackie Fitzgerald, Christine M. Freitag, Louise Gallagher, Iliaria Gori, Shlomi Haar, Liesbeth Hoekstra, Neda Jahanshad, Maria Jalbrzikowski, Joost Janssen, Jason Lerch, Beatriz Luna, Mauricio Moller Martinho, Jane McGrath, Filippo Muratori, Clodagh M. Murphy, Declan G. M. Murphy, Kirsten O'Hearn, Bob Oranje, Mara Parellada, Alessandra Retico, Pedro Rosa, Katya Rubia, Devon Shook, Margot Taylor, Paul M. Thompson, Michela Tosetti, Gregory L. Wallace, Fengfeng Zhou, and Jan K. Buitelaar. Cortical and subcortical brain morphometry differences between patients with autism spectrum disorder and healthy individuals across the lifespan: Results from the enigma asd working group. *The American Journal of Psychiatry*, 175(4):359–369, 4 2018. ISSN 1535-7228. doi: 10.1176/appi.ajp.2017.17010100. PMID: 29145754 PMCID: PMC6546164.
- [137] Christopher D. Whelan, Andre Altmann, Juan A. Botía, Neda Jahanshad, Derrek P. Hibar, Julie Absil, Saud Alhusaini, Marina K. M. Alvim, Pia Auvinen, Emanuele Bartolini, Felipe P. G. Bergo, Tauana Bernardes, Karen Blackmon, Barbara Braga, Maria Eugenia Caligiuri, Anna Calvo, Sarah J. Carr, Jian Chen, Shuai Chen, Andrea Cherubini, Philippe David, Martin Domin, Sonya Foley, Wendy França, Gerrit Haaker, Dmitry Isaev, Simon S. Keller, Raviteja Kotikalapudi, Magdalena A. Kowalczyk, Ruben Kuzniecky, Soenke Langner, Matteo Lenge, Kelly M. Leyden, Min Liu, Richard Q. Loi, Pascal Martin, Mario Mascalchi, Marcia E. Morita, Jose C. Pariente, Raul Rodríguez-Cruces, Christian Rummel, Taavi Saavalainen, Mira K. Semmelroch, Mariasavina Severino, Rhys H. Thomas, Manuela Tondelli, Domenico Tortora, Anna Elisabetta Vaudano, Lucy Vivash, Felix von Podewils, Jan Wagner, Bernd Weber, Yi Yao, Clarissa L. Yasuda, Guohao Zhang, Nuria Bargalló, Benjamin Bender, Neda Bernasconi, Andrea Bernasconi, Boris C. Bernhardt, Ingmar Blümcke, Chad Carlson, Gianpiero L. Cavalleri, Fernando Cendes, Luis Concha, Norman Delanty, Chantal Depondt, Orrin Devinsky, Colin P. Doherty, Niels K. Focke, Antonio Gambardella, Renzo Guerrini, Khalid Hamandi, Graeme D. Jackson, Reetta Kälviäinen, Peter Kochunov, Patrick Kwan, Angelo Labate, Carrie R. McDonald, Stefano Meletti, Terence J. O'Brien, Sebastien Ourselin, Mark P. Richardson, Pasquale Striano, Thomas Thesen, Roland Wiest, Junsong Zhang, Anamaria Vezzani, Mina Ryten, Paul M. Thompson, and Sanjay M. Sisodiya. Structural brain abnormalities in the common epilepsies assessed in a worldwide enigma study. *Brain: A Journal of Neurology*, 141(2):391–408, 2 2018. ISSN 1460-2156. doi:10.1093/brain/awx341. PMID: 29365066 PMCID: PMC5837616.
- [138] L. Schmaal, D. P. Hibar, P. G. Sämann, G. B. Hall, B. T. Baune, N. Jahanshad, J. W. Cheung, T. G. M. van Erp, D. Bos, M. A. Ikram, M. W. Vernooij, W. J. Niessen, H. Tiemeier, A. Hofman, K. Wittfeld, H. J. Grabe, D. Janowitz, R. Bülow, M. Selonke, H. Völzke, D. Grotegerd, U. Dannlowski, V. Arolt, N. Opel, W. Heindel, H. Kugel, D. Hoehn, M. Czisch, B. Couvy-Duchesne, M. E. Rentería, L. T. Strike, M. J. Wright, N. T. Mills, G. I. de Zubicaray, K. L. McMahon, S. E. Medland, N. G. Martin, N. A. Gillespie, R. Goya-Maldonado, O. Gruber, B. Krämer, S. N. Hatton, J. Lagopoulos, I. B. Hickie, T. Frodl, A. Carballo, E. M. Frey, L. S. van Velzen, B. W. J. H. Penninx, M.-J. van Tol, N. J. van der Wee, C. G. Davey, B. J. Harrison, B. Mwangi, B. Cao, J. C. Soares, I. M. Veer, H. Walter, D. Schoepf, B. Zurowski, C. Konrad, E. Schramm, C. Normann, K. Schnell, M. D. Sacchet, I. H. Gotlib, G. M. MacQueen, B. R. Godlewska, T. Nickson, A. M. McIntosh, M. Pappmeyer, H. C. Whalley, J. Hall, J. E. Sussmann, M. Li, M. Walter, L. Aftanas, I. Brack, N. A. Bokhan, P. M. Thompson, and D. J. Veltman. Cortical abnormalities in adults and adolescents with major depression based on brain scans from 20 cohorts worldwide in the enigma major depressive disorder working group. *Molecular Psychiatry*, 22(6):900–909, 6 2017. ISSN 1476-5578. doi:10.1038/mp.2016.60. PMID: 27137745 PMCID: PMC5444023.
- [139] Premika S. W. Boedhoe, Lianne Schmaal, Yoshinari Abe, Pino Alonso, Stephanie H. Ameis, Alan Anticevic, Paul D. Arnold, Marcelo C. Batistuzzo, Francesco Benedetti, Jan C. Beucke, Irene Bollettini, Anushree Bose, Silvia Brem, Anna Calvo, Rosa Calvo, Yuqi Cheng, Kang Ik K. Cho, Valentina Ciullo, Sara Dallaspezia, Damiaan Denys, Jamie D. Feusner, Kate D. Fitzgerald, Jean-Paul Fouché, Egill A. Frideirsson, Patricia Gruner, Gregory L. Hanna, Derrek P. Hibar, Marcelo Q. Hoexter, Hao Hu, Chaim Huyser, Neda Jahanshad, Anthony James, Norbert Kathmann, Christian Kaufmann, Kathrin Koch, Jun Soo Kwon, Luisa Lazaro, Christine Lochner, Rachel Marsh, Ignacio Martínez-Zalacain, David Mataix-Cols, José M. Menchón, Luciano Minuzzi, Astrid Morer, Takashi Nakamae, Tomohiro Nakao, Janardhanan C. Narayanaswamy, Seiji Nishida, Erika Nurmi, Joseph O'Neill, John Piacentini, Fabrizio Piras, Federica Piras, Y. C. Janardhan Reddy, Tim J. Reess, Yuki Sakai, Joao R. Sato, H. Blair Simpson, Noam Soreni, Carles Soriano-Mas, Gianfranco Spalletta, Michael C. Stevens, Philip R. Szeszko, David F. Tolin, Guido A. van Wingen, Ganesan Venkatasubramanian, Susanne Walitza, Zhen Wang, Je-Yeon Yun, ENIGMA-OCD Working Group, Paul M. Thompson, Dan J. Stein, Odile A. van den Heuvel, and ENIGMA OCD Working Group. Cortical abnormalities associated with pediatric and adult obsessive-compulsive disorder: Findings from the enigma obsessive-compulsive disorder working group. *The American Journal of Psychiatry*, 175(5):453–462, 5 2018. ISSN 1535-7228. doi: 10.1176/appi.ajp.2017.17050485. PMID: 29377733 PMCID: PMC7106947.
- [140] Theo G. M. van Erp, Esther Walton, Derrek P. Hibar, Lianne Schmaal, Wenhao Jiang, David C. Glahn, Godfrey D. Pearlson, Nailin Yao, Masaki Fukunaga, Ryota Hashimoto, Naohiro Okada, Hidenaga Yamamori, Juan R. Bustillo, Vincent P. Clark, Ingrid Agartz, Bryon A. Mueller, Wiepke Cahn, Sonja M. C. de Zwarte, Hilleke E. Hulshoff Pol, René S. Kahn, Roel A. Ophoff, Neeltje E. M. van Haren, Ole A. Andreassen, Anders M. Dale, Nhat Trung Doan, Tiril P. Gurholt, Ce-

- cilie B. Hartberg, Unn K. Haukvik, Kjetil N. Jørgensen, Trine V. Lagerberg, Ingrid Melle, Lars T. Westlye, Oliver Gruber, Bernd Kraemer, Anja Richter, David Zilles, Vince D. Calhoun, Benedicto Crespo-Facorro, Roberto Roiz-Santiañez, Diana Tordesillas-Gutiérrez, Carmel Loughland, Vaughan J. Carr, Stanley Catts, Vanessa L. Copley, Janice M. Fullerton, Melissa J. Green, Frans A. Henskens, Assen Jablensky, Rhoshel K. Lenroot, Bryan J. Mowry, Patricia T. Michie, Christos Pantelis, Yann Quidé, Ulrich Schall, Rodney J. Scott, Murray J. Cairns, Marc Seal, Paul A. Tooney, Paul E. Rasser, Gavin Cooper, Cynthia Shannon Weickert, Thomas W. Weickert, Derek W. Morris, Elliot Hong, Peter Kochunov, Lauren M. Beard, Raquel E. Gur, Ruben C. Gur, Theodore D. Satterthwaite, Daniel H. Wolf, Aysenil Belger, Gregory G. Brown, Judith M. Ford, Fabio Macciardi, Daniel H. Mathalon, Daniel S. O'Leary, Steven G. Potkin, Adrian Preda, James Voyvodic, Kelvin O. Lim, Sarah McEwen, Fude Yang, Yunlong Tan, Shuping Tan, Zhiren Wang, Fengmei Fan, Jingxu Chen, Hong Xiang, Shiyu Tang, Hua Guo, Ping Wan, Dong Wei, Henry J. Bockholt, Stefan Ehrlich, Rick P. F. Wolthuisen, Margaret D. King, Jody M. Shoemaker, Scott R. Sponheim, Lieuwe De Haan, Laura Koenders, Marise W. Machielsen, Therese van Amelsvoort, Dick J. Veltman, Francesca Assogna, Nerisa Banaj, Pietro de Rossi, Mariangela Iorio, Fabrizio Piras, Gianfranco Spalletta, Peter J. McKenna, Edith Pomarol-Clotet, Raymond Salvador, Aiden Corvin, Gary Donohoe, Sinead Kelly, Christopher D. Whelan, Erin W. Dickie, David Rotenberg, Aristotle N. Voineskos, Simone Ciufolini, Joaquim Radua, Paola Dazzan, Robin Murray, Tiago Reis Marques, Andrew Simmons, Stefan Borgwardt, Laura Egloff, Fabienne Harrisberger, Anita Riecher-Rössler, Renata Smieskova, Kathryn I. Alpert, Lei Wang, Erik G. Jönsson, Sanne Koops, Iris E. C. Sommer, Alessandro Bertolino, Aurora Bonvino, Annabella Di Giorgio, Emma Neilson, Andrew R. Mayer, Julia M. Stephen, Jun Soo Kwon, Je-Yeon Yun, Dara M. Cannon, Colm McDonald, Irina Lebedeva, Alexander S. Tomyshev, Tolibjohn Akhadov, Vasily Kaleda, Helena Fatouros-Bergman, Lena Flyckt, Karolinska Schizophrenia Project, Geraldo F. Busatto, Pedro G. P. Rosa, Mauricio H. Serpa, Marcus V. Zanetti, Cyril Hoschl, Antonin Skoch, Filip Spaniel, David Tomecek, Saskia P. Hagenaars, Andrew M. McIntosh, Heather C. Whalley, Stephen M. Lawrie, Christian Knöchel, Viola Oertel-Knöchel, Michael Stäblein, Fleur M. Howells, Dan J. Stein, Henk S. Temmingh, Anne Uhlmann, Carlos Lopez-Jaramillo, Danai Dima, Agnes McMahon, Joshua I. Faskowitz, Boris A. Gutman, Neda Jahanshad, Paul M. Thompson, and Jessica A. Turner. Cortical brain abnormalities in 4474 individuals with schizophrenia and 5098 control subjects via the enhancing neuro imaging genetics through meta analysis (enigma) consortium. *Biological Psychiatry*, 84(9):644–654, 11 2018. ISSN 1873-2402. doi:10.1016/j.biopsych.2018.04.023. PMID: 29960671 PMCID: PMC6177304.
- [141] D. P. Hibar, L. T. Westlye, N. T. Doan, N. Jahanshad, J. W. Cheung, C. R. K. Ching, A. Versace, A. C. Bilderbeck, A. Uhlmann, B. Mwangi, B. Krämer, B. Overs, C. B. Hartberg, C. Abé, D. Dima, D. Grotegerd, E. Sprooten, E. Bøen, E. Jimenez, F. M. Howells, G. Delvecchio, H. Temmingh, J. Starke, J. R. C. Almeida, J. M. Goikolea, J. Houenou, L. M. Beard, L. Rauer, L. Abramovic, M. Bonnin, M. F. Ponteduro, M. Keil, M. M. Rive, N. Yao, N. Yalin, P. Najt, P. G. Rosa, R. Redlich, S. Trost, S. Hagenaars, S. C. Fears, S. Alonso-Lana, T. G. M. van Erp, T. Nickson, T. M. Chaim-Avancini, T. B. Meier, T. Elvsåshagen, U. K. Haukvik, W. H. Lee, A. H. Schene, A. J. Lloyd, A. H. Young, A. Nugent, A. M. Dale, A. Pfennig, A. M. McIntosh, B. Lafer, B. T. Baune, C. J. Ekman, C. A. Zarate, C. E. Bearden, C. Henry, C. Simhandl, C. McDonald, C. Bourne, D. J. Stein, D. H. Wolf, D. M. Cannon, D. C. Glahn, D. J. Veltman, E. Pomarol-Clotet, E. Vieta, E. J. Canales-Rodriguez, F. G. Nery, F. L. S. Duran, G. F. Busatto, G. Roberts, G. D. Pearlson, G. M. Goodwin, H. Kugel, H. C. Whalley, H. G. Ruhe, J. C. Soares, J. M. Fullerton, J. K. Rybakowski, J. Savitz, K. T. Chaim, M. Fatjó-Vilas, M. G. Soeiro-de Souza, M. P. Boks, M. V. Zanetti, M. C. G. Otaduy, M. S. Schaufelberger, M. Alda, M. Ingvar, M. L. Phillips, M. J. Kempton, M. Bauer, M. Landén, N. S. Lawrence, N. E. M. van Haren, N. R. Horn, N. B. Freimer, O. Gruber, P. R. Schofield, P. B. Mitchell, R. S. Kahn, R. Lenroot, R. Machado-Vieira, R. A. Ophoff, S. Sarró, S. Frangou, T. D. Satterthwaite, T. Hajek, U. Dannlowski, U. F. Malt, V. Arolt, W. F. Gattaz, W. C. Drevets, X. Caseras, I. Agartz, P. M. Thompson, and O. A. Andreassen. Cortical abnormalities in bipolar disorder: an mri analysis of 6503 individuals from the enigma bipolar disorder working group. *Molecular Psychiatry*, 23(4):932–942, 4 2018. ISSN 1476-5578. doi:10.1038/mp.2017.73. PMID: 28461699 PMCID: PMC5668195.
- [142] Max A. Laansma, Joanna K. Bright, Sarah Al-Bachari, Tim J. Anderson, Tyler Ard, Francesca Assogna, Katherine A. Baquero, Henk W. Berendse, Jamie Blair, Fernando Cendes, John C. Dalrymple-Alford, Rob M.A. de Bie, Ines Debove, Michiel F. Dirx, Jason Druzgal, Hedley C.A. Emsley, Gäetan Garraux, Rachel P. Guimarães, Boris A. Gutman, Rick C. Helmich, Johannes C. Klein, Clare E. Mackay, Corey T. McMillan, Tracy R. Melzer, Laura M. Parkes, Fabrizio Piras, Toni L. Pitcher, Kathleen L. Poston, Mario Rango, Letícia F. Ribeiro, Cristiane S. Rocha, Christian Rummel, Lucas S.R. Santos, Reinhold Schmidt, Petra Schwingenschuh, Gianfranco Spalletta, Letizia Squarcina, Odile A. van den Heuvel, Chris Vriend, Jiun-Jie Wang, Daniel Weintraub, Roland Wiest, Clarissa L. Yasuda, Neda Jahanshad, Paul M. Thompson, Ysbrand D. van der Werf, and The ENIGMA-Parkinson's Study. International multicenter analysis of brain structure across clinical stages of parkinson's disease. *Movement Disorders*, 36(11):2583–2594, 2021. ISSN 1531-8257. doi:10.1002/mds.28706. preprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1002/mds.28706>.
- [143] Ditte Demontis, Raymond K Walters, Joanna Martin, Manuel Mattheisen, Thomas D Als, Esben Agerbo, Gísli Baldursson, Rich Belliveau, Jonas Bybjerg-Grauholm, Marie Bækvad-Hansen, et al. Discovery of the first genome-wide significant risk loci for attention deficit/hyperactivity disorder. *Nature genetics*, 51(1): 63–75, 2019.
- [144] Jakob Grove, Stephan Ripke, Thomas D Als, Manuel Mattheisen, Raymond K Walters, Hyejung Won, Jonatan Pallesen, Esben Agerbo, Ole A Andreassen, Richard Anney, et al. Identification of common genetic risk variants for autism spectrum disorder. *Nature genetics*, 51(3): 431–444, 2019.

- [145] Eli A Stahl, Gerome Breen, Andreas J Forstner, Andrew McQuillin, Stephan Ripke, Vassily Trubetskoy, Manuel Mattheisen, Yunpeng Wang, Jonathan RI Coleman, H  l  na A Gaspar, et al. Genome-wide association study identifies 30 loci associated with bipolar disorder. *Nature genetics*, 51(5):793–803, 2019.
- [146] Antonio F Pardi  as, Peter Holmans, Andrew J Pocklington, Valentina Escott-Price, Stephan Ripke, Noa Carrera, Sophie E Legge, Sophie Bishop, Darren Cameron, Marian L Hamshire, et al. Common schizophrenia alleles are enriched in mutation-intolerant genes and in regions under strong background selection. *Nature genetics*, 50(3): 381–389, 2018.
- [147] Janet Pi  ero,   lex Bravo, N  ria Queralt-Rosinach, Alba Guti  rrez-Sacrist  n, Jordi Deu-Pons, Emilio Centeno, Javier Garc  a-Garc  a, Ferran Sanz, and Laura I Furlong. Disgenet: a comprehensive platform integrating information on human disease-associated genes and variants. *Nucleic acids research*, page gkw943, 2016.
- [148] Janet Pi  ero, Juan Manuel Ram  rez-Anguita, Josep Sa  ch-Pitarch, Francesco Ronzano, Emilio Centeno, Ferran Sanz, and Laura I Furlong. The disgenet knowledge platform for disease genomics: 2019 update. *Nucleic acids research*, 48(D1):D845–D855, 2020.
- [149] Kenneth K Kwong, John W Belliveau, David A Chesler, Inna E Goldberg, Robert M Weisskoff, Brigitte P Poncelet, David N Kennedy, Bernice E Hoppel, Mark S Cohen, and Robert Turner. Dynamic magnetic resonance imaging of human brain activity during primary sensory stimulation. *Proceedings of the National Academy of Sciences*, 89(12):5675–5679, 1992.
- [150] 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.
- [151] Matthew F Glasser, Stamati  s N Sotiropoulos, Anthony Wilson, Timothy S Coalson, Bruce Fischl, Jesper L Andersson, Junqian Xu, Saad Jbabdi, Matthew Webster, Jonathan R Polimeni, David C Van Essen, and Mark Jenkinson. The minimal preprocessing pipelines for the human connectome project. *Neuroimage*, 80:105–124, 2013. doi:10.1016/j.neuroimage.2013.04.127.
- [152] Susan Whitfield-Gabrieli and Alfonso Nieto-Castanon. Conn: A Functional Connectivity Toolbox for Correlated and Anticorrelated Brain Networks. *Brain Connectivity*, 2(3):125–141, 2012. ISSN 21580022. doi: 10.1089/brain.2012.0073. Citation Key: Whitfield-Gabrieli.
- [153] Behzadi Y, Restom K, Liau J, and Liu TT. A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *NeuroImage*, 37:90–101, 2007. [Online; accessed 2019-06-26].
- [154] Sharna D Jamadar, Phillip GD Ward, Emma X Liang, Edwina R Orchard, Zhaolin Chen, and Gary F Egan. Metabolic and hemodynamic resting-state connectivity of the human brain: a high-temporal resolution simultaneous bold-fMRI and fdg-fPET multimodality study. *Cerebral Cortex*, 31(6):2855–2867, 2021.
- [155] Katharina Voigt, Emma X Liang, Bratislav Misi  c, Phillip G D Ward, Gary F Egan, and Sharna D Jamadar. Metabolic and functional connectivity provide unique and complementary insights into cognition-connectome relationships. *Cerebral Cortex*, 2022.
- [156] Ninon Burgos, M Jorge Cardoso, Kris Thielemans, Marc Modat, Stefano Pedemonte, John Dickson, Anna Barnes, Rebekah Ahmed, Colin J Mahoney, Jonathan M Schott, et al. Attenuation correction synthesis for hybrid PET-MR scanners: application to brain studies. *IEEE transactions on medical imaging*, 33(12):2332–2341, 2014.
- [157] Mark Jenkinson, Peter Bannister, Michael Brady, and Stephen Smith. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage*, 17(2):825–841, 2002.
- [158] Sharna D Jamadar, Phillip GD Ward, Thomas G Close, Alex Fornito, Malin Premaratne, Kieran O’Brien, Daniel St  b, Zhaolin Chen, N Jon Shah, and Gary F Egan. Simultaneous bold-fMRI and constant infusion fdg-PET data of the resting human brain. *Scientific data*, 7(1):1–12, 2020.
- [159] Fran  ois Tadel, Sylvain Baillet, John C Mosher, Dimitrios Pantazis, and Richard M Leahy. Brainstorm: a user-friendly application for MEG/EEG analysis. *Computational intelligence and neuroscience*, 2011, 2011.
- [160] Giles L Colclough, Mark W Woolrich, PK Tewarie, Matthew J Brookes, Andrew J Quinn, and Stephen M Smith. How reliable are MEG resting-state connectivity metrics? *Neuroimage*, 138:284–293, 2016.
- [161] Ross D Markello, Aurina Arnatkeviciute, Jean-Baptiste Poline, Ben D Fulcher, Alex Fornito, and Bratislav Misi  c. Standardizing workflows in imaging transcriptomics with the abagen toolbox. *Elife*, 10:e72129, 2021.
- [162] Simon Kaller, Michael Rullmann, Marianne Patt, Georg Becker, Julia Luthardt, Johanna Girhardt, Philipp Meyer, Peter Werner, Henryk Barthel, Anke McLeod, Thomas Fritz, Swen Hesse, and Osama Sabri. Test-retest measurements of dopamine D1-type receptors using simultaneous PET/MRI imaging. *European Journal of Nuclear Medicine and Molecular Imaging*, 44, 6 2017. doi: 10.1007/s00259-017-3645-0.
- [163] Christine M. Sandiego, Jean-Dominique Gallezot, Keunpoong Lim, Jim Ropchan, Shu-fei Lin, Hong Gao, Evan D. Morris, and Kelly P. Cosgrove. Reference region modeling approaches for amphetamine challenge studies with [11C]flb 457 and PET. *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism*, 35(4):623–629, 3 2015. ISSN 1559-7016. doi:10.1038/jcbfm.2014.237. PMID: 25564239 PMCID: PMC4420880.
- [164] Mark Slifstein, Elsmarieke van de Giessen, Jared Van Snellenberg, Judy L. Thompson, Rajesh Narendran, Roberto Gil, Elizabeth Hackett, Ragy Girgis, Najate Ojeil, Holly Moore, Deepak D’Souza, Robert T. Malison, Yiyun Huang, Keunpoong Lim, Nabeel Nabulsi, Richard E. Carson, Jeffrey A. Lieberman, and Anissa Abi-Dargham. Deficits in prefrontal cortical and extrastriatal dopamine release in schizophrenia: a positron emission tomographic functional magnetic resonance imaging study. *JAMA psychiatry*, 72(4):316–324, 4 2015. ISSN 2168-6238. doi:10.1001/jamapsychiatry.2014.2414. PMID: 25651194 PMCID: PMC4768742.
- [165] Christopher T. Smith, Jennifer L. Crawford, Linh C. Dang, Kendra L. Seaman, M. Danica San Juan, Aishwarya Vijay, Daniel T. Katz, David Matuskey, Ronald L. Cowan, Evan D. Morris, David H. Zald, and Gregory R. Samanez-Larkin. Partial-volume correction increases estimated dopamine D2-like receptor binding potential

- and reduces adult age differences. *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism*, 39(5):822–833, 5 2019. ISSN 1559-7016. doi:10.1177/0271678X17737693. PMID: 29090626 PMCID: PMC6498753.
- [166] Yasmin Zakiniaieiz, Ansel T. Hillmer, David Matuskey, Nabeel Nabulsi, Jim Ropchan, Carolyn M. Mazure, Marina R. Picciotto, Yiyun Huang, Sherry A. McKee, Evan D. Morris, and Kelly P. Cosgrove. Sex differences in amphetamine-induced dopamine release in the dorsolateral prefrontal cortex of tobacco smokers. *Neuropsychopharmacology: Official Publication of the American College of Neuropsychopharmacology*, 44(13):2205–2211, 12 2019. ISSN 1740-634X. doi:10.1038/s41386-019-0456-y. PMID: 31269510 PMCID: PMC6897943.
- [167] Juergen Dukart, Štefan Holiga, Christopher Chatham, Peter Hawkins, Anna Forsyth, Rebecca McMillan, Jim Myers, Anne R. Lingford-Hughes, David J. Nutt, Emilio Merlo-Pich, Celine Risterucci, Lauren Boak, Daniel Umbricht, Scott Schobel, Thomas Liu, Mitul A. Mehta, Fernando O. Zelaya, Steve C. Williams, Gregory Brown, Martin Paulus, Garry D. Honey, Suresh Muthukumaraswamy, Joerg Hipp, Alessandro Bertolino, and Fabio Sambataro. Cerebral blood flow predicts differential neurotransmitter activity. *Scientific Reports*, 8(1): 4074, 3 2018. ISSN 2045-2322. doi:10.1038/s41598-018-22444-0. PMID: 29511260 PMCID: PMC5840131.
- [168] Renata Belfort-DeAguiar, Jean-Dominique Gallezot, Janice J Hwang, Ahmed Elshafie, Catherine W Yeckel, Owen Chan, Richard E Carson, Yu-Shin Ding, and Robert S Sherwin. Noradrenergic activity in the human brain: A mechanism supporting the defense against hypoglycemia. *The Journal of Clinical Endocrinology and Metabolism*, 103(6):2244–2252, 6 2018. ISSN 0021-972X. doi:10.1210/jc.2017-02717.
- [169] Yu-Shin Ding, Tarun Singhal, Beata Planeta-Wilson, Jean-Dominique Gallezot, Nabeel Nabulsi, David Labaree, Jim Ropchan, Shannan Henry, Wendol Williams, Richard E. Carson, Alexander Neumeister, and Robert T. Malison. Pet imaging of the effects of age and cocaine on the norepinephrine transporter in the human brain using (s,s)-[(11)c]o-methylreboxetine and hrtr. *Synapse (New York, N.Y.)*, 64(1):30–38, 1 2010. ISSN 1098-2396. doi:10.1002/syn.20696. PMID: 19728366 PMCID: PMC3727644.
- [170] Chiang-shan R. Li, Marc N. Potenza, Dianne E. Lee, Beata Planeta, Jean-Dominique Gallezot, David Labaree, Shannan Henry, Nabeel Nabulsi, Rajita Sinha, Yu-Shin Ding, Richard E. Carson, and Alexander Neumeister. Decreased norepinephrine transporter availability in obesity: Positron emission tomography imaging with (s,s)-[(11)c]o-methylreboxetine. *NeuroImage*, 86:306–310, 2 2014. ISSN 1095-9572. doi: 10.1016/j.neuroimage.2013.10.004. PMID: 24121204 PMCID: PMC3947246.
- [171] Elizabeth Sanchez-Rangel, Jean-Dominique Gallezot, Catherine W. Yeckel, Wai Lam, Renata Belfort-DeAguiar, Ming-Kai Chen, Richard E. Carson, Robert Sherwin, and Janice J. Hwang. Norepinephrine transporter availability in brown fat is reduced in obesity: a human pet study with [11c] mrb. *International Journal of Obesity*, 44(4):964–967, 4 2020. ISSN 1476-5497. doi: 10.1038/s41366-019-0471-4. number: 4 publisher: Nature Publishing Group.
- [172] Markus Savli, Andreas Bauer, Markus Mitterhauser, Yu-Shin Ding, Andreas Hahn, Tina Kroll, Alexander Neumeister, Daniela Haeusler, Johanna Ungersboeck, Shannan Henry, Sanaz Attaripour Isfahani, Frank Rattay, Wolfgang Wadsak, Siegfried Kasper, and Rupert Lanzenberger. Normative database of the serotonergic system in healthy subjects using multi-tracer pet. *NeuroImage*, 63(1):447–459, 10 2012. ISSN 1053-8119. doi: 10.1016/j.neuroimage.2012.07.001.
- [173] Jean-Dominique Gallezot, Nabeel Nabulsi, Alexander Neumeister, Beata Planeta-Wilson, Wendol A. Williams, Tarun Singhal, Sunhee Kim, R. Paul Maguire, Timothy McCarthy, J. James Frost, Yiyun Huang, Yu-Shin Ding, and Richard E. Carson. Kinetic modeling of the serotonin 5-HT_{1B} receptor radioligand [(11)c]p943 in humans. *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism*, 30(1):196–210, 1 2010. ISSN 1559-7016. doi:10.1038/jcbfm.2009.195. PMID: 19773803 PMCID: PMC2949107.
- [174] David Matuskey, Zubin Bhagwagar, Beata Planeta, Brian Pittman, Jean-Dominique Gallezot, Jason Chen, Jane Wanyiri, Soheila Najafzadeh, Jim Ropchan, Paul Geha, Yiyun Huang, Marc N. Potenza, Alexander Neumeister, Richard E. Carson, and Robert T. Malison. Reductions in brain 5-HT_{1B} receptor availability in primarily cocaine-dependent humans. *Biological Psychiatry*, 76(10):816–822, 11 2014. ISSN 1873-2402. doi: 10.1016/j.biopsych.2013.11.022. PMID: 24433854 PMCID: PMC4037398.
- [175] James W. Murrough, Christoph Czermak, Shannan Henry, Nabeel Nabulsi, Jean-Dominique Gallezot, Ralitzia Gueorguieva, Beata Planeta-Wilson, John H. Krystal, John F. Neumaier, Yiyun Huang, Yu-Shin Ding, Richard E. Carson, and Alexander Neumeister. The effect of early trauma exposure on serotonin type 1b receptor expression revealed by reduced selective radioligand binding. *Archives of General Psychiatry*, 68(9):892–900, 9 2011. ISSN 1538-3636. doi: 10.1001/archgenpsychiatry.2011.91. PMID: 21893657 PMCID: PMC3244836.
- [176] Christopher Pittenger, Thomas G. Adams, Jean-Dominique Gallezot, Michael J. Crowley, Nabeel Nabulsi, null James Ropchan, Hong Gao, Stephen A. Kichuk, Ryan Simpson, Eileen Billingslea, Jonas Hannestad, Michael Bloch, Linda Mayes, Zubin Bhagwagar, and Richard E. Carson. Ocd is associated with an altered association between sensorimotor gating and cortical and subcortical 5-HT_{1B} receptor binding. *Journal of Affective Disorders*, 196:87–96, 5 2016. ISSN 1573-2517. doi: 10.1016/j.jad.2016.02.021. PMID: 26919057 PMCID: PMC4808438.
- [177] Aybala Saricicek, Jason Chen, Beata Planeta, Barbara Ruf, Kalyani Subramanyam, Kathleen Maloney, David Matuskey, David Labaree, Lorenz Deserno, Alexander Neumeister, John H. Krystal, Jean-Dominique Gallezot, Yiyun Huang, Richard E. Carson, and Zubin Bhagwagar. Test-retest reliability of the novel 5-HT_{1B} receptor pet radioligand [11c]p943. *European Journal of Nuclear Medicine and Molecular Imaging*, 42(3):468–477, 3 2015. ISSN 1619-7089. doi:10.1007/s00259-014-2958-5. PMID: 25427881.
- [178] Vincent Beliveau, Melanie Ganz, Ling Feng, Brice

- Ozenne, Liselotte Højgaard, Patrick M. Fisher, Claus Svarer, Douglas N. Greve, and Gitte M. Knudsen. A high-resolution in vivo atlas of the human brain's serotonin system. *The Journal of Neuroscience*, 37(1):120, 1 2017. doi:10.1523/JNEUROSCI.2830-16.2016. PMID: 28053035 publisher: Society for Neuroscience.
- [179] Rajiv Radhakrishnan, Nabeel Nabulsi, Edward Gaiser, Jean-Dominique Gallezot, Shannan Henry, Beata Planeta, Shu-fei Lin, Jim Ropchan, Wendol Williams, Evan Morris, Deepak Cyril D'Souza, Yiyun Huang, Richard E. Carson, and David Matuskey. Age-related change in 5-HT₆ receptor availability in healthy male volunteers measured with 11C-gsk215083 PET. *Journal of Nuclear Medicine*, 59(9):1445–1450, 9 2018. ISSN 0161-5505. doi:10.2967/jnumed.117.206516. PMID: 29626125 PMCID: PMC6126437.
- [180] Rajiv Radhakrishnan, David Matuskey, Nabeel Nabulsi, Edward Gaiser, Jean-Dominique Gallezot, Shannan Henry, Beata Planeta, Shu-Fei Lin, Jim Ropchan, Yiyun Huang, Richard E. Carson, and Deepak Cyril D'Souza. In vivo 5-HT₆ and 5-HT_{2A} receptor availability in antipsychotic treated schizophrenia patients vs. unmedicated healthy humans measured with [11C]gsk215083 PET. *Psychiatry Research: Neuroimaging*, 295:111007, 1 2020. ISSN 1872-7506. doi:10.1016/j.psychres.2019.111007. PMID: 31760336.
- [181] Stephen R. Baldassarri, Ansel T. Hillmer, Jon Mikael Anderson, Peter Jatlow, Nabeel Nabulsi, David Labaree, Kelly P. Cosgrove, Stephanie S. O'Malley, Thomas Eissenberg, Suchitra Krishnan-Sarin, and Irina Esterlis. Use of electronic cigarettes leads to significant beta₂-nicotinic acetylcholine receptor occupancy: Evidence from a PET imaging study. *Nicotine and Tobacco Research: Official Journal of the Society for Research on Nicotine and Tobacco*, 20(4):425–433, 3 2018. ISSN 1469-994X. doi:10.1093/ntr/ntx091. PMID: 28460123 PMCID: PMC5896427.
- [182] A. T. Hillmer, I. Esterlis, J. D. Gallezot, F. Bois, M. Q. Zheng, N. Nabulsi, S. F. Lin, R. L. Papke, Y. Huang, O. Sabri, R. E. Carson, and K. P. Cosgrove. Imaging of cerebral alpha₄beta₂* nicotinic acetylcholine receptors with (-)-[(18)F]flubatine PET: Implementation of bolus plus constant infusion and sensitivity to acetylcholine in human brain. *NeuroImage*, 141:71–80, 11 2016. ISSN 1095-9572. doi:10.1016/j.neuroimage.2016.07.026. PMID: 27426839 PMCID: PMC5026941.
- [183] Mika Naganawa, Nabeel Nabulsi, Shannan Henry, David Matuskey, Shu-Fei Lin, Lawrence Sliker, Adam J. Schwarz, Nancy Kant, Cynthia Jesudason, Kevin Rulley, Antonio Navarro, Hong Gao, Jim Ropchan, David Labaree, Richard E. Carson, and Yiyun Huang. First-in-human assessment of 11C-lsn3172176, an m₁ muscarinic acetylcholine receptor PET radiotracer. *Journal of Nuclear Medicine: Official Publication, Society of Nuclear Medicine*, 62(4):553–560, 4 2021. ISSN 1535-5667. doi:10.2967/jnumed.120.246967. PMID: 32859711 PMCID: PMC8049371.
- [184] M. Aghourian, C. Legault-Denis, J.-P. Soucy, P. Rosa-Neto, S. Gauthier, A. Kostikov, P. Gravel, and M.-A. Bédard. Quantification of brain cholinergic denervation in Alzheimer's disease using PET imaging with [18F]-feobv. *Molecular Psychiatry*, 22(11):1531–1538, 11 2017. ISSN 1476-5578. doi:10.1038/mp.2017.183. PMID: 28894304.
- [185] Marc-Andre Bedard, Meghmik Aghourian, Camille Legault-Denis, Ronald B. Postuma, Jean-Paul Soucy, Jean-François Gagnon, Amélie Pelletier, and Jacques Montplaisir. Brain cholinergic alterations in idiopathic REM sleep behaviour disorder: a PET imaging study with 18F-feobv. *Sleep Medicine*, 58:35–41, 6 2019. ISSN 1878-5506. doi:10.1016/j.sleep.2018.12.020. PMID: 31078078.
- [186] Jonathan M. DuBois, Olivier G. Rousset, Jared Rowley, Manuel Porras-Betancourt, Andrew J. Reader, Aurelie Labbe, Gassan Massarweh, Jean-Paul Soucy, Pedro Rosa-Neto, and Eliane Kobayashi. Characterization of age/sex and the regional distribution of mGluR5 availability in the healthy human brain measured by high-resolution [(11)C]abp688 PET. *European Journal of Nuclear Medicine and Molecular Imaging*, 43(1):152–162, 1 2016. ISSN 1619-7089. doi:10.1007/s00259-015-3167-6. PMID: 26290423.
- [187] Kelly Smart, Sylvia M. L. Cox, Stephanie G. Scala, Maria Tippler, Natalia Jaworska, Michel Boivin, Jean R. Séguin, Chawki Benkelfat, and Marco Leyton. Sex differences in [11C]abp688 binding: a positron emission tomography study of mGluR5 receptors. *European Journal of Nuclear Medicine and Molecular Imaging*, 46(5):1179–1183, 2019. ISSN 1619-7070. doi:10.1007/s00259-018-4252-4. PMID: 30627817 PMCID: PMC6451701.
- [188] Marian Galovic, Adam Al-Diwani, Umesh Vivekananda, Francisco Torrealdea, Kjell Erlandsson, Tim D. Fryer, Young T. Hong, Benjamin A. Thomas, Colm J. McGinnity, Evan Edmond, Kerstin Sander, Erik Årstad, Ilija Jelcic, Franklin I. Aigbirhio, Ashley M. Groves, Kris Thielemans, Brian Hutton, Alexander Hammers, John S. Duncan, Jonathan P. Coles, Anna Barnes, Charlotte J. Stagg, Matthew C. Walker, Sarosh R. Irani, Matthias J. Koepp, and for the NEST Investigators. In vivo NMDA receptor function in people with NMDA receptor antibody encephalitis. *medRxiv*, 12 2021. doi:10.1101/2021.12.04.21267226. URL <https://www.medrxiv.org/content/10.1101/2021.12.04.21267226v1>. page: 2021.12.04.21267226.
- [189] Marian Galovic, Kjell Erlandsson, Tim D. Fryer, Young T. Hong, Roido Manavaki, Hasan Sari, Sarah Chetcuti, Benjamin A. Thomas, Martin Fisher, Selena Sephton, Roberto Canales, Joseph J. Russell, Kerstin Sander, Erik Årstad, Franklin I. Aigbirhio, Ashley M. Groves, John S. Duncan, Kris Thielemans, Brian F. Hutton, Jonathan P. Coles, Matthias J. Koepp, and NEST investigators. Validation of a combined image derived input function and venous sampling approach for the quantification of [18F]GE-179 PET binding in the brain. *NeuroImage*, 237:118194, 8 2021. ISSN 1095-9572. doi:10.1016/j.neuroimage.2021.118194. PMID: 34023451.
- [190] Martin Nørgaard, Vincent Beliveau, Melanie Ganz, Claus Svarer, Lars Pinborg, Sune Keller, Peter Jensen, Douglas Greve, and Gitte Knudsen. A high-resolution in vivo atlas of the human brain's benzodiazepine binding site of GABA_A receptors. *NeuroImage*, 232:117878, 2021. doi:10.1101/2020.04.10.035352. ISBN: 4535456720.
- [191] Jean-Dominique Gallezot, Beata Planeta, Nabeel Nabulsi, Donna Palumbo, Xiaoxi Li, Jing Liu, Carolyn Rowinski, Kristin Chidsey, David Labaree, Jim Ropchan, Shu-Fei Lin, Aarti Sawant-Basak, Timothy J. McCarthy, Anne W. Schmidt, Yiyun Huang, and Richard E. Carson.

- Determination of receptor occupancy in the presence of mass dose: [11c]gsk189254 pet imaging of histamine h3 receptor occupancy by pf-03654746. *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism*, 37(3):1095–1107, 3 2017. ISSN 1559-7016. doi:10.1177/0271678X16650697. PMID: 27207170 PMCID: PMC5363483.
- [192] Deepak Cyril D'Souza, Jose A. Cortes-Briones, Mohini Ranganathan, Halle Thurnauer, Gina Creatura, Toral Surti, Beata Planeta, Alexander Neumeister, Brian Pittman, Marc Normandin, Michael Kapinos, Jim Ropchan, Yiyun Huang, Richard E. Carson, and Patrick D. Skosnik. Rapid changes in cb1 receptor availability in cannabis dependent males after abstinence from cannabis. *Biological Psychiatry. Cognitive Neuroscience and Neuroimaging*, 1(1):60–67, 1 2016. ISSN 2451-9022. doi:10.1016/j.bpsc.2015.09.008. PMID: 26858993 PMCID: PMC4742341.
- [193] Alexander Neumeister, Marc D. Normandin, James W. Murrough, Shannan Henry, Christopher R. Bailey, David A. Luckenbaugh, Keri Tuit, Ming-Qiang Zheng, Isaac R. Galatzer-Levy, Rajita Sinha, Richard E. Carson, Marc N. Potenza, and Yiyun Huang. Positron emission tomography shows elevated cannabinoid cb1 receptor binding in men with alcohol dependence. *Alcoholism, Clinical and Experimental Research*, 36(12):2104–2109, 12 2012. ISSN 1530-0277. doi:10.1111/j.1530-0277.2012.01815.x. PMID: 22551199 PMCID: PMC3418442.
- [194] Marc D. Normandin, Ming-Qiang Zheng, Kuo-Shyan Lin, N. Scott Mason, Shu-Fei Lin, Jim Ropchan, David Labaree, Shannan Henry, Wendol A. Williams, Richard E. Carson, Alexander Neumeister, and Yiyun Huang. Imaging the cannabinoid cb1 receptor in humans with [11c]omar: assessment of kinetic analysis methods, test-retest reproducibility, and gender differences. *Journal of Cerebral Blood Flow and Metabolism: Official Journal of the International Society of Cerebral Blood Flow and Metabolism*, 35(8):1313–1322, 8 2015. ISSN 1559-7016. doi:10.1038/jcbfm.2015.46. PMID: 25833345 PMCID: PMC4528005.
- [195] Mohini Ranganathan, Jose Cortes-Briones, Rajiv Radhakrishnan, Halle Thurnauer, Beata Planeta, Patrick Skosnik, Hong Gao, David Labaree, Alexander Neumeister, Brian Pittman, Toral Surti, Yiyun Huang, Richard E. Carson, and Deepak Cyril D'Souza. Reduced brain cannabinoid receptor availability in schizophrenia. *Biological Psychiatry*, 79(12):997–1005, 6 2016. ISSN 1873-2402. doi:10.1016/j.biopsych.2015.08.021. PMID: 26432420 PMCID: PMC4884543.
- [196] Tatu Kantonen, Tomi Karjalainen, Janne Isojärvi, Pirjo Nuutila, Jouni Tuisku, Juha Rinne, Jarmo Hietala, Valtteri Kaasinen, Kari Kallioikoski, Harry Scheinin, Jussi Hirvonen, Aki Vehtari, and Lauri Nummenmaa. Interindividual variability and lateralization of mu-opioid receptors in the human brain. *NeuroImage*, 217:116922, 8 2020. ISSN 1053-8119. doi:10.1016/j.neuroimage.2020.116922.
- [197] Katrin Amunts, Claude Lepage, Louis Borgeat, Hartmut Mohlberg, Timo Dickscheid, Marc-Étienne Rousseau, Sebastian Bludau, Pierre-Louis Bazin, Lindsay B Lewis, Ana-Maria Oros-Peusquens, et al. Bigbrain: an ultrahigh-resolution 3d human brain model. *science*, 340(6139):1472–1475, 2013.
- [198] Casey Paquola, Jessica Royer, Lindsay B Lewis, Claude Lepage, Tristan Glatard, Konrad Wagstyl, Jordan DeKraker, Paule-J Toussaint, Sofie L Valk, Louis Collins, et al. The bigbrainwarp toolbox for integration of big-brain 3d histology with multimodal neuroimaging. *Elife*, 10:e70119, 2021.
- [199] David C Van Essen, Stephen M Smith, Deanna M Barch, Timothy E.J. Behrens, Essa Yacoub, and Kamil Ugurbil. The wu-minn human connectome project: An overview. *NeuroImage*, 80:62–79, 2013. ISSN 10538119. doi:10.1016/j.neuroimage.2013.05.041. PMID: 23684880.
- [200] Fang-Cheng Yeh, Sandip Panesar, David Fernandes, Antonio Meola, Masanori Yoshino, Juan C Fernandez-Miranda, Jean M Vettel, and Timothy Verstynen. Population-averaged atlas of the macroscale human structural connectome and its network topology. *NeuroImage*, 178:57–68, 2018. ISSN 10959572. doi:10.1016/j.neuroimage.2018.05.027. PMID: 29758339.
- [201] Fang-Cheng Yeh, Van Jay Wedeen, and Wen-Yih Isaac Tseng. Estimation of fiber orientation and spin density distribution by diffusion deconvolution. *NeuroImage*, 55(3):1054–1062, 4 2011. ISSN 1053-8119. doi:10.1016/J.NEUROIMAGE.2010.11.087. publisher: Academic Press.
- [202] Fang-Cheng Yeh, T D Verstynen, Y Wang, J C Fernández-Miranda, and W-Yi Tseng. Deterministic diffusion fiber tracking improved by quantitative anisotropy. *PLoS ONE*, 8(11):80713, 2013. doi:10.1371/journal.pone.0080713.
- [203] Andrea I Luppi and Emmanuel A Stamatakis. Combining network topology and information theory to construct representative brain networks. *Network Neuroscience*, 5(1):96–124, 2021. ISSN 2472-1751. doi:10.1162/netn_a_00170.
- [204] Rahul S Desikan, Florent Ségonne, Bruce Fischl, Brian T Quinn, Bradford C Dickerson, Deborah Blacker, Randy L Buckner, Anders M Dale, R Paul Maguire, Bradley T Hyman, et al. An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest. *Neuroimage*, 31(3):968–980, 2006.
- [205] Richard F Betzel, Alessandra Griffo, Patric Hagmann, and Bratislav Mišić. Distance-dependent consensus thresholds for generating group-representative structural brain networks. *Network neuroscience*, 3(2):475–496, 2019.
- [206] Bratislav Mišić, Richard F Betzel, Azadeh Nematzadeh, Joaquin Goni, Alessandra Griffo, Patric Hagmann, Alessandro Flammini, Yong-Yeol Ahn, and Olaf Sporns. Cooperative and competitive spreading dynamics on the human connectome. *Neuron*, 86(6):1518–1529, 2015.
- [207] Jordy Tasserie, Antoine Grigis, Lynn Uhrig, Morgan Dupont, Alexis Amadon, and Béchir Jarraya. Pypreclin: An automatic pipeline for macaque functional MRI preprocessing. *NeuroImage*, 207, 2020. ISSN 10959572. doi:10.1016/j.neuroimage.2019.116353. URL <https://doi.org/10.1016/j.neuroimage.2019.116353>. PMID: 31743789.
- [208] Kelly Shen, Gleb Bezgin, Michael Schirner, Petra Ritter, Stefan Everling, and Anthony R McIntosh. A macaque connectome for large-scale network simulations in the virtualbrain. *Scientific data*, 6(1):123, 2019.
- [209] Rembrandt Bakker, Thomas Wachtler, and Markus Dies-

Supplementary Materials

- mann. Cocomac 2.0 and the future of tract-tracing databases. *Frontiers in neuroinformatics*, 6:30, 2012.
- [210] Ao Chen, Sha Liao, Mengnan Cheng, Kailong Ma, Liang Wu, Yiwei Lai, Xiaojie Qiu, Jin Yang, Jiangshan Xu, Shijie Hao, et al. Spatiotemporal transcriptomic atlas of mouse organogenesis using dna nanoball-patterned arrays. *Cell*, 185(10):1777–1792, 2022.
- [211] Joseph E. Knox, Kameron Decker Harris, Nile Gradis, Jennifer D. Whitesell, Hongkui Zeng, Julie A. Harris, Eric Shea-Brown, and Stefan Mihalas. High-resolution data-driven model of the mouse connectome. *Network Neuroscience*, 3(1):217–236, dec 1 2018. doi:10.1162/NETN_A_00066. URL <https://direct.mit.edu/netn/article/3/1/217/2194/High-resolution-data-driven-model-of-the-mouse>. publisher: MIT Press.
- [212] Joshua B. Burt, Murat Demirtaş, William J. Eckner, Natasha M. Navejar, Jie Lisa Ji, William J. Martin, Alberto Bernacchia, Alan Anticevic, and John D. Murray. Hierarchy of transcriptomic specialization across human cortex captured by structural neuroimaging topography. *Nature Neuroscience*, 21(9):1251–1259, 9 2018. ISSN 15461726. doi:10.1038/s41593-018-0195-0. PMID: 30082915 publisher: Nature Publishing Group.
- [213] Chad J Donahue, Stamatios N Sotiropoulos, Saad Jbabdi, Moises Hernandez-Fernandez, Timothy E Behrens, Tim B Dyrby, Timothy Coalson, Henry Kennedy, Kenneth Knoblauch, David C Van Essen, et al. Using diffusion tractography to predict cortical connection strength and distance: a quantitative comparison with tracers in the monkey. *Journal of Neuroscience*, 36(25):6758–6770, 2016.
- [214] Ross D Markello and Bratislav Misic. Comparing spatial null models for brain maps. *NeuroImage*, 236:118052, 2021.
- [215] Anthony Randal McIntosh and Nancy J Lobaugh. Partial least squares analysis of neuroimaging data: applications and advances. *Neuroimage*, 23:S250–S263, 2004.
- [216] Anthony R McIntosh and Bratislav Mišić. Multivariate statistical analyses for neuroimaging data. *Annu Rev Psychol*, 64:499–525, 2013.
- [217] František Váša and Bratislav Mišić. Null models in network neuroscience. *Nature Reviews Neuroscience*, pages 1–12, 2022.
- [218] Sergei Maslov and Kim Sneppen. Specificity and stability in topology of protein networks. *Science*, 296(5569):910–913, 5 2002. doi:10.1126/science.1065103. publisher: American Association for the Advancement of Science.
- [219] Richard F Betzel and Danielle S Bassett. Specificity and robustness of long-distance connections in weighted, interareal connectomes. *Proc. Natl. Acad. Sci. USA*, 2018. doi:10.1073/pnas.1720186115. [Online; accessed 2019-07-08].
- [220] Mark EJ Newman. Analysis of weighted networks. *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, 70(5):056131, 2004.
- [221] Mark EJ Newman and Michelle Girvan. Finding and evaluating community structure in networks. *Physical review E*, 69(2):026113, 2004.

Supplementary Results

Whereas Hansen et al. (110) focused on the relationship between cognition and molecular architecture, another study (109) investigated how transitions between different brain states (represented by the NeuroSynth maps of various cognitive operations, and replicated using maps of parameter estimates from in-scanner tasks) depend on the network architecture of the human structural connectome. Using principles of network control theory, they showed that the human connectome minimises the average energy required to transition between pairs of brain states, being more energetically efficient than any topologically random network, even when the number of connections (node degree) of each region is the same. Use of a more stringent null model that preserves not only the number of connections but also their length (thereby preserving the overall geometric embedding of the brain) led to an improvement over the degree-preserving networks, but still without reaching the energetic efficiency of the empirical human connectome – thereby demonstrating that network topology and geometry both play a role in the energetic efficiency of the human brain’s wiring diagram.

As with the analysis replicating the study of Hansen et al. (110), here we proceed in two steps. First, we repli-

cate the human results after replacing the NeuroSynth maps with brain maps from automated expert consensus (Fig. S8a,b). Transitions between the same 17 representative cognitive operations described above become significantly energetically more difficult if the connectome topology is disrupted, and even more difficult if both topology and geometry are randomised. Hence, NeuroSynth and automated expert consensus once again provides the same insights about human brain function (Fig. S8c).

Second, we repeat the same analysis in macaque and mouse, using LLM-derived cognitive maps and species-specific connectomes reconstructed from ex vivo tract-tracing (rather than in vivo diffusion MRI tractography). We recover the exact same pattern of results in both species: the empirical connectome is the most energetically efficient for supporting transitions between cognitive brain maps (Fig. S8d,e). Disrupting network topology incurs an energetic premium, and an even greater energetic cost is observed when both topology and geometry are randomised.

Supplementary Figures

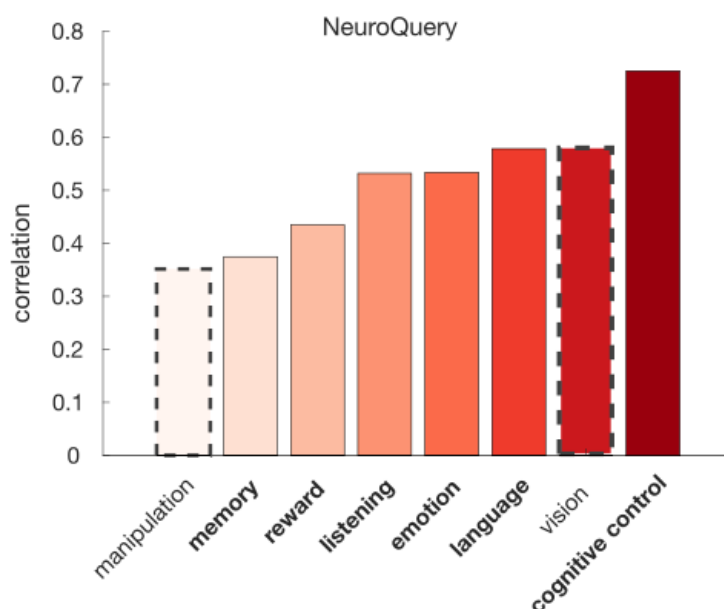


Figure S1. **LLM-generated brain maps of cognition correlate with NeuroQuery maps** | Correlation between meta-analytic maps and LLM-generated maps is replicated when using NeuroQuery instead of NeuroSynth to summarise the human neuroimaging literature. Significance of correlations is assessed against a population of null maps with preserved spatial autocorrelation generated using Moran spectral randomisation.

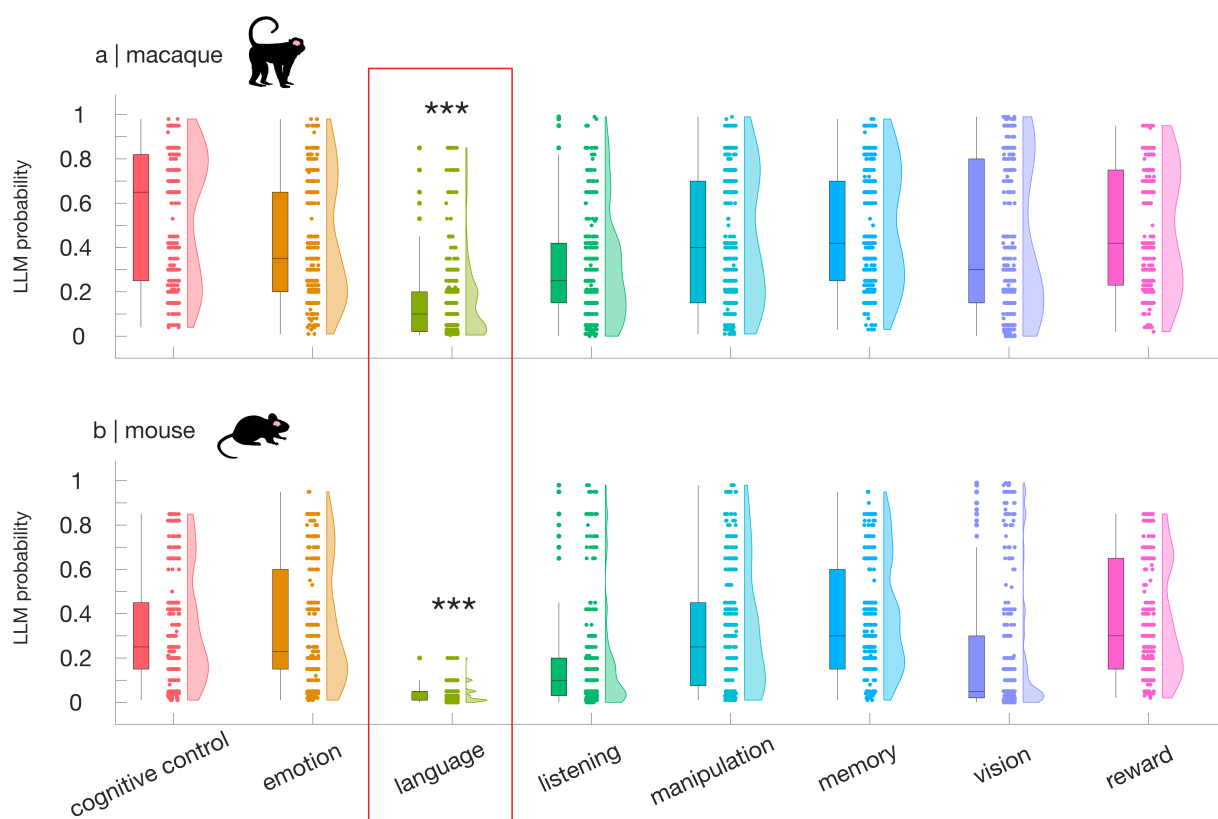


Figure S2. **LLM-estimated probability of regional involvement with “language” for macaque and mouse is significantly lower than any other cognitive domain** | (a) Macaque. (b) Mouse. For each species, data are combined across all regions and all LLMs.

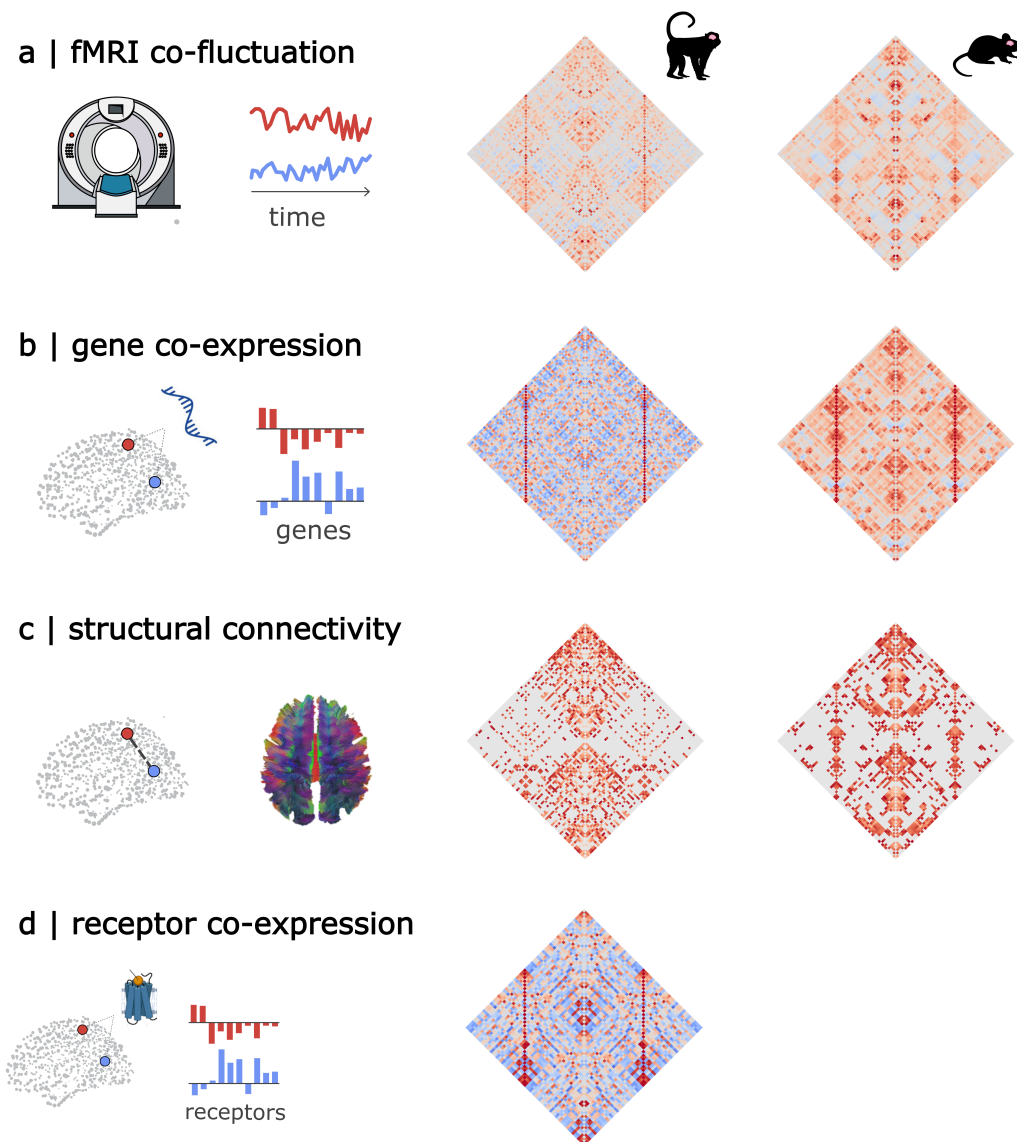


Figure S3. Networks of structural and functional interactions for macaque and mouse | (a) Co-fluctuation (correlation over time) of spontaneous activity from fMRI haemodynamics. (b) Correlated patterns of gene expression from stereo-seq (macaque) and in situ hybridization (mouse). (c) Structural connectivity from combined diffusion tractography and tract-tracing (macaque) and tract-tracing (mouse). (d) Correlated patterns of receptor density from in vitro autoradiography (macaque only). Note that macaque receptor density is not available for all regions, and regions without receptor expression were therefore excluded from the analysis.

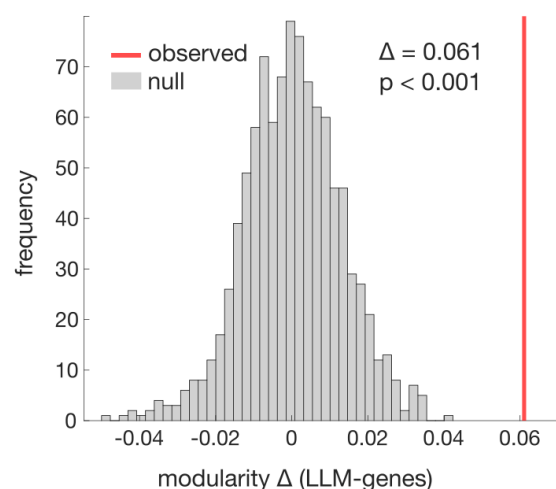


Figure S4. **Clustering of disease maps from LLMs aligns with symptom dimensions better than transcriptomically derived maps of molecular risk** | Modularity is computed using the Newman-Girvan algorithm (220, 221). The observed difference in modular structure is significantly greater than for randomly reorganised matrices.

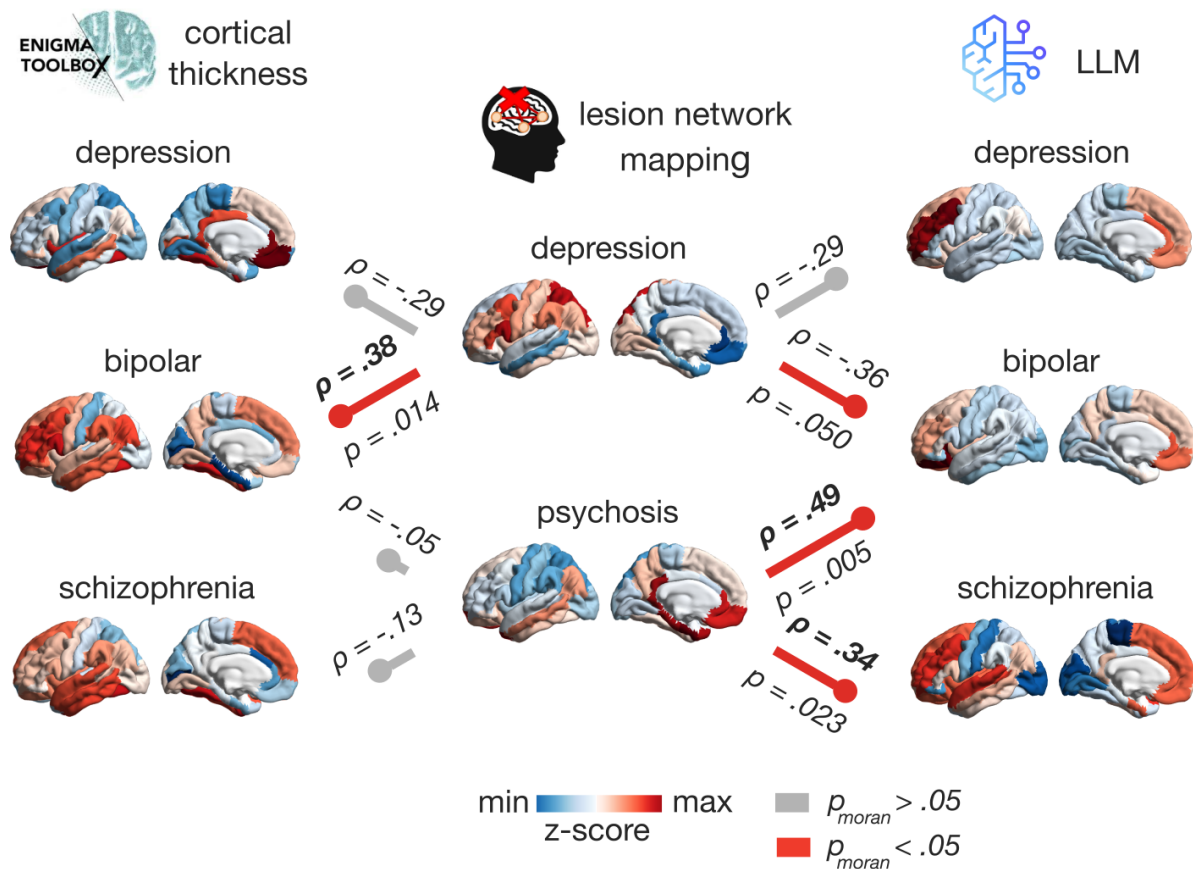


Figure S5. LLMs recapitulate circuits of cognitive dysfunction from lesion network mapping. | We use two maps of lesion circuits: (1) a psychosis circuit, derived from regions whose lesion induces psychosis symptoms (50) ($N = 153$ lesion cases), and (2) a depression circuit, derived from the convergence of regions whose lesion induces depressive symptoms, and whose brain stimulation (using DBS or TMS) alleviates depressive symptoms (49) (14 studies; $N = 461$ lesion cases and $N = 251$ stimulation cases). In other words, the two maps provide a causal mapping from regions to symptoms (5). We compare these lesion maps with the ENIGMA and LLM-derived maps for depression and schizophrenia, respectively, as well as bipolar disorder, which involves symptoms of both. Indeed, the LLM map for bipolar is significantly correlated with both depression and psychosis lesion maps, reflecting the nature of the disorder. The LLM map for schizophrenia is also significantly correlated with the lesion map for psychosis. Significance of correlations is assessed against a population of null maps with preserved spatial autocorrelation generated using Moran spectral randomisation.

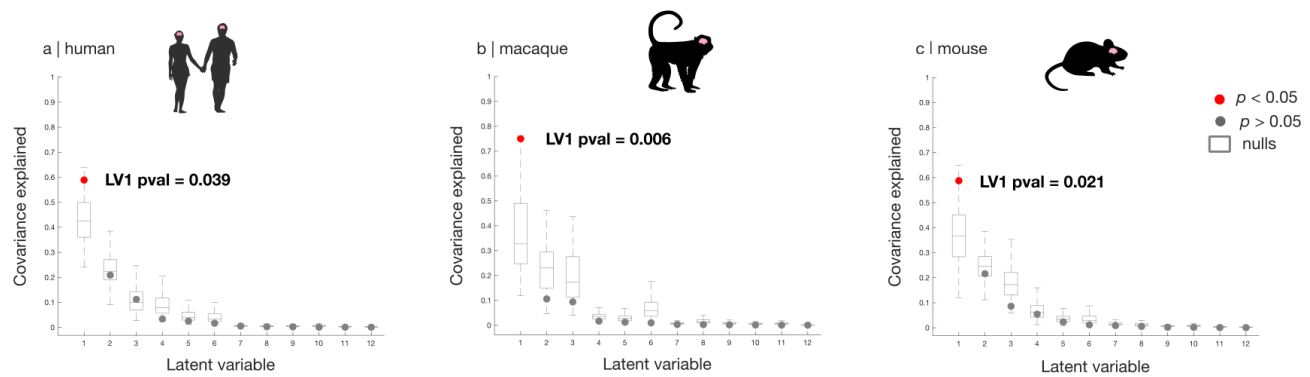


Figure S6. **AI-generated species-specific cognitive maps are significantly associated with species-specific transcriptomic patterns** | (a) Significant PLS LV1 in human. (b) Significant PLS LV1 in macaque. (c) Significant PLS LV1 in mouse. In each species, significance is assessed against a null distribution of species-specific spatial autocorrelation-preserving null maps generated using Moran spectral randomisation

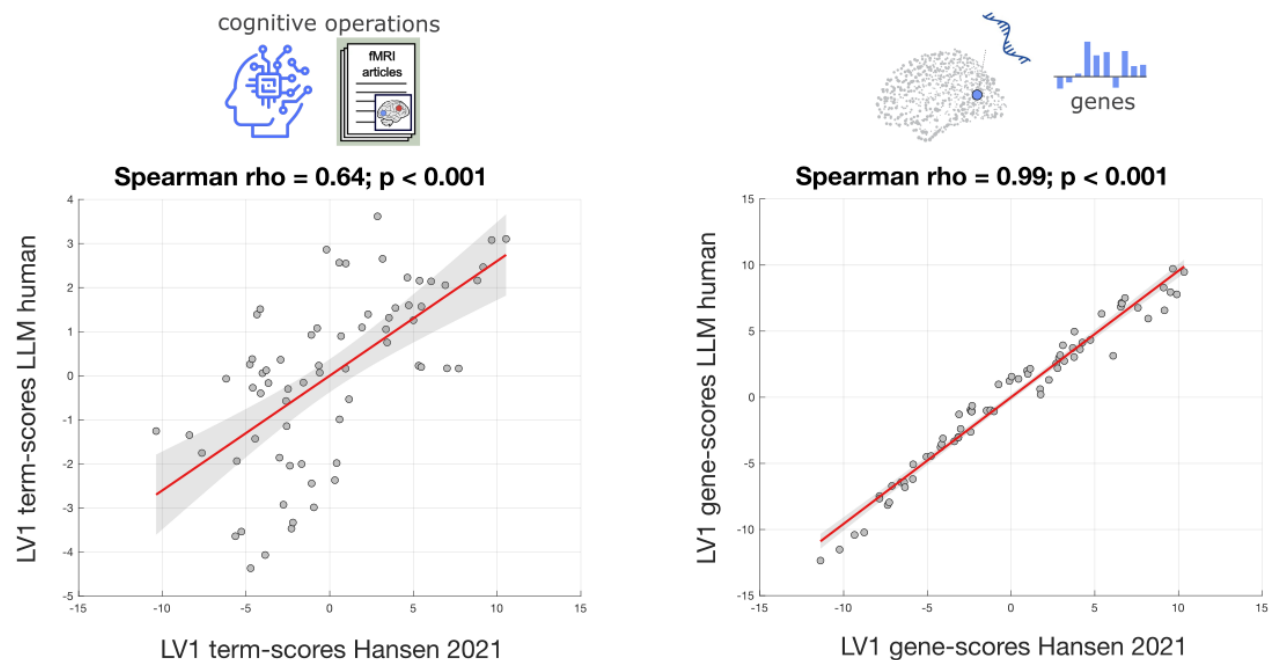


Figure S7. **AI-generated cognitive maps recapitulate gene-cognition associations obtained using NeuroSynth** | (Left) Significant correlation between PLS LV1 scores for cognitive terms obtained from NeuroSynth (as per (110); abscissa) and from AI experts (ordinate). (Right) Significant correlation between PLS LV1 scores for 81 brain-related genes obtained from NeuroSynth (as per (110); abscissa) and from AI experts (ordinate).

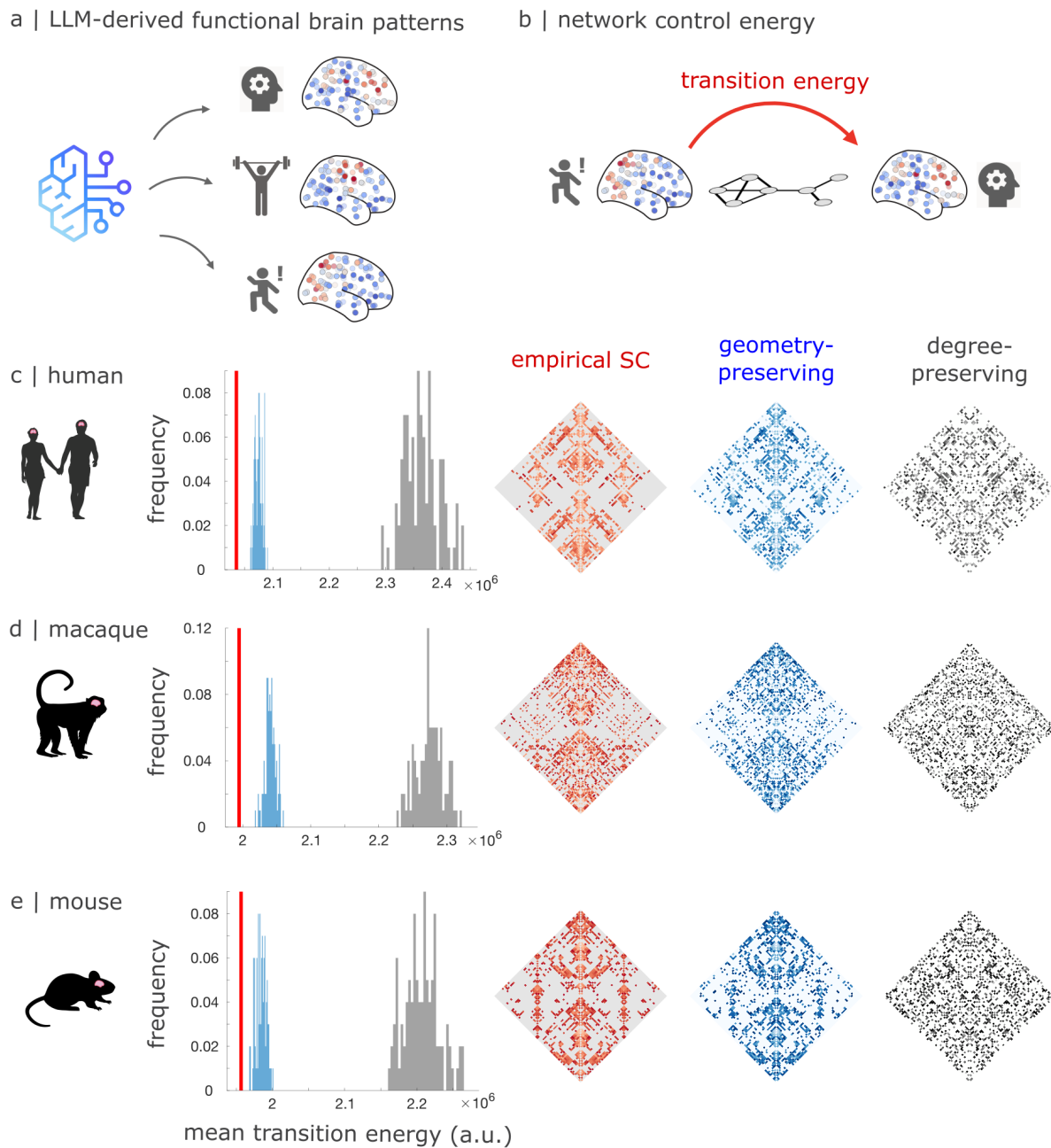


Figure S8. AI experts generalise the energy efficiency of human brain networks to macaque and mouse | (a) We use LLMs to obtain brain maps pertaining to 17 representative cognitive operations in human, macaque, and mouse. (b) We apply principles of network control theory with species-specific connectomes to compute the energy required to transition between each pair of cognitive brain maps. (c) Consistently across all three species, the mean transition energy across all pairs is significantly lower for the real connectome than for geometry-preserving rewired networks, which in turn outperform degree-preserving nulls.

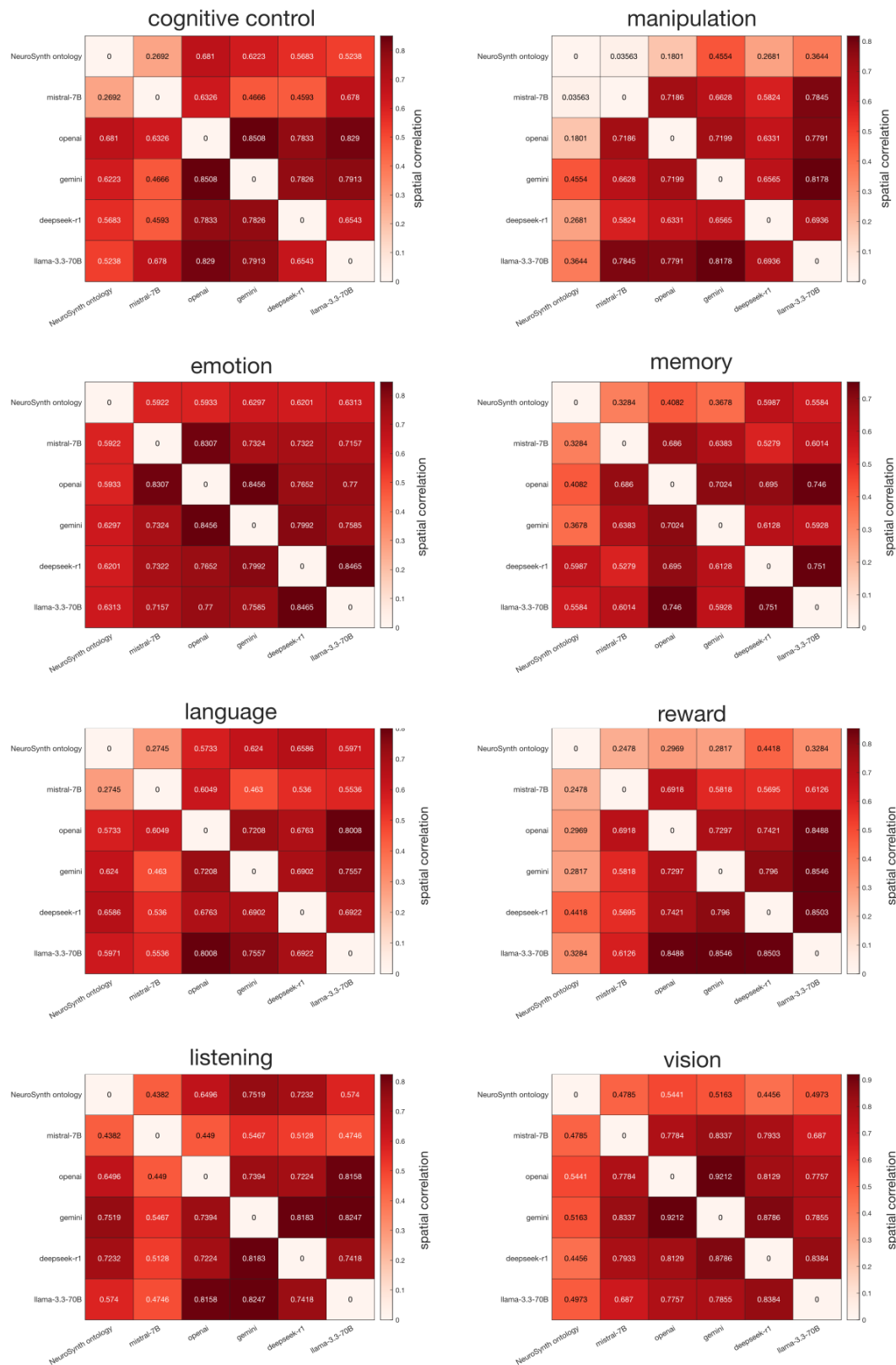


Figure S9. **Similarity of brain maps generated by different LLMs** | For 8 representative domains of cognition, we show the spatial correlation between the corresponding brain maps generated by NeuroSynth meta-analysis, and by 5 LLMs (Mistral, OpenAI's ChatGPT, Gemini, DeepSeek, and Llama).

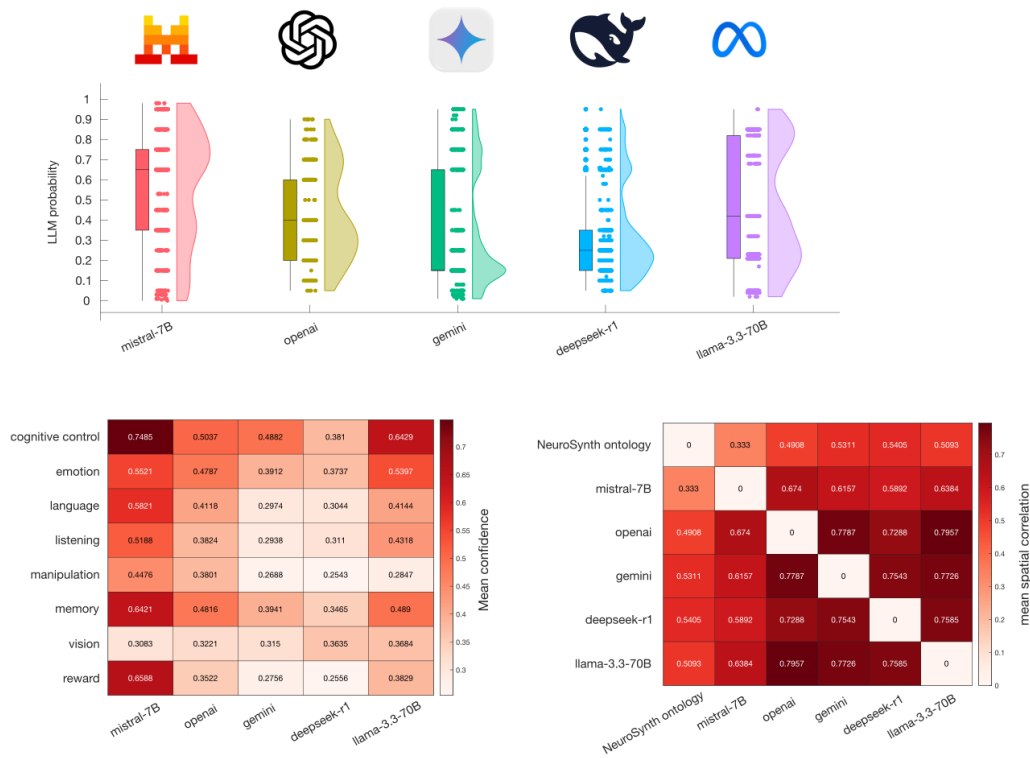


Figure S10. **Comparison of different LLMs** | (a) Distribution of LLMs' confidence across region-function pairs, for Mistral, OpenAI's ChatGPT, Gemini, DeepSeek, and Llama. (b) Breakdown of LLM confidence by LLM model and 8 representative domains of cognition, averaged across regions. (c) Similarity (spatial correlation) between brain maps generated by NeuroSynth and by different LLMs, averaged across the 8 domains of cognition in (b).

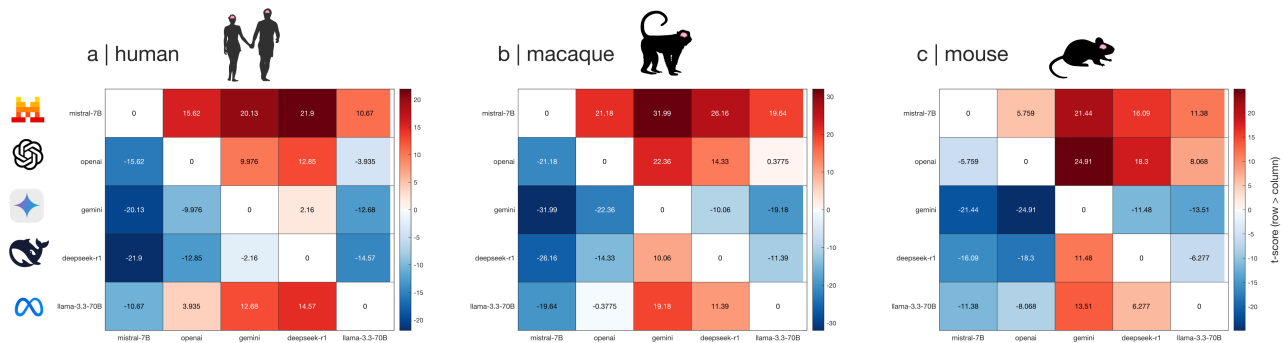


Figure S11. **Comparison of different LLMs' mean probability across region-function pairs** | (a) Human. (b) Macaque. (c) Mouse. Colorbar indicates the t-score from a paired-samples t-test, such that positive values indicate row > column.

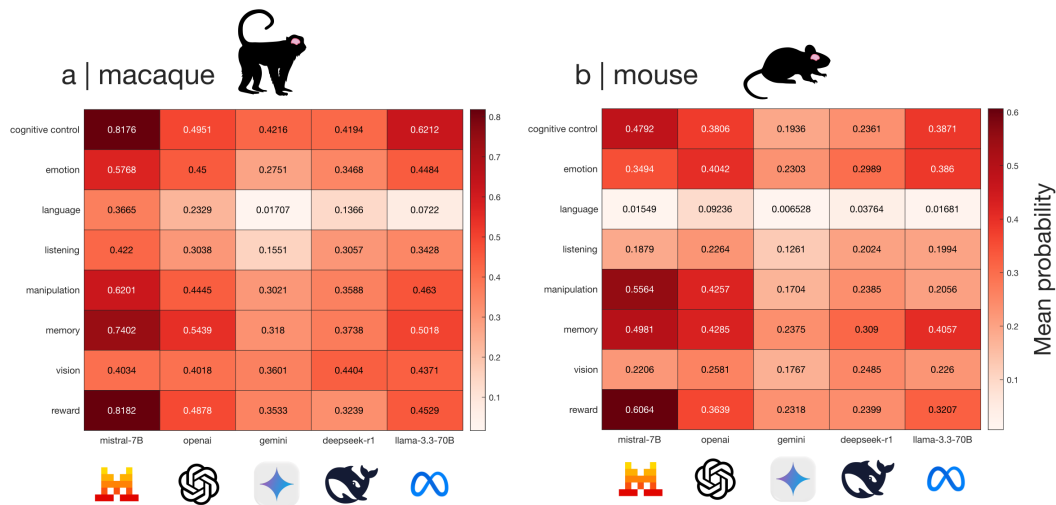


Figure S12. **Comparison of different LLMs for macaque and mouse** | (a) Breakdown of LLM confidence by LLM model and 8 representative domains of cognition, averaged across regions of the macaque brain. (b) Breakdown of LLM confidence by LLM model and 8 representative domains of cognition, averaged across regions of the mouse brain. Both show a clear drop in estimated probability for “language”.

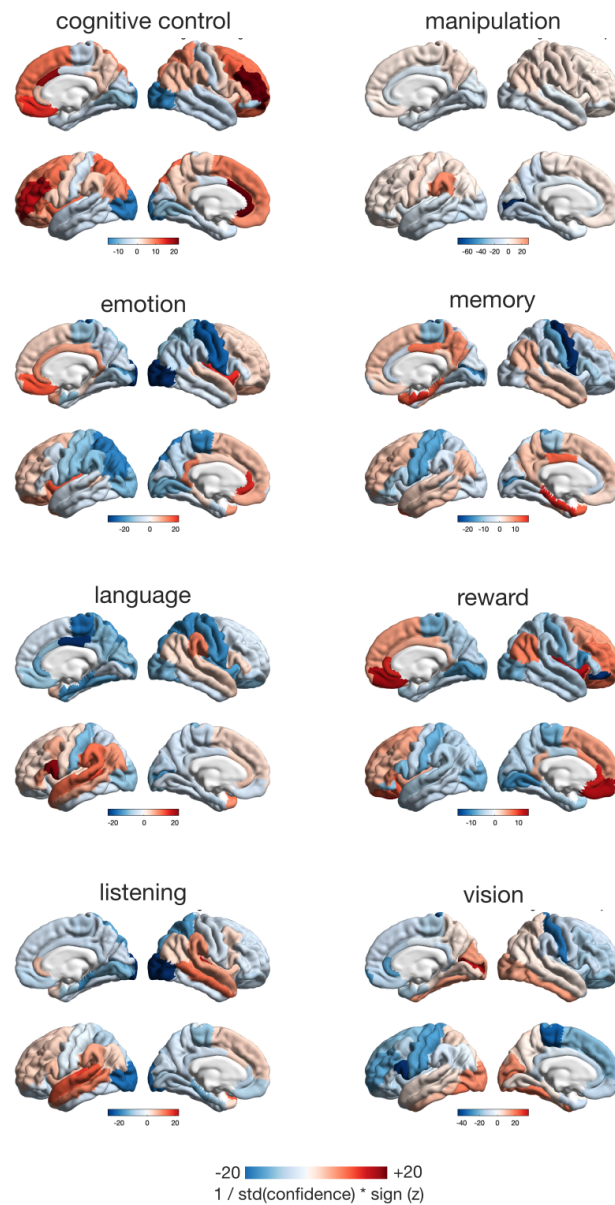


Figure S13. **Regional agreement of brain maps generated by different LLMs** | For 8 representative domains of cognition, each brain region shows the convergence (inverse of the standard deviation) between the values attributed by 5 different LLMS (Mistral, OpenAI's ChatGPT, Gemini, DeepSeek, and Llama). Therefore greater magnitude reflects greater agreement among LLMS. Sign is obtained from the mean of the z-scored maps. Altogether, a high positive value indicates that LLMS agree in attributing high relevance (strong association between region and function); high negative value indicates that LLMS agree in attributing low relevance (region is not associated with the function of interest); and values closer to zero indicate more divergence among models.

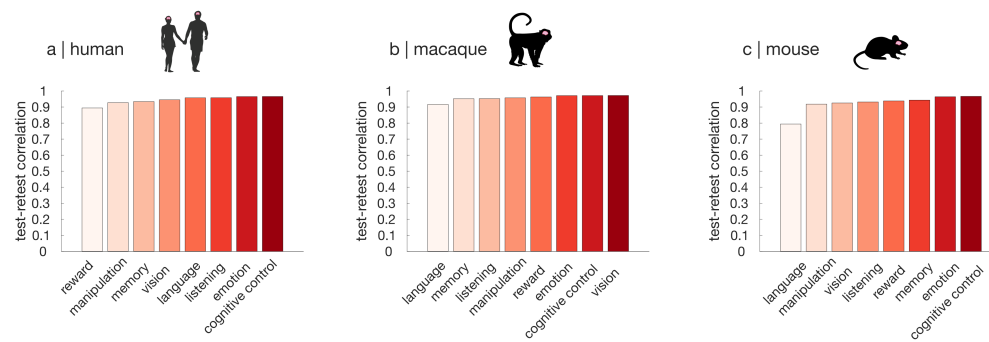


Figure S14. LLM-generated maps are consistent across different iterations of the workflow | (a) Human. (b) Macaque. (c) Mouse. All correlations are significant against a null distribution of spatial autocorrelation-preserving null maps generated using Moral spectral randomisation.

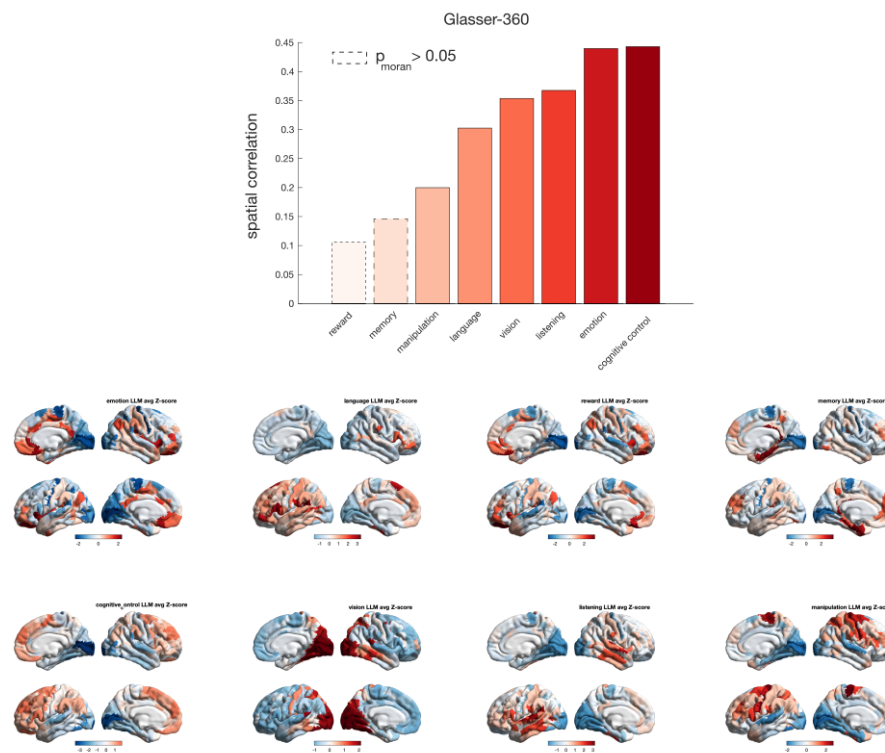


Figure S15. LLM-generated brain maps of cognition are robust to parcellation choice | (Top) Correlation between NeuroSynth meta-analytic maps and LLM-generated maps is replicated when using the finer-grain Glasser parcellation, with 360 regions instead of 68. Note that more than half of the regions in this parcellation (194 out of 360) were not previously described in the literature (114). Significance of correlations is assessed against a population of null maps with preserved spatial autocorrelation generated using Moran spectral randomisation. (Bottom) LLM-generated brain maps for the Glasser parcellation.

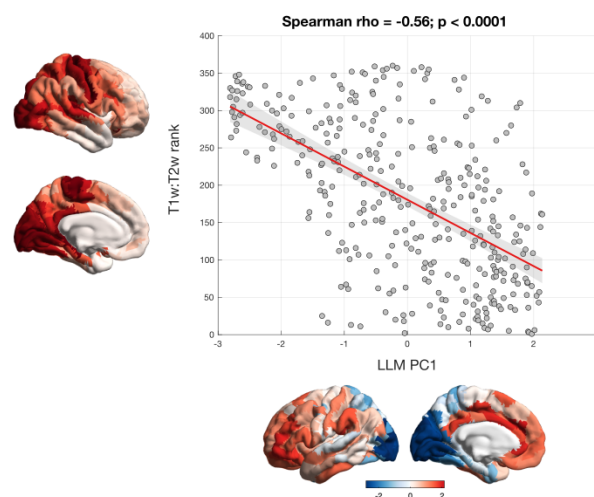


Figure S16. Cortical hierarchy identified by LLMs is robust to choice of parcellation | Significant correlation between the PC1 of LLM-derived functional similarity obtained using the finer-grain Glasser parcellation with 360 regions, and hierarchical position estimated using in vivo T1w:T2w ratio. Significance of correlations is assessed against a population of null maps with preserved spatial autocorrelation generated using Moran spectral randomisation.

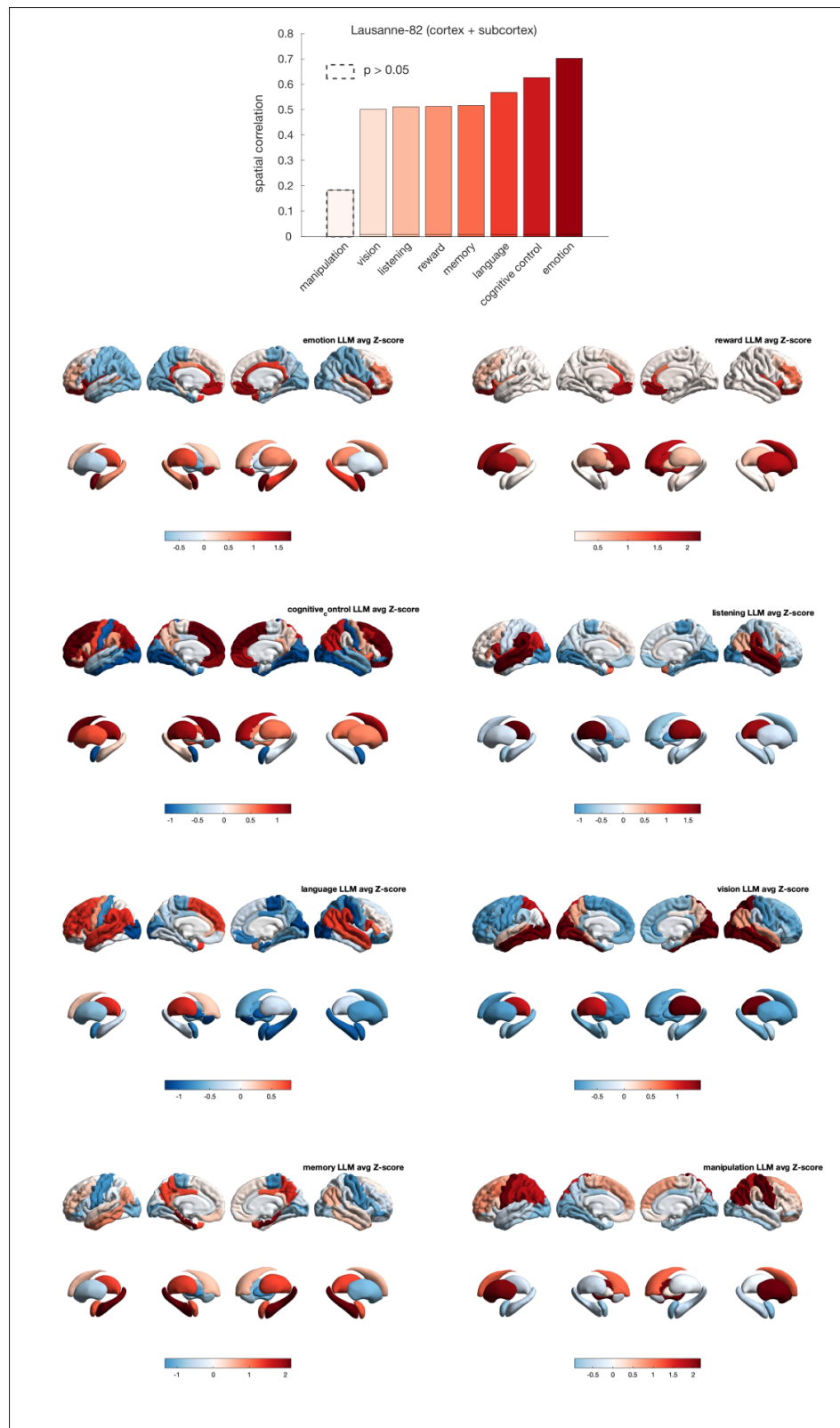


Figure S17. **LLMs generate brain maps of cognition including subcortex** | (Top) Correlation between NeuroSynth meta-analytic maps and LLM-generated maps is replicated when including subcortical structures. Significance of correlations is assessed against a population of null maps with preserved spatial autocorrelation generated using Moran spectral randomisation. (Bottom) LLM-generated brain maps with subcortex.

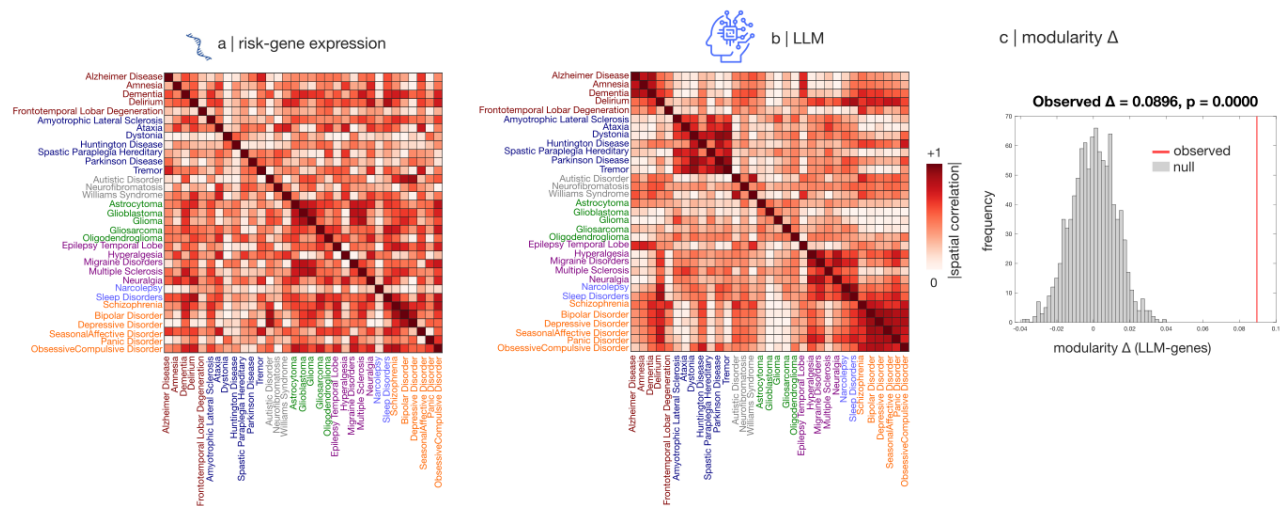


Figure S18. Clustering of disease maps from LLMs aligns with symptom dimensions better than transcriptomically derived maps of molecular risk even when including subcortex | (a) Similarity (magnitude of spatial correlation) between disease maps from regional expression of risk genes, including subcortical structures. (b) Similarity (magnitude of spatial correlation) between disease maps from LLM expert consensus, including subcortical structures. (c) Difference in modular structure based on assignment into 7 disease categories from the International Classification of Diseases manual: Neurocognitive disorders and Dementias; Neurodegenerative and movement disorders; Neurodevelopmental disorders; Brain tumors; Neurological disorders; Sleep and Wake disorders; and Mood and Behaviour disorders. The observed difference in modular structure is significantly greater than for randomly reorganised matrices.