Local and distributed information coding in the ventral stream

Running title: Local and distributed coding

Paul C. Bogdan[1,*], Cortney Howard[1], Kirsten Gillette[1],
Roberto Cabeza[1], Simon W. Davis[1,2]

[1] Department of Psychology and Neuroscience, Duke University, Durham, NC, 27708

[2] Department of Neurology, Duke University School of Medicine, Durham, NC, 27708

* Please address correspondence to:

Cabeza Lab, Duke University
LSRC B254, Box 90999
Durham, NC 27708
Tel: (919) 668-2299
Fax: (919) 681-0815
Email: paul.bogdan@duke.edu

Abstract

Neuroscience is awash with studies showing how virtually every cognitive or neural function is distributed across many regions. Yet, it is left unclear whether findings of widespread information processing imply large-scale integrative computations or instead should be seen as showing multitudes of localized modules. To investigate this distinction, we used fMRI data (N=60) from participants viewing objects across four tasks, and we examined the relationships between regions in terms of information coding. We demonstrate that coding in the occipital lobe is overwhelmingly localized and small-scale population codes are the foundation of early perceptual representation. In contrast, different portions of the inferior temporal lobe coordinate as a distributed unit to produce multi-regional population codes for semantic information. No other brain area – neither the parietal nor prefrontal cortices – shows the preference for distributed coding seen in the inferior temporal lobe. The inferior temporal lobe additionally displays uniquely low levels of redundancy in its information coding across its constituent regions, and we demonstrate how this supports lobe-wide semantic processing. Taken together, these results outline a framework of how the ventral stream transitions from local to distributed neuronal coding as information progresses from perceptual to semantic representations.

1. Introduction

One of neuroscience's oldest debates is whether the brain encodes information in a localized or distributed fashion. Early lesion studies were formative in localizing cognition.[1,2] But even in early work, distributed processing schemes were sown with hypotheses on neurological disorders deriving from the disconnection of neural regions.[3–5] Task-fMRI brought new tools to explore this distributed view, and techniques like representational similarity analysis (RSA) show how stimuli of all sorts prompt information coding across the entire brain.[6–8] Yet, the original question about the brain's functional architecture remains largely unanswered. If a stimulus elicits information coding in many regions, are these regions communicating to form distributed but integrated units of computation? Alternatively, do widespread patterns imply multitudes of modules, each independently representing different properties of a stimulus' content? As an analogy, a factory may be organized so the manufacture of its different products leverages overlapping systems, or the factory may instead be structured such that each product is made independently. Integrated systems can be more efficient but also create challenges. In the brain, whether different regions compose separate or unified computational units is a question that exists at multiple levels (**Figures 1A & 1B**). The present research seeks a tractable answer to this question, examining how localization and integration may vary across regions and depend on whether the regions encode perceptual or semantic content.

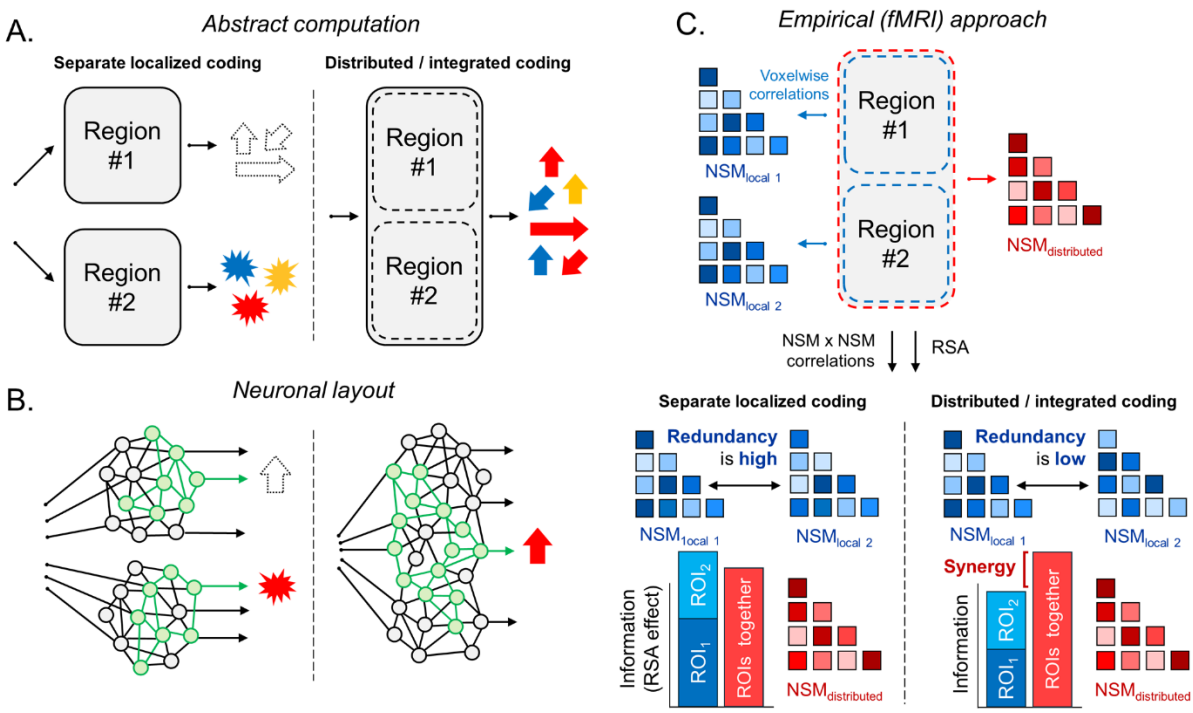

Figure 1. Illustration of the concepts targeted here. ***A.*** *Different features can either be extracted separately or simultaneously, and regions can be conceptualized as localized modules or distributed but integrated units. For example, Region 1 may code for motion, while Region 2 codes for color; alternatively, these regions may collectively contribute a distributed, synergistic code achieves an integrated representation.* ***B.*** *Neuronally, features can be parsed separately via multiple small-scale population codes or simultaneously with larger-scale population codes.* ***C.*** *To assess whether regions encode information in a local or distributed manner, representational similarity analysis (RSA) can be used. First, distributed units operate as more than the sum of their parts. Neural similarity matrices (NSMs) can be extracted from regions separately ($NSM_{local\ 1}$ & $NSM_{local\ 2}$) or while combining their data (NSM_{dist}). These NSMs can be submitted to RSA to evaluate the quantities of information encoded locally by individual ROIs or synergistically by the collective. Second, distributed units are organized in a non-redundant fashion, meaning distributed units should show weak correlations between its constituent regions.*

These questions about distributed or localized coding can be formalized and investigated by applying concepts from information theory.[9] Formally, encoding a feature means that a brain region distinguishes items across that feature; we say V4 encodes color because it displays different responses across differently colored objects and similar responses across similarly colored objects. In turn, the relationships between regions can be described in terms of two quantities: redundancy and synergy. *Synergy* depends on whether two regions coordinate to become more than the sum of their parts. Alongside V4 encoding color, consider how V5 encodes motion.[10] If V4 lesions cause both color and motion coding to collapse, this indicates that V4 and V5 coordinate to encode

information via distributed population codes. Alternatively, if V4 lesions largely spare motion processing, this points to localized codes. *Redundancy* describes the overlap in information coding between two regions. For example, if red objects usually move fast and V4 represents redness, then V5 representing speed would be partially redundant. In contrast, an imaginary distributed V4/5 unit that simultaneously parses both color and shape would avoid dedicating resources to duplicate coding (i.e., the Efficient Coding Hypothesis[11]). Note that the brain will virtually never showcase pure redundancy – e.g., in the color example, V4 would still be useful for non-red objects – but partial redundancy can still exist and point to localized codes. Conversely, if a set of regions are synergistic and non-redundant in their information processing, this suggests that they should be seen as composing an integrated computational unit with distributed population codes.

Under this theoretical lens, the distributed or localized organization of information processing can be studied within fMRI data and the RSA paradigm. In RSA, correspondence between a theoretical representational similarity matrix (RSM) and a measured neural similarity matrix (NSM) indicates that the information described in the RSM is represented by a brain area. Significant RSA effects, in turn, shed light on regions' neuron-level population coding.[12,13] To study whether regions coordinate synergistically to form a distributed unit, the NSM can be defined while pooling voxels across regions (**Figure 1C**). If the resulting NSM better fits the theorized RSM better than expected from analyzing the regions individually, this means the regions interact to encode a super-additive amount of information. To study redundancy, analyses can correlate regions' NSMs (**Figure 1C**). Weak correlations between pairs of regions mean these regions largely distinguish different items and thus minimally encode redundant information, which is also consistent with these regions operating as a distributed unit.

In sum, our research adapts traditional RSA techniques to study how collections of regions encode information in distributed or localized fashions. Our study uses fMRI data (N = 60) from a

memory paradigm where 114 real-world objects were repeatedly shown across four stages involving (i) passive naming, (ii) contextual encoding, (iii) visual recognition, and (iv) conceptual recognition. Here, we focus on patterns common across different stages to ensure generalizability across processing demands. Analyses were as follows: *First*, we examined the distributed versus localized distinction by investigating how pooling voxels from multiple regions may magnify RSA effects, reflecting synergy. We also consider how the degree of local or distributed coding may vary across the representation of perceptual or semantic information. *Second*, we measured correlations between regions to study overlaps in the information encoded in each, such that high overlaps point to redundancy and thus localized coding. *Third*, we empirically assessed the link between synergy and redundancy, expecting to find an inverse relationship that broadly supports this information theoretic perspective on neural computation.

2. Results

2.1. Standard representational similarity analysis

Before distinguishing areas along the localized/distributed distinction, we performed a traditional RSA analysis as a reference point. We parcellated the brain into 210 neocortical ROIs using the Brainnetome Atlas (mean size = 4.9 cm^3). For each ROI, an NSM (Neural Similarity Matrix) was generated via across-trial correlations of the ROI's voxels. Each ROI's NSM was then correlated with a model RSM corresponding to perceptual features (early convolutional neural network activation) or semantic features (word2vec embedding). Next, second-order RSM × NSM_{ROI} correlations were computed separately for each participant and each of the experiment's four stages. The correlation coefficients were averaged across stages by participant and submitted to one-sample t-tests for the group-level analysis (**Figure 2**). Perceptual information was overwhelmingly represented in posterior areas, while semantic representations were more anterior and spatially diffuse – consistent with earlier work.[14–16] However, more elaborate analyses are

needed to disentangle whether these patterns within individual ROIs wholly capture population codes or just components of larger ones.

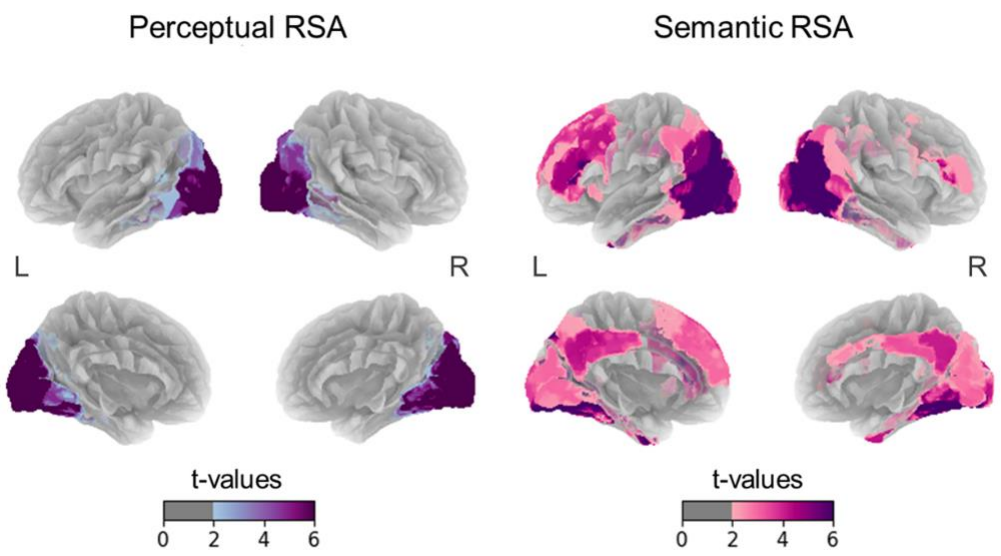

Figure 2. Traditional RSA results for perceptual or semantic coding. *RSM (perceptual or semantic) × NSM_{ROI} correlations were performed separately for 210 neocortical. After averaging the correlation coefficients across the four stages, the data were submitted to one-sample t-tests whose results are shown here. Although 210 ROIs were tested, no correction for multiple comparisons in significance testing was applied, as these maps are simply meant as reference for the later analyses.*

2.2. Distributed versus localized coding and RSA synergy

Our next analysis examined distributed and localized coding across four large-scale areas: the occipital lobe, inferior temporal lobe, parietal lobe, and prefrontal cortex (PFC) (volumes 138-288 cm^3). We contrasted RSA effects between (i) a baseline NSM capturing the within-ROI coding, and (ii) a NSM that also accounts for distributed codes spanning multiple ROIs. For an area's baseline NSM, we averaged its constituent ROIs' NSMs to form $NSM_{m\text{-}local}$ (**Figure 3A**). Then, to model synergistic relationships, the voxels from its multiple ROIs were pooled (effectively producing one huge ROI), and voxel-wise correlations were used to define NSM_{dist}. The two NSMs were then separately submitted to second-order correlations. For the results, if an area's NSM_{dist} better tracks a model RSM, this would indicate that the area holistically encodes more information than the sum of its parts.

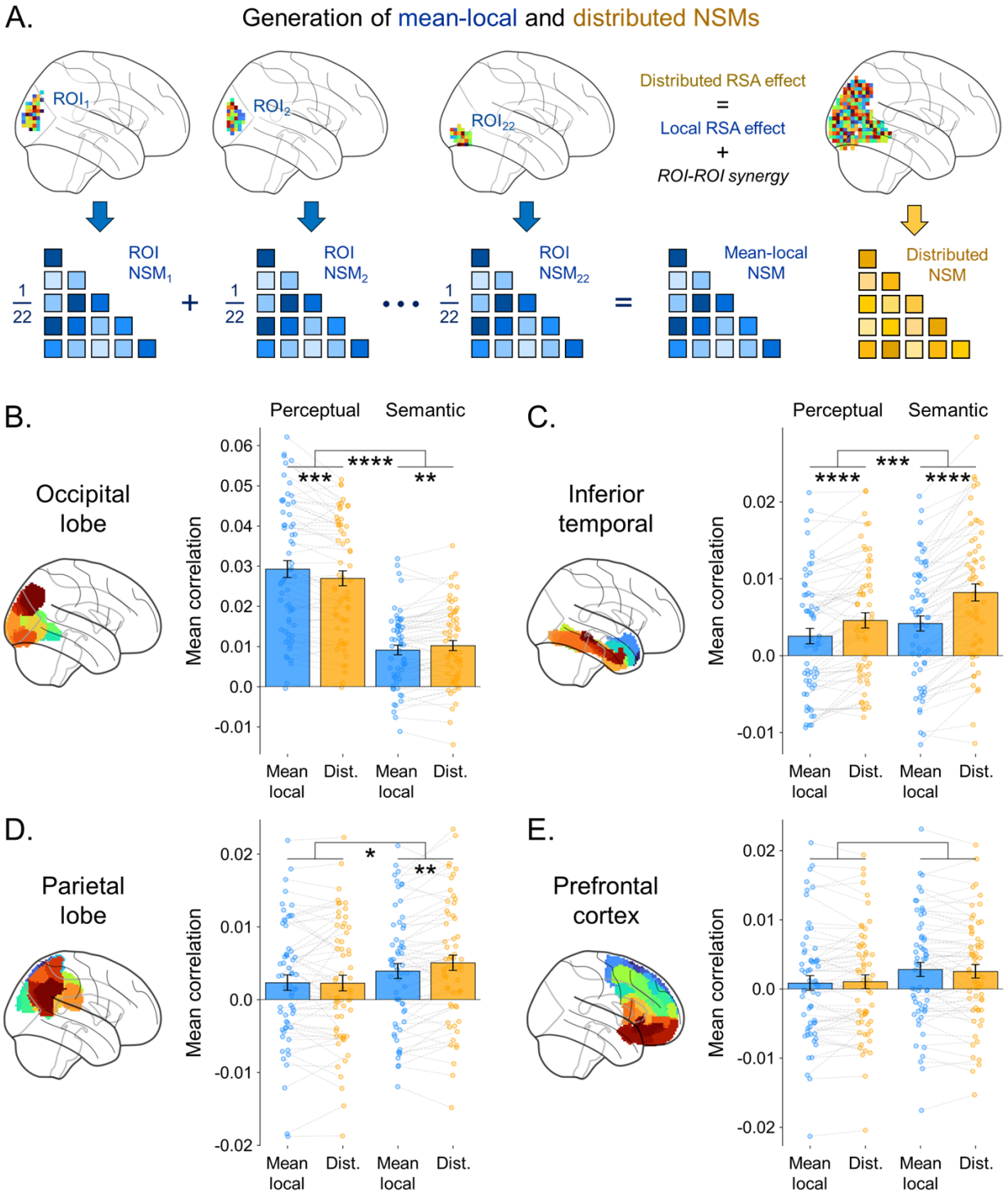

Figure 3. RSA results for small or large coding. ***A.*** *Illustration of the methodology used to generate the local/distributed NSMs. This example is based on the 22 occipital ROIs. We include a high-level equation: "Distributed RSA effect = local RSA effect + synergy", this presumes that both the distributed and local effects are equally impacted by redundancy, and the synergy addition emerges from the distributed NSM considering interactions across ROIs; redundancy will be investigated later.* ***B-E.*** *Results of the RSM (perceptual or semantic) × NSM (local or distributed) correlations for the four large areas tested. Dots represent individual participants, and the dashed lines indicate pairs from the same participant. Significance stars reflect paired t-tests or interactions from repeated-measure analyses of variance (ANOVAs). *, $p < .05$; **, $p < .01$; ***, $p < .001$; ****, $p < .0001$.*

The occipital lobe emerged as a core for processing visual features but showed no signs of encoding information in a distributed fashion (**Figure 3B**). In fact, particularly strong RSA effects emerged specifically for small-scale coding, suggesting the occipital lobe was characterized by particularly localized modules. This is a double dissociation relative to the inferior temporal lobe; the two-way Area (occipital/inferior temporal) × Size (Local/Distributed) interaction for perceptual information is significant ($F[1, 59] = 42.89$, $p < .0001$). By contrast, the strongest inferior temporal lobe effects were seen for semantic representation and distributed coding (**Figure 3C**). Measuring the Area × Size interaction now for semantic information is again significant ($F[1, 59] = 32.69$, $p < .0001$). In the parietal lobe and prefrontal cortex, differences between coding sizes were smaller or non-existent (**Figures 3D & 3E**), pointing to the uniqueness of the occipital lobe's localized coding and the inferior temporal lobe's distributed coding.

Several further tests were done to confirm the distributed coding synergy in the inferior temporal lobe. The first test focused on disentangling RSA effects common to the within-ROI and area-wide results. The area-wide NSM was revised to be now based on across-trial correlations between ROIs' average activations (i.e., essentially treating each ROI as one voxel). For example, there were 32 inferior temporal lobe ROIs, so the $NSM_{dist\text{-}avgs}$ correlations were taken between 32-length vectors. Averaging by ROI eliminated any within-ROI patterns. Along with this change, the original RSM × NSM strategy was revised to regress the model RSMs on small-scale and large-scale NSMs together ($RSM \sim NSM_{m\text{-}local} + NSM_{dist\text{-}avgs}$) (**Figure 4A**). This produced the same double dissociations as above with significant Area × Size interactions for perceptual information ($F[1, 59] = 32.15$, $p < .0001$) and significant Area × Size interactions for semantic information ($F[1, 59] = 10.31$, $p = .002$) (**Figures 4B & 4C**). There were no significant differences in the parietal lobe or PFC (**Figures 4D & 4E**), meaning that these changes in coding size specifically concern how the ventral stream processes information.

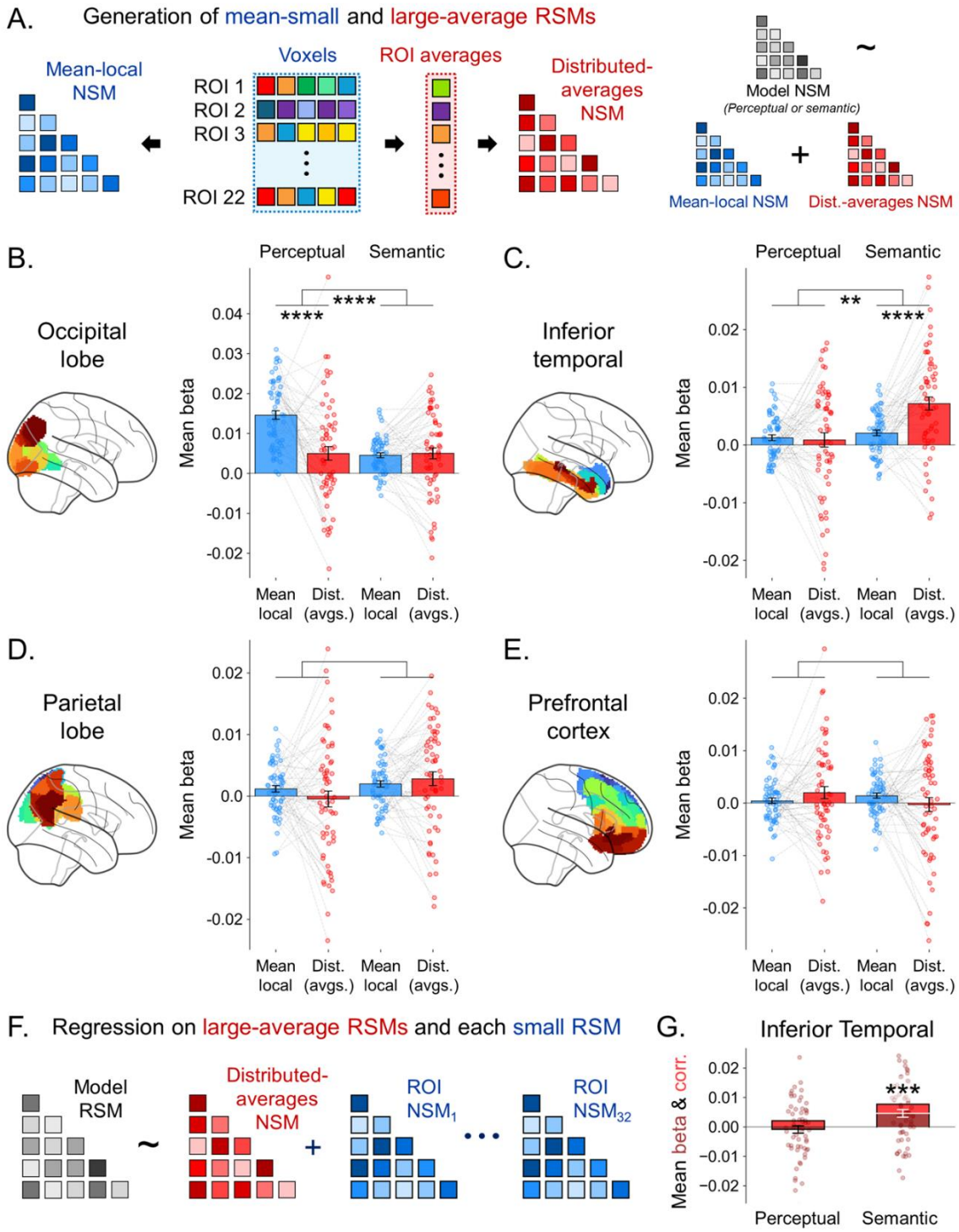

Figure 4. RSA results for local or distributed (ROI averages) coding. ***A.*** *Unlike the Figure 3 analyses, here a distributed NSM is defined with correlations between constituent ROIs' average activities. Also, whereas the Figures 2 and 3 methodologies performed RSA with RSM × NSM correlations, here it is done with a regression of RSM ~ $NSM_{m\text{-}local}$ + $NSM_{dist\text{-}a}$.* ***B-E****. Group-mean bars, participant dots, and paired t-test/ANOVA interaction results are reported as in Figure 3.* ***F.*** *A separate analysis focused on just the inferior temporal lobe and regressed a model RSM on the large-averages NSM along with all 32 inferior temporal lobe small-scale NSMs.* ***G.*** *The results of this many-predictor regression are shown in dark red. The three white stars indicate that the $NSM_{dist\text{-}a}$ beta significantly surpasses zero. For reference, the results of RSM × $NSM_{dist\text{-}a}$ correlations are also shown in lighter red. *, $p < .05$; **, $p < .01$; ***, $p < .001$.*

Yet stricter analyses specifically attempted to identify a distributed synergy effect in the inferior-temporal lobe while fully accounting for within-ROI coding. Here, model RSMs were regressed on $NSM_{dist\text{-}avgs}$ along with all 32 constituent ROIs' small-scale NSMs ($RSM \sim NSM_{dist\text{-}avgs} + NSM_{ROI\ 1} + \ldots + NSM_{ROI\ 32}$) (**Figure 4F**). Our focus is on the $NSM_{dist\text{-}avgs}$ beta coefficient, and whether it surpasses zero while the other predictors ruled out small-scale coding in any one ROI. Group-level analyses indeed showed that the inferior temporal lobe's $NSM_{dist\text{-}avgs}$ still significantly tracked semantic information (**Figure 4G**). This is strong evidence of distributed coding – i.e., information encoded collectively that cannot be explained by the activity within any individual ROI. Hence, the inferior temporal lobe is organized such that relatively far-flung neuronal elements coordinate to represent semantic information as an integrated unit. To be clear, this does not imply that the inferior temporal lobe is homogeneous but that it is organized around cooperative information processing. This aspect would be further tested in analyses of redundancy in coding.

2.3. Distributed coding and low redundancy

If a set of regions form integrated computational units with distributed population codes, then it may be expected that they will organize to minimize redundancy in their information processing, per the efficient coding hypothesis.[11] To examine distributed processing under this lens, we first computed $NSM_{ROI} \times NSM_{ROI}$ correlations between ROIs.[7] These correlations were markedly decreased within the inferior temporal lobe compared to other brain systems, suggesting that different inferior temporal regions specialize in coding different information (**Figure 5A**). As a further test, we probed item-wise neural-pattern similarity – e.g., if one ROI shows a consistent voxel-wise signature to a "cat" stimulus across two task stages, then the ROI is said to represent the item "cat". As our task included four stages, neural-pattern similarity was computed between all six possible pairs of stages. We repeated this process for all 114 objects, generating a 114-length vector for each ROI, which was then correlated between ROIs. These correlations were again strikingly

decreased among inferior temporal regions (**Figure 5B**). For reference, computing functional connectivity – the correlation between ROIs' mean univariate activities – yielded less pronounced inferior-temporal drops (**Figure 5C**), indicating that these other trends specifically concern information coding. Altogether, these low correlations within the inferior temporal lobe are consistent with the area performing coordinated distributed computations. As the earlier results suggest (**Figures 3 & 4**), these likely support semantic representation.

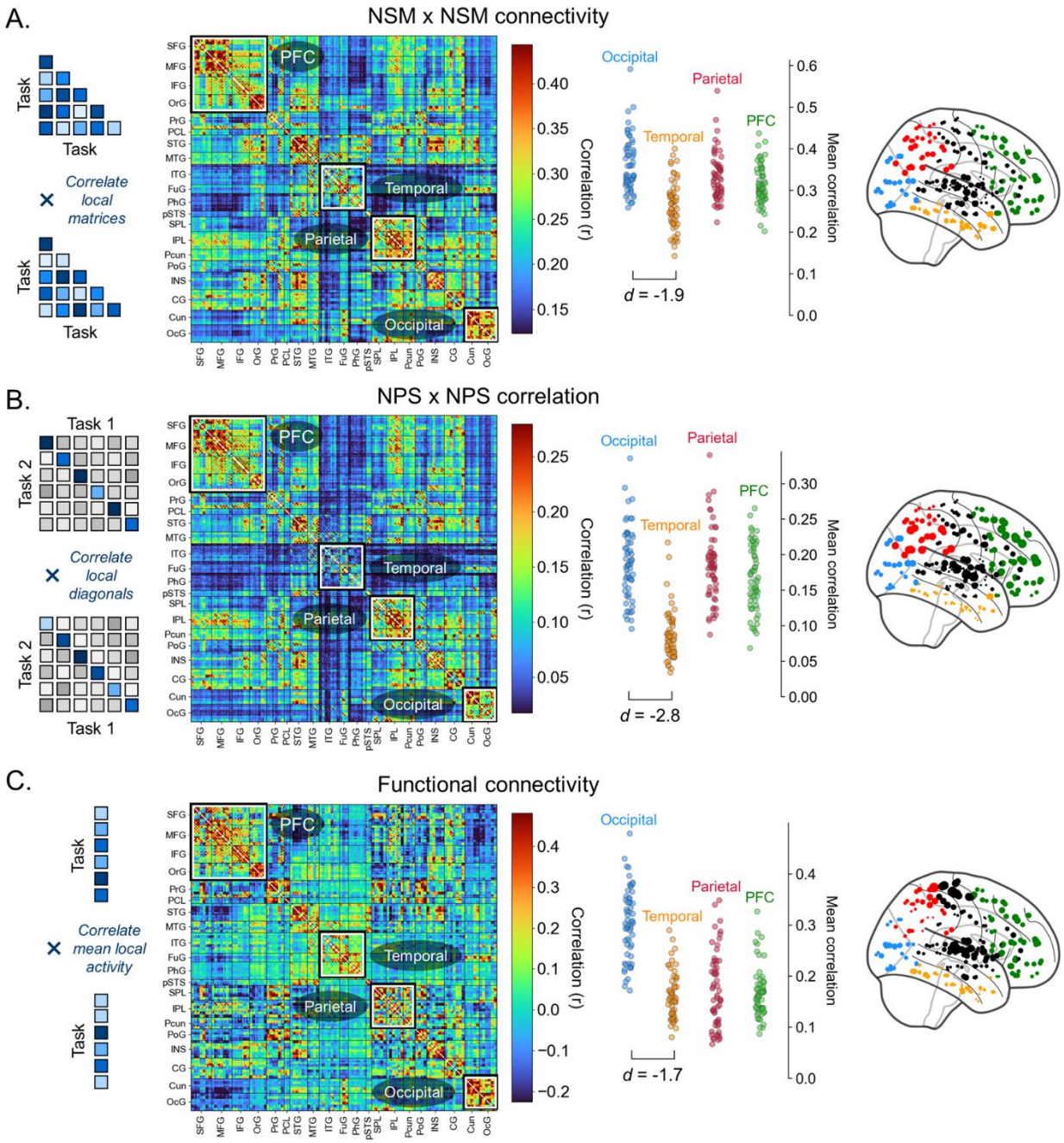

Figure 5. Heterogeneity in temporal lobe processing. *The three rows represent ROI-ROI correlations between (**A.**) small-scale NSMs, correlated separately for the four tasks then averaged, (**B.**) neural pattern similarity (NPS) measures, computed and correlated separately for the six possible pairs of two tasks, then averaged (**C.**) for reference, functional connectivity, correlated separately for the four tasks then averaged. Each dot in jitter plots next to each matrix represents one participant's average correlation within the specified anatomical region. Cohen's ds are reported from paired t-tests comparing the occipital and inferior temporal lobes. In the glass brains on the right, each marker corresponds to one ROI and the marker size represents the ROI's mean correlation with its 10 physically closest neighbors in the same hemisphere.*

Voxel-wise analyses were also performed correlating neighboring areas more precisely and to rule out the possibility that the above findings reflect differences in ROIs' shapes or areas' sizes. Gray matter was segmented into 2×2×2 voxel cubes (6×6×6 mm). For each cube, a voxel-wise NSM_{local} was defined and correlated with the NSM_{local} of every other cube within 20 mm. Each cube's correlations with its neighbors were averaged and plotted (**Figure 6**), which showed patterns consistent with the earlier correlation matrices. Most relevant to the present focus is the ventral red strip in **Figures 6A & 6B**. It corresponds to the inferior temporal gyrus, which had distinctly lower correlations among neighbors than occipital lobe areas. This is consistent with the inferior temporal sub-regions organizing their information processing to operate better as an integrated unit. Interestingly, a new pattern emerged where the orbitofrontal cortex (OFC) displayed lower redundancy than elsewhere in the PFC (**Figure 6C**), which will briefly be further investigated later.

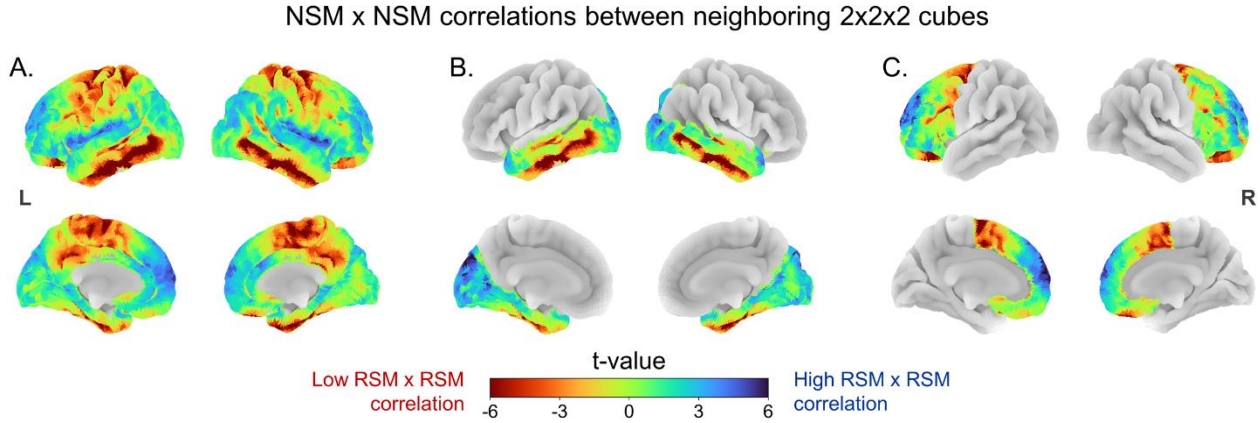

Figure 6. Cube-wise NSM × NSM correlations. *Colors are based on the mean NSM × NSM correlation between a 2×2×2 voxel cube at a given location and neighboring cubes within 20 mm. The mean correlations were computed independently for the left and right hemispheres, so no cross-hemisphere interplay is considered, as this could bias medial versus lateral comparisons. For each participant, NSM × NSM correlations were z-standardized across the brain – e.g., if there were just three locations with values 4.0, 3.0, and 2.0, these would be standardized to +1, 0, and -1. The standardized values were averaged across the four stages and submitted to a one-sample t-test, generating the shown t-values. Hence, each location's t-value effectively represents a comparison to other examined locations; this produces slight differences in the t-values across the analyses of (**A.**) the whole neocortex, (**B.**) just the occipital and temporal lobes, and (**C.**) just the prefrontal cortex.*

Taken together with the earlier RSA findings, the inferior temporal lobe ROIs were synergistic in information coding and minimally redundant – both patterns pointing to the inferior temporal lobe producing distributed population codes. The opposite was the case in the occipital lobe, where ROIs were not synergistic and encoded information with considerable redundancy, pointing to highly localized coding. Our final analyses would attempt to give further empirical support for this interpretation of our results and our rationale on the link between synergy and redundancy.

2.4. Distributed representation and local heterogeneity

Two analyses were conducted to justify that distributed population coding of semantics (synergy between ROIs) is intertwined with region-level heterogeneity (limited redundancy). First, following this logic, aside from just the inferior temporal lobe, other areas showing low *NSM × NSM* correlations should likewise display distributed coding of semantics. Hence, we return to that pattern noted earlier, whereby the OFC showed weaker correlations among neighbors than elsewhere in the PFC (**Figure 6C**). Performing the local/distributed RSA analysis on the OFC shows that the OFC indeed displayed a proclivity for distributed semantic coding, whereas the rest of the PFC does not (**Figure 7A**). Although the statistical effect is much weaker than the inferior temporal patterns, this is consistent with the logic laid out thus far.

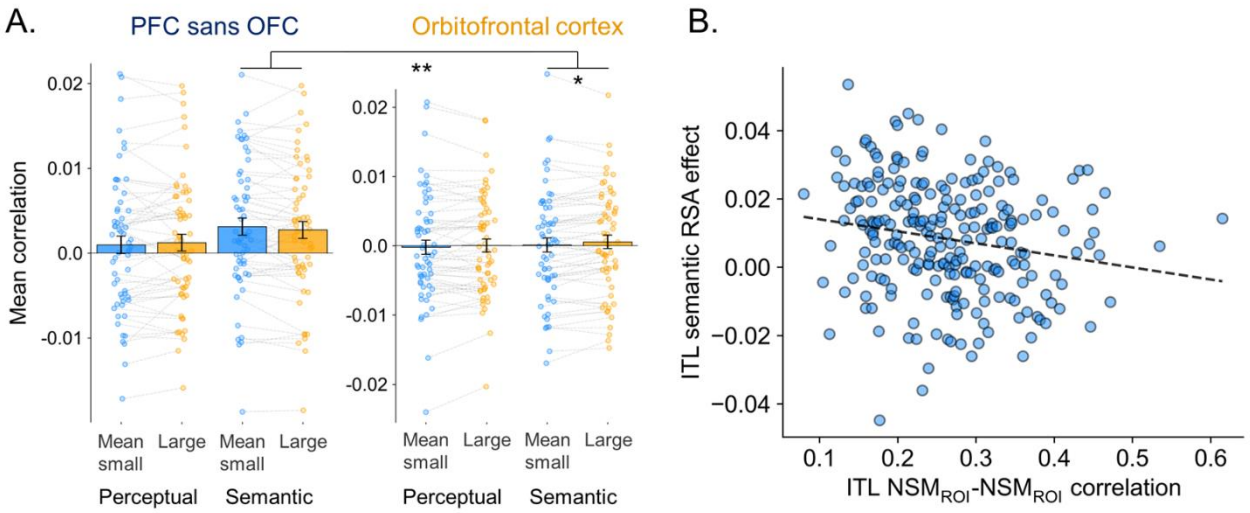

Figure 7. Assessing the link between distributed representation and regional correlations. ***A.*** *The analysis from Figure 3 was conducted for (i) one area defined as the superior frontal gyrus, the middle frontal gyrus, the inferior frontal gyrus, and the anterior cingulate cortex, and (ii) one area defined as just the orbitofrontal gyrus. Paired t-tests were performed for every small/large comparison, and a repeated-measure ANOVA was used for the illustrated across-region interaction effect.* ***B.*** *Each dot corresponds to the data of one session from one participant. The x-axis corresponds to the mean $NSM_{ROI} \times NSM_{ROI}$ correlation, averaged across the inferior temporal lobe as in Figure 6A. The y-axis corresponds to the distributed RSA effect, computed as in Figure 3. The dashes show the negatively sloped line of best fit. *, $p < .05$; **, $p < .01$.*

Finally, we returned to the inferior temporal lobe and assessed whether stronger large-scale RSA effects are indeed linked to lower $NSM_{ROI} \times NSM_{ROI}$ correlations (i.e., does higher synergy parallel lower redundancy). For each participant and separately for each of the four study stages, we calculated the inferior temporal large-scale semantic RSA effect from **Figure 3C** along with the mean inferior temporal lobe $NSM_{ROI} \times NSM_{\mathrm{ROI}}$ from **Figure 5A**. Linear regression showed that the former indeed significantly predicts the latter (β = -.55 [-.91, -.19], p = .003; **Figure 7B**). This effect only emerged for analyses of the inferior temporal lobe and semantics, being insignificant when considering perceptual representation in the inferior temporal lobe (β = -.22 [-.54, .11], p = .19) or semantic representation in the occipital lobe (β = -.09 [-.44, .26], p = .61). Hence, uniquely strong distributed representation in the inferior temporal lobe relates to the area's uniquely low level of redundancy between its regions' information processing.

3. Discussion

The present research aimed to clarify the organization of computation and the size of population codes across the brain. The research yielded three conclusions. *First,* we performed RSA at different spatial scales, which indicated that, in the occipital lobe, perceptual processing is split across local population codes. In contrast, inferior temporal information processing is distributed in a way not found anywhere else in the cortex, such that distant inferior temporal regions synergistically interact to encode semantics. *Second*, we assessed the relationship between inferior temporal regions and found that the lobe is uniquely characterized by its constituent regions specializing in representing and distinguishing different objects. *Third*, we conducted analyses linking these different levels and showed that individual regions specializing in the objects they encode promotes distributed representation of semantic information. Overall, these results describe how the architecture of object processing in the ventral stream supports efficient information processing.

Perceptual information was most strongly represented in the occipital lobe and best modeled as local codes (**Figures 3 & 4**). Our analyses defined perceptual information based on the first pooling layer of a convolutional neural network, which captures features like color, shape, and texture [17]. Prior work has shown how different visual functions are relegated to distinct areas.[18,19] Our findings add that perceptual coding is not synergistic between regions, suggesting that communication within the occipital lobe is fairly minimal. Indeed, V4 lesions destroy color processing but leave motion processing largely intact,[20] while V5 lesions do the opposite.[21] It is interesting to speculate on why localized communication within the occipital lobe may be beneficial. In principle, more interconnected networks can encode more information.[22] Yet, as the computer vision literature shows, fully connected layers are often sub-optimal at early stages, and

restricted convolutional neurons ease training.[23] In the human brain, a predisposition for modularity in early vision may similarly benefit perceptual development.

Contrasting the occipital patterns, semantic representation in the inferior temporal lobe involved diffuse interactions, evidenced by coding synergy between ROIs (**Figures 3 & 4**). Interestingly, prior work mapping semantic coding across the brain more often yields effects in parietal, frontal, or lateral temporal lobes.[24–27] Further, research on how information is linked across sensory modalities has found limited inferior temporal patterns.[28–30] Yet, when we modeled large-scale representations that account for synergistic interactions between regions, the inferior temporal lobe's semantic profile yielded strong effects. This marked distribution in representation was not seen in any other wide-scale brain area, meaning that this is a powerful angle through which to see the inferior temporal lobe specifically. Potentially, this functionality relates to the inferior temporal lobe's central role in semantic memory encoding and retrieval. Research on superposition shows how one large network can better support a large semantic space than multiple smaller networks [31]. For instance, a network with 1000 neurons can represent more concepts than two 500-neuron networks, and thus a large interconnected structure may be particularly useful in the inferior temporal lobe.

We additionally analyzed how a distributed system may organize its constituents to minimize redundancy. This involves correlating the representational signatures of neighboring regions (**Figures 5 & 6**). Within the inferior temporal lobe, NSM × NSM and NPS × NPS correlations are low between regions – markedly decreased compared to any other brain system. That is, inferior temporal ROIs highly vary in which objects they predominantly encode (NPS results) and which pairs of objects they predominantly distinguish (NSM results). This speaks to the synergy results on the inferior temporal lobe functioning as one or a small number of large computational units with great interconnection. Additionally, this finding aligns with prominent

theories of inferior temporal organization focusing on its many category-specialized structures – e.g., for faces, locations, body parts, numbers, and letters.[32,33] Our stimuli were diverse and not organized around known categories, and the present results suggest heterogeneity and specialization are general properties of the inferior temporal lobe.

These ROI-level properties support synergistic interactions and wider representation per our final correlations between these levels (**Figure 7B**). Our "integration account" of the inferior temporal lobe is not at odds with existing category-specialization theories. Instead, these are different sides of the same coin. For instance, rather than viewing the FFA and PPA as independently processing faces and locations respectively, we can conceptualize them as coordinating to encode both faces and places collectively. This is a special form of specialization, which differs from how the occipital lobe processes information in small population codes. Indeed, whereas the V4's involvement in color coding and V5's involvement in motion processing are narrow, the FFA plays a rich role in processing objects more generally.[34,35] This is consistent with the FFA contributing to collective representations rather than operating in isolation.

In summary, the brain is a highly distributed representational system. Yet, the nature of this distribution varies according to tractable spatial and informational patterns. In some areas, distributed processing involves segmenting information processing across different regions set off to operate fairly independently. In other areas, distributed processing reflects an area operating as a synergistic collective unit of information processing, where the loss of a region shatters a large-scale population code. This dimension is particularly valuable for understanding the ventral stream. Incoming information begins as local perceptual codes in the occipital lobe and advances into distributed semantic representation in the inferior temporal lobe. Future studies may also consider how representations may not be entirely organized into contiguous voxels but may emerge from synergistic interactions within a spatially disparate (non-contiguous) network. Further studies may

also benefit mapping distribution more precisely than just along small or large scales, potentially developing atlases based on information synergy and redundancy properties. Broadly, the brain varies in the size of its computational units, and understanding this dimension is key to seeing how neural structures promote efficient information coding.

4. Methods

4.1. Participants

For this four-stage study, participants were recruited from the local community and screened to ensure they were native or fluent English speakers and no history of brain injury or mild cognitive impairment. In total, 38 younger adults (aged 18-30) and 38 older adults (aged 65-85) participated in the study. Of these, 16 people did not complete at least one of the four scans – e.g., due to withdrawing from the study after the first stage. The final set of 60 participants who provided fMRI data for all four stages consisted of a younger adult subset (N = 35, M_{age} = 22.8 [SD = 3.3], 64% female, 36% male) and an older adult subset (N = 31, M_{age} = 71.5 [SD = 4.5], 64% female, 36% male). There were no exclusions. The research was approved by the Duke University Institutional Review Board.

4.2. Task design

Participants completed four stages of a study design on object processing, semantics, and memory. In each stage, participants viewed stimuli for the same set of 114 objects, always separated across three runs. In the first stage, participants completed a task wherein each trial they were shown an object (e.g., a tractor) with a label (e.g., "tractor"). Participants used a 4-point scale to rate how accurately the label described the object (1="*does not describe the object*"; 4="*exact description*"). On average, the task was designed to elicit high ratings (M = 3.60).

One week later, participants completed the second stage, wherein each trial they were shown a familiar scene (e.g., a farm) followed by an object (e.g., a tractor). Participants reported how likely it would be to find the object in the scene using a 4-point scale (1 = "*Very unlikely*", 4 = "*Very*

likely”). The scenes and objects were organized such that the scene-object pairs were distinctly incongruent in one-third of trials (mean rating = 1.28), were unrelated in one-third of trials (M = 2.11), and were distinctly congruent in one-third of trials (M = 3.67). The scene-object pairs were counterbalanced across participants such that every scene and every object appeared in all three conditions.

One day after the second stage, participants completed the third stage (conceptual retrieval task) and fourth stage (visual retrieval task) sequentially. For the conceptual retrieval task, in each trial, participants were presented 144 words; 114 were the object labels from the first stage and 30 were new ones. For each word, participants indicated whether they previously saw it before using a 4-point scale (1 = “*definitely new*”, 2 = “*probably new*”, 3 = “*probably old*”, 4 = “*definitely old*”). In general, most objects yielded a hit response (hit rate [3- or 4-response to old images]: M = .76, SD = .16). For the visual retrieval task, participants were shown 126 images of objects. Among these, 96 images were the exact object images shown in the first two stages, 18 images were similar variations of those original ones (e.g., a picture of a blue tractor while the first two stages actually showed a red one), and 12 images were totally new. Participants responded to each image either: (i) “old”, (ii) “similar”, or (iii) “new”. As with the conceptual retrieval task, most objects yielded hits (old hit rate: M = 0.64, SD = 0.17; similar hit rate: M = 0.49, SD = 0.21).

4.3. MRI and preprocessing

4.3.1. MRI acquisition

MRI data were collected using a General Electric 3T MR750 scanner and an 8-channel head coil. Anatomical images were acquired using a T1-weighted echo-planar sequence (96 slices at $0.9\times0.9\times1.9$ mm^3). Functional images were acquired using an echo-planar imaging sequence (repetition time = 2000 ms, echo time = 30 ms, field of view = 19.2 cm, 36 oblique slices with voxel dimensions of $3\times3\times3$ mm). Stimuli were projected onto a mirror at the back of the scanner bore,

and responses were recorded using a four-button fiber-optic response box (Current Designs, Philadelphia, PA, USA). Functional resting-state images were collected from the participants using the same parameters (210 volumes, 7 minutes). The BOLD timeseries were resampled into standard space with a spatial resolution of 2×2×2 mm^3 or 97×115×97 voxels.

The below descriptions of anatomical and functional preprocessing were automatically generated by fMRIPrep with the express intention that users should copy and paste this text into their manuscripts unchanged.

4.3.2. Anatomical data preprocessing

A total of two T1-weighted (T1w) images were found within the input BIDS dataset. All of them were corrected for intensity non-uniformity (INU) with N4BiasFieldCorrection [36] distributed with ANTs 2.3.3.[37] The T1w-reference was then skull-stripped with a Nipype implementation of the antsBrainExtraction.sh workflow (from ANTs), using OASIS30ANTs as a target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM), and gray-matter (GM) was performed on the brain-extracted T1w using fast (FSL 6.0.5.1).[38] An anatomical T1w-reference map was computed after registration of two T1w images (after INU-correction) using mri_robust_template (FreeSurfer 7.3.2).[39]

Brain surfaces were reconstructed using recon-all (FreeSurfer 7.3.2),[40] and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of Mindboggle.[41] Volume-based spatial normalization to one standard space (MNI152NLin2009cAsym) was performed through nonlinear registration with antsRegistration (ANTs 2.3.3), using brain-extracted versions of both T1w reference and the T1w template. The following template was selected for spatial normalization and accessed with TemplateFlow (23.0.0):[42] ICBM 152 Nonlinear Asymmetrical template version 2009c.[43]

4.3.3. Functional data preprocessing

For each of the seven BOLD runs found per participant (across all tasks and sessions), the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using MCFLIRT (FSL 6.0.5.1, Jenkinson et al. 2002). BOLD runs were slice-time corrected to 0.972s (0.5 of slice acquisition range 0.00s-1.94s) using 3dTshift from AFNI.[44]

The BOLD time series (including slice-timing correction when applied) were resampled onto their original, native space by applying the transforms to correct for head motion. These resampled BOLD time-series will be referred to as preprocessed BOLD in original space, or just preprocessed BOLD. The BOLD reference was then co-registered to the T1w reference using bbregister (FreeSurfer) which implements boundary-based registration.[45] Co-registration was configured with six degrees of freedom.

Several confounding time series were calculated based on the preprocessed BOLD: framewise displacement (FD), DVARS, and three region-wise global signals. FD was computed using two formulations following Power et al. (absolute sum of relative motions)[46] and Jenkinson (relative root mean squared displacement between affines).[47] FD and DVARS are calculated for each functional run, both using their implementations in Nipype (following the definitions by Power et al.[46]). The three global signals are extracted within the CSF, the WM, and the whole-brain masks. Principal components are estimated after high-pass filtering the preprocessed BOLD time series (using a discrete cosine filter with 128s cut-off) for the two CompCor variants: temporal (tCompCor) and anatomical (aCompCor). tCompCor components are then calculated from the top 2% variable voxels within the brain mask. For aCompCor, three probabilistic masks (CSF, WM, and

combined CSF+WM) are generated in anatomical space. The implementation differs from that of Behzadi et al. in that instead of eroding the masks by two pixels on BOLD space, a mask of pixels that likely contain a volume fraction of GM is subtracted from the aCompCor masks. This mask is obtained by dilating a GM mask extracted from the FreeSurfer's aseg segmentation, and it ensures components are not extracted from voxels containing a minimal fraction of GM. Finally, these masks are resampled into BOLD space and binarized by thresholding at 0.99 (as in the original implementation). Components are also calculated separately within the WM and CSF masks. For each CompCor decomposition, the k components with the largest singular values are retained, such that the retained components' time series are sufficient to explain 50% of variance across the nuisance mask (CSF, WM, combined, or temporal). The remaining components are dropped from consideration.

The head-motion estimates calculated in the correction step were also placed within the corresponding confounds file. The confound time series derived from head motion estimates and global signals were expanded with the inclusion of temporal derivatives and quadratic terms for each.[48] Additional nuisance time series are calculated by means of principal components analysis of the signal found within a thin band (crown) of voxels around the edge of the brain, as proposed by Patriat, Reynolds, and Birn.[49] The BOLD time series were resampled into standard space, generating a preprocessed BOLD run in MNI152NLin2009cAsym space. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. All resampling can be performed with a single interpolation step by composing all the pertinent transformations (i.e., head-motion transform matrices, susceptibility distortion correction when available, and co-registrations to anatomical and output spaces). Gridded (volumetric) resampling was performed using antsApplyTransforms (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels.[50] Non-gridded (surface) resampling was performed

using mri_vol2surf (FreeSurfer). Many internal operations of fMRIPrep use Nilearn 0.9.1,[51] mostly within the functional processing workflow. For more details of the pipeline, see the section corresponding to workflows in fMRIPrep's documentation.

4.3.4. Single-trial activity modeling

Analyses required measuring each voxel's BOLD response in each trial. This was done using first-level general linear models with the Least Squares Separate approach by Mumford et al.,[52] which involves fitting a separate regression for each trial. The regression included a boxcar signal spanning the trial's object presentation period, and this regressor's coefficient represents the trial's BOLD response. The regression also included a boxcar signal covering the presentation time of all other stimuli (i.e., every other object and every scene). Additionally, the linear models included six translation/rotation regressors, three other covariates for head motion (FD, DVARS, and RSMD), and covariates for mean global, white-matter, and cerebrospinal signals. Hence, for all four tasks of each of the 60 participants (240 scans), 114 three-dimensional beta coefficient volumes were defined. These were submitted to RSA.

4.4. Representational similarity analysis

4.4.1. Generating the model-based representational similarity matrices

RSA was conducted with respect to both perceptual and semantic features by defining a model RSM for each of these two types of information. The creation of each model RSM involved defining a perceptual or semantic feature vector for each object. To produce the perceptual vector for a given trial, every object image was submitted to the VGG16 convolutional neural network.[17] For each object, all the activation values from neurons of the first VGG16 layer were extracted. For each object, its activation values (224x224x64 values) were flattened into a single vector (length = 3,211,264). The 114 objects' massive vectors were reduced to smaller vectors of length = 114 using PCA. Then, the perceptual RSM was generated via across-trial correlations of these vectors. For the

semantic RSM, each trial's vector was generated by submitting its object's label to a pre-trained word2vec model;[53] the public "word2vec-google-news-300" version from the *Gensim* Python package. The word2vec model yielded a 300-length vector for each object, which was correlated between trials to generate the semantic RSM.

4.4.2. Computing neural similarity matrices

All of the RSA done in the report was ROI-based. Analyses parcellated the brain into 210 neocortical ROIs using the Brainnetome Atlas.[54] The analyses also involved organizing a subset of the ROIs into four large anatomical areas based on Brainnetome labels: the *occipital lobe* (22 ROIs; cuneus & occipital gyrus; 139 cm^3, including white matter), the *inferior temporal lobe* (32 ROIs; fusiform, parahippocampal, inferior temporal gyri, & anterior temporal lobe [the anterior temporal lobe is not an official Brainnetome label, but is a custom one used by our group]; 138 cm^3), the *parietal lobe* (30 ROIs; precuneus, inferior parietal lobule, & super parietal lobule; 192 cm^3), and the *prefrontal cortex* (52 ROIs; orbital, inferior, middle, & superior frontal gyri; 288 cm^3). For each ROI separately, a small-scale NSM was computed via across-trial voxel-wise correlations. These were averaged to compute an area's mean-small NSM. In addition, for each area, a large-scale voxelwise NSM was computed via across-trial correlations while pooling constituent ROIs (i.e., effectively combining the constituents into one large ROI) (used for the analyses in **Figure 3**). Also, a different large-scale NSM was computed for each area via across-trial correlations with respect to the areas' ROIs' mean activations (i.e., effectively treating each ROI as a single voxel; used for the **Figure 4** analyses).

RSA was then performed either using Spearman correlations or linear regressions. For the former, second-order RSM × NSM Spearman correlations were computed for the different model RSMs and NSMs, done separately for each participant and each of the task stages. The correlation coefficients were then averaged by participant across task stages and then submitted to group-level

analyses (paired t-tests or repeated-measures ANOVAs). For the regression analysis, a model RSM was regressed on multiple NSMs. The resulting beta coefficients were extracted, averaged by participant across task stages and then submitted to group-level analyses.

4.5. Redundancy analysis correlations

We pursued a series of analyses based on ROI-to-ROI relationships to examine patterns linked to when multiple ROIs encode overlapping/redundant information. We also sought to evaluate how this perspective compares to traditional functional connectivity analyses to confirm this approach adds new information above and beyond traditional means of assessing region-to-region interactions. First, analyses examined the similarity in small-scale NSMs, correlating these between ROIs – also referred to as model-free representational connectivity.[7] The ROI-ROI correlations were calculated for each of the four stages and averaged. Second, analyses examined item specializing by calculating item-wise neural pattern similarity for each trial (also called the exemplar discriminability index);[55] computing the neural pattern similarity of an ROI for a given object involved measuring the similarity of the voxel-wise response across two stages, then subtracting the average level of similarity (average correlation between the response to the object in one stage and the response to every other object in another stage). Computing neural pattern similarity for each object yielded a 114-length vector per ROI, as there were 114 objects in the task. This vector was correlated between ROIs. Third, analyses examined functional connectivity as the correlation between ROI's trial-wise activations (averaged across voxels); often referred to as beta series functional connectivity.[56]

Subsequent redundancy analyses modeled NSMs for every 2 × 2 × 2 cube of voxels in the neocortex. Then correlations were performed between these cubes and neighboring cubes within 20 mm. Although not otherwise reported, analyses were also performed with respect to neighbors within 5 mm, 10 mm, or 30 mm, and all the conclusions put forth in the Results remain unchanged.

4.6. Linking large-scale representation and ROI-ROI correlations

The final set of analyses focused just on the inferior temporal lobe, as it was the only area to show large-scale representations. These analyses attempted to link those large representations to ROI-level information processing. For this, the analyses used the $NSM_{ROI} \times NSM_{ROI}$ correlation matrices generated just above. For each participant and each of the four stages, the mean correlation was computed among ROIs within the inferior temporal lobe. Given that there are 32 inferior temporal ROIs and thus 496 possible pairs of ROIs, the mean $NSM_{ROI} \times NSM_{ROI}$ correlation amounts to averaging across those 496 pairs. Along with computing this mean, for each participant and each of the four stages, the degree of large-scale representation was also calculated. This was done using the initial technique, which measured voxel-wise correlations while pooling every inferior temporal lobe ROIs' voxels and then performed a second-order correlation with a semantic model RSM. Thus, for the 60 participants and four stages, one measure of the $NSM_{ROI} \times NSM_{ROI}$ correlations and one measure of the degree of large-scale representation were computed. These 240 data points were submitted to a multilevel regression, linking the two measures: *Large scale RSA ~ 1 + ROI corr + (1 | participant)*. The significance of the fixed effect linking the measures was evaluated using the *lmerTest* package.[57] As points of comparison, the analyses of this paragraph were also performed for the occipital lobe and with regard to perceptual RSA effects.

Data availability. The data collected will be shared upon request via email to the corresponding author, Paul C. Bogdan, or principal investigators Roberto Cabeza and Simon W. Davis. Due to IRB restrictions, the collected data cannot be uploaded to a public repository.

References

1. Alexander, M. P., Naeser, M. A. & Palumbo, C. Broca's area aphasias: aphasia after lesions including the frontal operculum. *Neurology* **40**, 353–353 (1990).
2. Wernicke, C. The aphasic symptom-complex: a psychological study on an anatomical basis. *Archives of Neurology* **22**, 280–282 (1970).
3. Catani, M. & Ffytche, D. H. The rises and falls of disconnection syndromes. *Brain* **128**, 2224–2239 (2005).
4. Geschwind, N. Disconnexion syndromes in animals and man. *Brain* **88**, 585–585 (1965).
5. Mah, Y.-H., Husain, M., Rees, G. & Nachev, P. Human brain lesion-deficit inference remapped. *Brain* **137**, 2522–2531 (2014).
6. Haxby, J. V., Connolly, A. C. & Guntupalli, J. S. Decoding neural representational spaces using multivariate pattern analysis. *Annual review of neuroscience* **37**, 435–456 (2014).
7. Huang, S., De Brigard, F., Cabeza, R. & Davis, S. W. Connectivity analyses for task-based fMRI. *Physics of Life Reviews* (2024).
8. Williams, N. & Henson, R. N. Recent advances in functional neuroimaging analysis for cognitive neuroscience. *Brain and Neuroscience Advances* **2**, 2398212817752727 (2018).
9. Luppi, A. I., Rosas, F. E., Mediano, P. A., Menon, D. K. & Stamatakis, E. A. Information decomposition and the informational architecture of the brain. *Trends in Cognitive Sciences* (2024).
10. Seymour, K., Clifford, C. W., Logothetis, N. K. & Bartels, A. The coding of color, motion, and their conjunction in the human visual cortex. *Current Biology* **19**, 177–183 (2009).
11. Barlow, H. B. Possible principles underlying the transformation of sensory messages. *Sensory communication* **1**, 217–233 (1961).
12. Kriegeskorte, N., Mur, M. & Bandettini, P. A. Representational similarity analysis-connecting the branches of systems neuroscience. *Frontiers in systems neuroscience* **2**, 4 (2008).

13. Kriegeskorte, N., Cusack, R. & Bandettini, P. How does an fMRI voxel sample the neuronal activity pattern: compact-kernel or complex spatiotemporal filter? *Neuroimage* **49**, 1965–1976 (2010).

14. Clarke, A. & Tyler, L. K. Object-specific semantic coding in human perirhinal cortex. *Journal of Neuroscience* **34**, 4766–4775 (2014).

15. Davis, S. W. *et al.* Visual and semantic representations predict subsequent memory in perceptual and conceptual memory tests. *Cerebral Cortex* **31**, 974–992 (2021).

16. Martin, C. B., Douglas, D., Newsome, R. N., Man, L. L. & Barense, M. D. Integrative and distinctive coding of visual and conceptual object features in the ventral visual stream. *elife* **7**, e31873 (2018).

17. Simonyan, K. Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556* (2014).

18. Cavina-Pratesi, C., Kentridge, R., Heywood, C. & Milner, A. Separate channels for processing form, texture, and color: evidence from fMRI adaptation and visual object agnosia. *Cerebral cortex* **20**, 2319–2332 (2010).

19. Rolls, E. T., Deco, G., Huang, C.-C. & Feng, J. Multiple cortical visual streams in humans. *Cerebral Cortex* **33**, 3319–3349 (2023).

20. Zeki, S. A century of cerebral achromatopsia. *Brain* **113**, 1721–1777 (1990).

21. Newsome, W. T. & Pare, E. B. A selective impairment of motion perception following lesions of the middle temporal visual area (MT). *Journal of Neuroscience* **8**, 2201–2211 (1988).

22. Roudi, Y. & Latham, P. E. A balanced memory network. *PLoS computational biology* **3**, e141 (2007).

23. Krizhevsky, A., Sutskever, I. & Hinton, G. E. Imagenet classification with deep convolutional neural networks. *Advances in neural information processing systems* **25**, (2012).

24. Carota, F., Nili, H., Pulvermüller, F. & Kriegeskorte, N. Distinct fronto-temporal substrates of distributional and taxonomic similarity among words: evidence from RSA of BOLD signals. *NeuroImage* **224**, 117408 (2021).

25. Huth, A. G., De Heer, W. A., Griffiths, T. L., Theunissen, F. E. & Gallant, J. L. Natural speech reveals the semantic maps that tile human cerebral cortex. *Nature* **532**, 453–458 (2016).

26. Kidder, A., Silson, E. H., Nau, M. & Baker, C. I. Distributed Cortical Regions for the Recall of People, Places, and Objects. *eneuro* **12**, (2025).

27. Meersmans, K. *et al.* Representation of associative and affective semantic similarity of abstract words in the lateral temporal perisylvian language regions. *Neuroimage* **217**, 116892 (2020).

28. Fernandino, L., Tong, J.-Q., Conant, L. L., Humphries, C. J. & Binder, J. R. Decoding the information structure underlying the neural representation of concepts. *Proceedings of the National Academy of Sciences* **119**, e2108091119 (2022).

29. Fernandino, L. *et al.* Concept representation reflects multimodal abstraction: A framework for embodied semantics. *Cerebral cortex* **26**, 2018–2034 (2016).

30. Tong, J. *et al.* A distributed network for multimodal experiential representation of concepts. *Journal of Neuroscience* **42**, 7121–7130 (2022).

31. Elhage, N. *et al.* Toy models of superposition. *arXiv preprint arXiv:2209.10652* (2022).

32. Kanwisher, N. Functional specificity in the human brain: a window into the functional architecture of the mind. *Proceedings of the national academy of sciences* **107**, 11163–11170 (2010).

33. Pitcher, D., Charles, L., Devlin, J. T., Walsh, V. & Duchaine, B. Triple dissociation of faces, bodies, and objects in extrastriate cortex. *Current Biology* **19**, 319–324 (2009).

34. Haist, F., Lee, K. & Stiles, J. Individuating faces and common objects produces equal responses in putative face-processing areas in the ventral occipitotemporal cortex. *Frontiers in human neuroscience* **4**, 181 (2010).

35. Zachariou, V., Safiullah, Z. N. & Ungerleider, L. G. The fusiform and occipital face areas can process a nonface category equivalently to faces. *Journal of cognitive neuroscience* **30**, 1499–1516 (2018).

36. Tunison, E., Sylvain, R., Sterr, J., Hiley, V. & Carlson, J. M. No money, no problem: enhanced reward positivity in the absence of monetary reward. *Frontiers in human neuroscience* **13**, 41 (2019).

37. Avants, B. B., Epstein, C. L., Grossman, M. & Gee, J. C. Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. *Medical image analysis* **12**, 26–41 (2008).

38. Zhang, Y., Brady, M. & Smith, S. Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. *IEEE transactions on medical imaging* **20**, 45–57 (2001).

39. Reuter, M., Rosas, H. D. & Fischl, B. Highly accurate inverse consistent registration: a robust approach. *Neuroimage* **53**, 1181–1196 (2010).

40. Dale, A. M., Fischl, B. & Sereno, M. I. Cortical surface-based analysis: I. Segmentation and surface reconstruction. *Neuroimage* **9**, 179–194 (1999).

41. Klein, A. *et al.* Mindboggling morphometry of human brains. *PLoS computational biology* **13**, e1005350 (2017).

42. Ciric, R. *et al.* TemplateFlow: FAIR-sharing of multi-scale, multi-species brain models. *Nature Methods* **19**, 1568–1571 (2022).

43. Fonov, V. S., Evans, A. C., McKinstry, R. C., Almli, C. R. & Collins, D. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *NeuroImage* **47**, S102 (2009).

44. Cox, R. W. & Hyde, J. S. Software tools for analysis and visualization of fMRI data. *NMR in Biomedicine: An International Journal Devoted to the Development and Application of Magnetic Resonance In Vivo* **10**, 171–178 (1997).

45. Greve, D. N. & Fischl, B. Accurate and robust brain image alignment using boundary-based registration. *Neuroimage* **48**, 63–72 (2009).

46. Power, J. D. *et al.* Methods to detect, characterize, and remove motion artifact in resting state fMRI. *Neuroimage* **84**, 320–341 (2014).

47. Jenkinson, M., Bannister, P., Brady, M. & Smith, S. Improved optimization for the robust and accurate linear registration and motion correction of brain images. *Neuroimage* **17**, 825–841 (2002).

48. Satterthwaite, T. D. *et al.* An improved framework for confound regression and filtering for control of motion artifact in the preprocessing of resting-state functional connectivity data. *Neuroimage* **64**, 240–256 (2013).

49. Patriat, R., Reynolds, R. C. & Birn, R. M. An improved model of motion-related signal changes in fMRI. *Neuroimage* **144**, 74–82 (2017).

50. Lanczos, C. Evaluation of noisy data. *Journal of the Society for Industrial and Applied Mathematics, Series B: Numerical Analysis* **1**, 76–85 (1964).

51. Abraham, A. *et al.* Machine learning for neuroimaging with scikit-learn. *Frontiers in neuroinformatics* **8**, 14 (2014).

52. Mumford, J. A., Turner, B. O., Ashby, F. G. & Poldrack, R. A. Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. *Neuroimage* **59**, 2636–2643 (2012).

53. Mikolov, T., Chen, K., Corrado, G. & Dean, J. Efficient estimation of word representations in vector space. *arXiv preprint arXiv:1301.3781* (2013).

54. Fan, L. *et al.* The human brainnetome atlas: a new brain atlas based on connectional architecture. *Cerebral cortex* **26**, 3508–3526 (2016).

55. Nili, H., Walther, A., Alink, A. & Kriegeskorte, N. Inferring exemplar discriminability in brain representations. *Plos one* **15**, e0232551 (2020).

56. Rissman, J., Gazzaley, A. & D'Esposito, M. Measuring functional connectivity during distinct stages of a cognitive task. *Neuroimage* **23**, 752–763 (2004).

57. Kuznetsova, A., Brockhoff, P. B. & Christensen, R. H. B. lmerTest package: tests in linear mixed effects models. *J Stat Softw* **82**, (2017).