REVIEW ARTICLE | FOCUS
https://doi.org/10.1038/s41593-020-00726-z

nature neuroscience

Challenges and future directions for representations of functional brain organization

Janine Bijsterbosch[1,2] ✉, Samuel J. Harrison[2,3], Saad Jbabdi[2], Mark Woolrich[4], Christian Beckmann[5], Stephen Smith[2] and Eugene P. Duff[2,6] ✉

A key principle of brain organization is the functional integration of brain regions into interconnected networks. Functional MRI scans acquired at rest offer insights into functional integration via patterns of coherent fluctuations in spontaneous activity, known as functional connectivity. These patterns have been studied intensively and have been linked to cognition and disease. However, the field is fractionated. Diverging analysis approaches have segregated the community into research silos, limiting the replication and clinical translation of findings. A primary source of this fractionation is the diversity of approaches used to reduce complex brain data into a lower-dimensional set of features for analysis and interpretation, which we refer to as brain representations. In this Primer, we provide an overview of different brain representations, lay out the challenges that have led to the fractionation of the field and that continue to form obstacles for convergence, and propose concrete guidelines to unite the field.

The fundamental brain organizing principles of segregation into functionally specialized areas and integration into highly connected macroscale brain networks form the basis of neuroscientific research into brain structure and function[1]. Over recent decades, the ability to noninvasively measure inherent activity fluctuations from the whole brain using resting-state functional MRI (rfMRI) has offered many insights into this macroscale functional organization. In rfMRI, similarities (between different brain regions) in the spontaneous fluctuations of the blood oxygenation level dependent (BOLD) signals—termed functional connectivity—can be used to explore intrinsic functional relationships across brain areas and examine how these vary across healthy and disease states and through the population[2]. However, the field of rfMRI is fractionated, with ongoing debates regarding preprocessing strategies and brain parcellations, as well as extensive divergence of post-processing analysis methods and endpoints[3,4]. These challenges greatly impact the ability of the field to achieve reproducible progress.

A key source of these issues is the challenge of brain representation. Like many modern scientific fields, rfMRI generates large amounts of data per participant, comprising ongoing activity measurements from tens of thousands of voxels over a period of up to an hour. A key task of analysis is to distill the enormous complexity of the measured brain activity into an accessible and interpretable form. We use the term 'brain representation' here to refer to the combined methodological steps that are taken to derive a lower dimensional set of features from an individual's rfMRI data set for subsequent analysis and interpretation. Brain representations are multifaceted descriptions of the acquired rfMRI data that often encompass both a spatial definition of brain units (i.e., parcellation) and a summary measure that extracts interpretable features at the level of the brain units (for example, pairwise correlation between brain unit time series, discussed further in "A primer on brain representations" below). As the determinant of the building blocks for further analyses, the choice of how to represent brain data fundamentally underpins descriptions of brain connectivity and organization.

Representation of the brain is often considered a task of mapping, aiming to delineate boundaries of functionally and neuroanatomically distinct areas of the neural tissue[1,5]. However, the task of brain representation is broader than mapping, encompassing a more comprehensive range of representational forms and addressing additional aspects of how data are transformed into these representations. The goal of this article is to provide a primer on the representational challenges of rfMRI, with the intention of improving consensus and reproducibility in the field.

The lack of consensus on brain representations results in part from a lack of knowledge of the true underlying functional organization of the brain. In the absence of this ground-truth, further divergence arises from differences in opinion about what we want to achieve. The organization of the brain is multifaceted, dynamic and hierarchical, leading to many complexities and trade-offs for the question of brain representation. On the one hand, brain representations should strive to accurately reflect neural units to achieve biological interpretability[6]. On the other hand, the accuracy of clinically relevant predictions derived from representations may be prioritized[7]. A standardized brain representation that is generalizable across groups, ages and clinical populations would provide a common language for the communication of findings, facilitating replication and aggregation of results. However, standardization would come at the cost of optimization for the specific research question, study population, experimental protocol or individual participant. In this Primer, we aim to clarify these issues and provide guidelines for the choice of brain representation.

Approaches to representing rfMRI data are diverse, and we therefore start with a short primer on a number of distinct families of brain representations. Despite their differences, each family of representations faces similar fundamental challenges, ranging from subtle mathematical biases and confounds to the interpretability of their final outputs. The section "Challenges for brain representations"

[1]Mallinckrodt Institute of Radiology, Washington University in St Louis, Saint Louis, MO, USA. [2]Centre for Functional MRI of the Brain (FMRIB), Wellcome Centre for Integrative Neuroimaging, Nuffield Department of Clinical Neurosciences, University of Oxford. John Radcliffe Hospital, Oxford, UK. [3]Translational Neuromodeling Unit, University of Zurich & ETH Zurich, Zurich, Switzerland. [4]Oxford Centre for Human Brain Activity (OHBA), Wellcome Centre for Integrative Neuroimaging, Department of Psychiatry, University of Oxford, Warneford Hospital, Oxford, UK. [5]Donders Institute and Department of Cognitive Neurosciences, Radboud University Medical Centre, Nijmegen, The Netherlands. [6]Department of Paediatrics, University of Oxford, John Radcliffe Hospital, Oxford, UK. ✉e-mail: Janine.Bijsterbosch@wustl.edu; eugene.duff@ndcn.ox.ac.uk

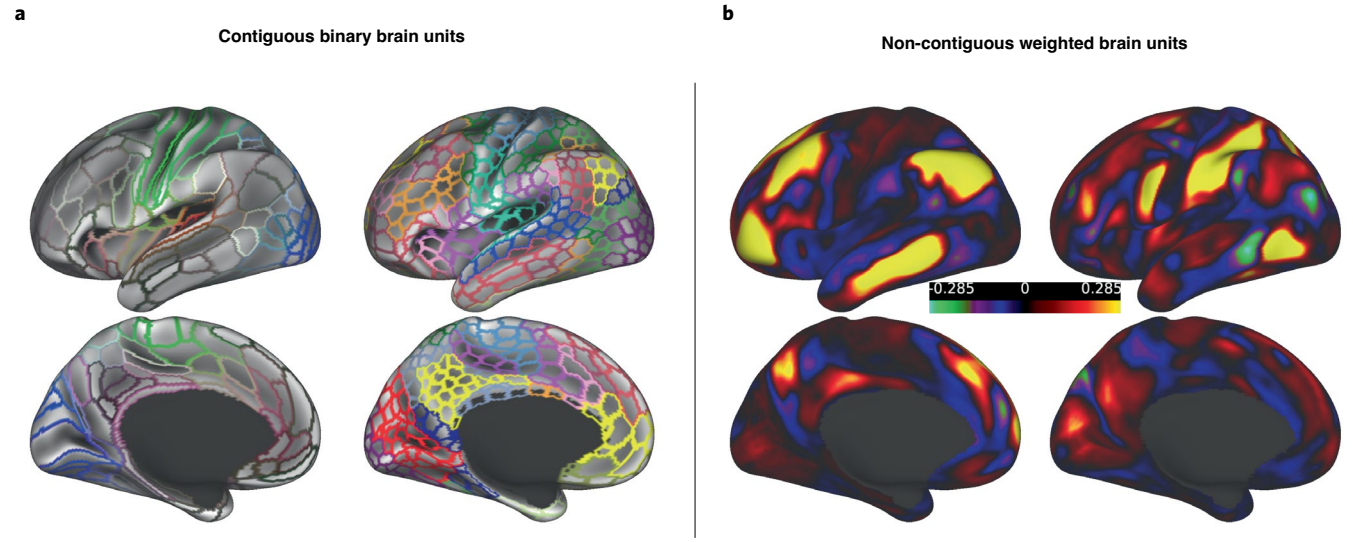

Fig. 1 | Example brain representations. a, Example binary parcellations adapted from Glasser[10] (left; Nature Publishing Group) and Schaefer[18] (right; Oxford Univ. Press). **b**, Example weighted resting-state networks obtained from a 50-dimensional independent component analysis (ICA) decomposition (publicly released Human Connectome Project maps estimated from 1,003 participants).

surveys and dissects these issues, which may be addressed in a variety of ways, contributing to the dispersion of brain representations. Finally, we put forward a set of explicit suggestions and guidelines to resolve the fractionation of the field and achieve a consensus that it is expected to improve reproducibility and collaboration ("Recommendations and future directions for brain representations"). As part of our recommendations, we advocate for the use of state-of-the-art preprocessing methods to avoid unnecessary discordance resulting from outdated practices and for improved transparency in reporting to inform and encourage debate about similarities and differences in brain representations. While we focus on rfMRI brain representations, many of the issues that we discuss are shared with other measurement domains such as diffusion MRI and local field potentials.

A primer on brain representations

Brain representations can take a variety of approaches to reduce the complexity of measured BOLD data into a set of features for analysis. A majority of brain representations identify (i) a low-dimensional set of brain units (for example, spatial parcels) and (ii) a summary measure that is applied at the level of brain units (for example, pairwise correlation between brain unit time series). Together, these reduce the rfMRI data down to the features used in subsequent statistical or prediction analyses. We use the term 'brain unit' to indicate any spatially defined neural entity that can be treated as a basic functional processing unit. The concept of a brain unit therefore generalizes beyond the rfMRI domain, with strong parallels to for example, Hebbian cell assemblies[8]. We define the 'summary measure' of a brain representation as a method for computing features, defined in relation to brain units, from the rfMRI data. The computed features vary freely across participants and/or conditions and are used to answer the research question. The summary measures of a brain representation are therefore relatively domain-specific and informed by the measurement type. A smaller number of brain representations do not use brain units and summary measures and instead estimate features that represent complex spatiotemporal patterns of activity (see section "Complex spatiotemporal brain representations").

Defining a brain unit. As a whole-brain noninvasive imaging modality, the spatial measurement resolution of rfMRI easily reaches $2 \times 2 \times 2$ mm in modern scans. This results in about 100,000 voxels within gray matter across the adult human brain. In rfMRI, these voxels (or vertices when representing the cortical gray matter surface as a tessellated mesh[9]) are the smallest measurable brain units. However, the voxel or vertex unit, while ideally chosen with reference to brain anatomy such as cortical thickness[10], does not represent any particular level of the neuroanatomical hierarchy. It is therefore common to group voxels or vertices together into a smaller set of brain units to achieve a meaningful lower-rank brain representation. The size of this set varies greatly in the literature, from fewer than ten large networks to hundreds of smaller brain units[5,11,12].

A brain unit may be spatially contiguous (i.e., adjacent voxels or vertices) or non-contiguous (i.e., made up of several spatially distinct regions spread throughout the brain). Contiguous brain units are consistent with the concept of functionally specialized cortical areas (Fig. 1a)[1,13]. Conversely, non-contiguous brain units are able to capture complex network structure within the hierarchically organized and largely hemispherically symmetric brain (Fig. 1b; see Box 1 for a discussion on brain unit nomenclature). In addition, brain units may be binary, such that each voxel or vertex is assigned uniquely and fully to one unit (i.e., a hard parcellation), or weighted, where each voxel or vertex can contribute to multiple units as described by its weights (i.e., a soft parcellation).

A wide variety of approaches can be used to define brain units. Parcellations based on established atlases defined from histology, lesions, gyrification or other features are obvious choices[14–16]. However, these atlases are typically derived from a small number of participants, and the anatomically defined brain unit boundaries do not necessarily match the functional organization. A recent comparison between anatomically and functionally defined brain units in the contexts of rfMRI predictions reported consistent improvements in prediction accuracy when using functionally defined brain units[17]. A wide variety of methods have been developed that use rfMRI data to generate parcellations. These may provide brain units that are more targeted to the specific resolution and contrast of rfMRI, and can provide greater flexibility in the parameters of the parcellation.

Clustering and related approaches have produced a number of high-quality rfMRI-derived binary parcellations (that are often also contiguous)[18–22]. Algorithms that parcellate rfMRI data into

Box 1 | Brain unit nomenclature

In this article we define a 'brain unit' as a basic functional processing entity. Within the rfMRI literature, brain units have been described in a number of different ways. We briefly clarify the semantic distinctions between some of these different conceptualizations of a brain unit:

- **Cortical area:** a cortical area benefits from having the clearest definition, being a region of cortex that is distinct from neighboring cortex in terms of its function, cytoarchitecture, connectivity or topographic organization[93].
- **Parcellation:** the verb 'parcellate' refers to the process of subdividing the brain into a number of brain units. By extension, an individual brain unit may also be referred to as a parcel.
- **Region of interest (ROI):** in some sub-disciplines of rfMRI, a brain unit may also be described as a 'region of interest' or ROI. For example, one possible rfMRI brain representation is to define one brain unit (ROI) as the seed and estimate the whole-brain (voxel-wise) correlation map with the seed ROI[94].
- **Edge:** one way to study interactions between brain units is to use the mathematical structure of a graph, where each individual brain unit is referred to as a 'node' and pairwise connections between brain units are known as 'edges'[28]. In practice for rfMRI, a node often (but not always) describes a contiguous binary brain unit that is part of a high-dimensional (50+) parcellation.
- **Network:** another common naming that fits with our intuitive understanding for a grouping of items is a 'network'. In rfMRI, 'network' may refer to the brain as a whole (i.e., in 'network neuroscience')[95], to a set of multiple brain units that are similar to each other (for example, defined using hierarchical clustering)[28] or to a single non-contiguous weighted brain unit (popularly known as 'resting state networks', or RSNs)[83]. Overall, the term 'network' tends to refer to the highest organizational level of the brain into a relatively small number of macroscale functional patterns (<25).

weighted units (that are often also non-contiguous) include principal component analysis (PCA), independent component analysis (Fig. 1b; spatial or temporal), non-negative matrix factorization, probabilistic functional modes and dictionary learning[23–26]. A number of these methods, in particular independent component analysis, have been extensively employed in rfMRI studies.

Parcellations defined from rfMRI rely on the ability of the spontaneous BOLD fluctuations to resolve brain units, which possibly limits their wider applicability. For example, specific fMRI task activity may reveal further division of brain units not apparent during rest. Multimodal approaches that use combinations of structural, rest and task functional imaging may provide parcellations with improved generalizability across different modalities. However, these multimodal parcellations may generalize less well to new out-of-sample rfMRI data, compared with unimodal (for example, purely rfMRI) parcellations, reflecting the trade-off between optimal modality-specific fit versus accounting for cross-modal disagreement[10].

Defining the summary measures. *Functional connectivity summary measures.* The most common type of information that is of interest in rfMRI is functional connectivity, which is defined as statistical similarities between signals measured from different brain regions and is thought to be indicative of functional integration[27]. For example, graph-based connectomics methods capture functional connectivity information by conceptualizing each individual brain unit as a node within a graph (for reviews see refs. [28,29]). Pairwise functional connectivity between nodes is commonly estimated as the correlation coefficient between summary time series from different brain units, but a variety of other summary measures can be used (Box 2). Connectomics refers to the study of all possible pairwise node-to-node functional connections (edges), which can be summarized in matrix form (network matrix; Fig. 2a)[30,31].

Alternatively, in representations that use weighted brain units, functional connectivity information may be captured in the spatial voxel or vertex weights, which can be used as features that vary freely across participants (for example, using dual regression as the summary measure to estimate participant-specific maps[32]). Depending on the dimensionality and definition of the brain units, functional connectivity can therefore be represented either temporally (i.e., in terms of the temporal correlation between a pair of separate brain units) or spatially (i.e., by combining multiple brain regions into the same non-contiguous brain unit and using the voxel-wise spatial weights to capture the relative strength of connectivity). This ambiguity leads to challenges in the interpretation of brain representations as discussed further in the "Representational ambiguity" section.

While the majority of summary measures consider static functional connectivity (i.e., averaged across the full rfMRI scan), a number of dynamic brain representations explicitly aim to capture time-varying functional connectivity in the summary measures[33,34]. While static functional connectivity estimates are informed by time-varying signal fluctuations, the key difference is that only one estimate of connectivity is obtained across the entire scan, whereas dynamic connectivity methods calculate multiple estimates of connectivity separately for different periods of time over the course of the scan. Many dynamic functional connectivity methods aim to identify distinct dynamic states, for example by performing clustering on sliding window estimates[35] or using a hidden Markov model (Fig. 2b)[36]. These are then used to generate the features for further analysis.

A further variant of functional connectivity summary measures is to aim for causal inference by estimating the directed connection from one brain unit to another (Fig. 2c; also referred to as 'effective connectivity'[27]). Indeed, a recent article by Reid and colleagues proposed causal neural interactions as a unifying conceptualization of functional connectivity[3]. While the conceptual considerations and ambitions of this aspirational framework are of great importance, their suggestions may appear somewhat abstract to investigators grappling with their own rfMRI data, in part because some of the key analytic decisions that researchers need to make are not fully addressed.

Univariate (node-based) summary measures. While the summary measures of most rfMRI brain representations assess functional connectivity (i.e., integration) in some form, several alternative summary measures describe aspects of the data at each brain unit. Examples include the localized signal amplitude (i.e., BOLD signal strength)[37–39], the size of brain units or the spatial overlap of weighted brain units[40,41]. Despite their different nature, these univariate measures are often not independent from functional connectivity. For example, changes in signal amplitude may directly influence functional connectivity (see "Representational ambiguity" below)[42].

Complex spatiotemporal brain representations. While all approaches discussed so far start with the definition of a brain unit, some brain representations avoid this step and estimate complex spatiotemporal patterns from the full data. For example, rfMRI data can be represented as one (or more) connectivity gradient(s) that capture variation in functional connectivity along a continuous axis of spatial location[43]. This approach can be used to identify overlapping modes of organization within a predefined brain unit[44] or to map

Box 2 | Example summary measures

A variety of summary measures can be used to calculate features of interest from the preprocessed rfMRI data that has been parcellated into brain units. Some example summary measures that have been used in the literature are summarized here.

- Edge-based measures describe functional connectivity between pairs of brain units, which can be combined to form network matrices. Edge-based summary measures include full or partial correlation[28] and mutual information.
- Directed (effective) connectivity measures are edge-based measures that aim to estimate the directionality as well as the strength of connectivity between pairs of brain units. Example measures include dynamic causal models[96], Patel's conditional dependence measures[97], pairwise likelihood ratios (LiNGAM)[98], Bayes nets[99] and dynamic graphical models[100].
- Dynamic functional connectivity measures reflect variations in functional connectivity over time (typically without attempting to estimate causality). One option for investigating dynamic functional connectivity is to estimate edge-based measures listed above from short, sliding time windows (instead of from the entire scan) and cluster the resulting network matrices into brain states[35]. To overcome challenges with window-based dynamic analyses[33], statistical approaches such as the hidden Markov model offer alternative summary measures for estimating dynamic features[92].
- Univariate (node-based) measures, such as the amplitude of the BOLD signal, can be estimated at the brain unit level or at the voxel level, based on the time series standard deviation[39] or the relative strength of low-frequency power[37].
- Spatial maps describe the topographical organization of brain units. Inter-individual variability in spatial properties can also be used as features, for example using summary measures that estimate participant-specific weighted resting-state networks[32], or participant-specific brain unit boundaries in binary parcellations[69].

the principal global (whole-brain) pattern of cortical organization from primary sensorimotor to multimodal association cortex (Fig. 2d)[45]. As a further example of complex brain representations, some dynamic methods directly take voxel or vertex rfMRI data as input to estimate a movie-like brain representation of spatially and temporally varying patterns, known as 'propagating waves'[46].

Challenges for brain representations

The current divergence of brain representations (i.e., combined descriptions of brain units and summary measures) of rfMRI data is natural and expected as part of an initial exploration phase that has been echoed in other disciplines[47]. However, now that the field is maturing to include biomarker discovery, efforts toward convergence to validated representations are needed to build a cumulative scientific framework.

Such efforts are complicated by the absence of a gold standard metric for the validation and comparison of brain representations. The underlying neural organization in an individual is not fully accessible by noninvasive technologies (and such detailed circuitry data would provide major representational challenges of its own[48]). Representations derived from rfMRI may therefore correspond to different features of this underlying neural organization. As such, comparisons across different representations necessarily rely on indirect metrics such as behavioral prediction accuracy, genetic heritability, within-parcel homogeneity, variance explained, test–retest reliability, comparisons against other data modalities such as structural imaging and histology, and simulations. In this section we lay out a number of challenges for brain representations that come about as a result of this lack of ground truth knowledge, with the goal of promoting awareness of the issues that are in some cases rarely explicitly considered or taught.

Heterogeneity and dimensionality of brain units. A common assumption inherent to the majority of brain representations is that an individual brain unit is functionally homogeneous, such that its relevant activity can be accurately reflected in a single summary time series. However, functional heterogeneity within a brain unit can be produced by measurement noise, structured artifacts, between-participant variability and 'true' heterogeneity at the level of neural processing. In addition to heterogeneity, functional multiplicity occurs when the same part of cortex encodes different types of information. Examples of this multiplicity can be seen in visual cortex, which simultaneously encodes a retinotopic map and stimulus orientation, or in parietal cortex, where different somatotopic maps converge and overlap[49]. The inevitable presence of some neuronal functional heterogeneity and/or multiplicity within brain units is typically acknowledged or even assumed[1], yet the implications for brain representations are rarely considered or accounted for.

One way to potentially reduce problems with functional heterogeneity and multiplicity is to split parcels up into ever-smaller brain units, to achieve a more fine-grained brain representation. However, an excessively fine parcellation, with multiple brain units representing the same functional entities, may lead to complexities with modelling and interpretation. For example, if a functional region is inappropriately divided into multiple brain units, this can lead to incorrect estimates of functional connectivity when using partial correlation and can detrimentally affect causal connectivity models. Determining the optimal number of brain units in a brain representation that balances the trade-offs between homogeneity and model complexity is challenging.

No consensus has been reached on the question of the optimal dimensionality of brain representations, with recent suggestions varying from six macroscale systems[12] to several hundred parcels[13,17]. This wide range is in part due to the hierarchical organization of the brain, such that it can be meaningfully represented at multiple different levels of granularity, depending on the research question[50]. For example, topological features of functional brain organization can be studied at different dimensionalities[29], and different patterns of within- and between-participant variability may dominate at different scales[51]. However, increased dimensionality estimates can also occur due to inadequate handling of between-participant variability, resulting in a misleadingly detailed granularity (see section "Dealing with variability"). Notably, the effective dimensionality of rfMRI data is subject to biological limits imposed by dependencies of the hemodynamically mediated BOLD signal on the architecture and latency of the brain's microvasculature[52].

These challenges of heterogeneity point to a disconnect that exists between the best model of the human brain and the best model of the rfMRI measures that are acquired. There is ample evidence in humans and other species for the presence of functionally specialized neural populations organized into neuroanatomically distinct cortical areas. Following this evidence, a binary parcellation into contiguous brain units may therefore be the best macroscale model of the brain[1]. However, despite rapid advances in recent years with the help of accelerated acquisition methods[53], the spatial and temporal resolution of fMRI are many orders of magnitude removed from the scale of neural populations and action potentials. Similarly, the physiology of the hemodynamic response implies limits on resolution, independent of advances in MR image

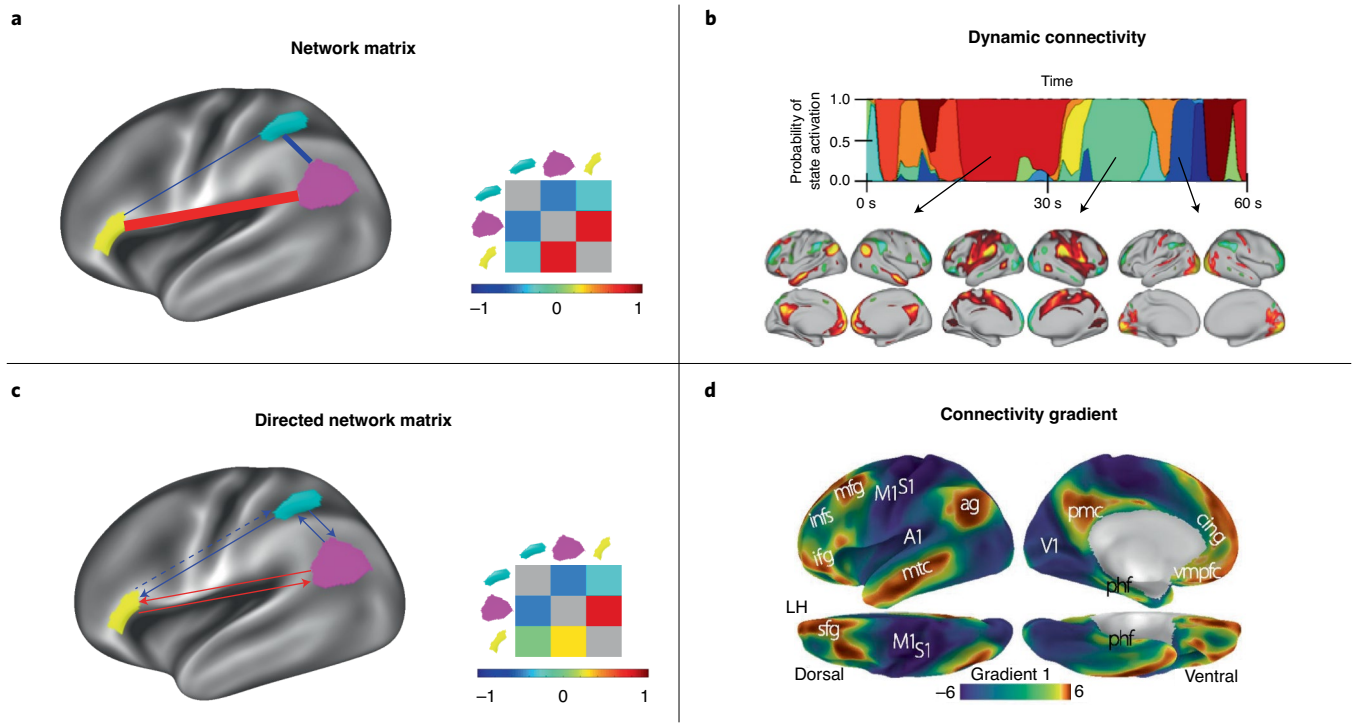

Fig. 2 | Different functional-connectivity-based versions of summary measures in different brain representations. a, Edges between nodes are shown in graphical form on the left (positive and negative correlations in red and blue, respectively; line thickness indicates correlation strength) and in matrix form on the right. **b**, Example of different dynamic brain states and their fluctuations over time derived using a hidden Markov model (figure adapted from ref. [92], Elsevier). **c**, Directed edges may have different reciprocal connections strengths, shown as separate arrows in the graphical form, and resulting in a non-symmetrical network matrix. **d**, Principal global connectivity gradient from primary sensorimotor regions (blue) to multimodal association cortex (red) (figure adapted from ref. [45], US National Academy of Sciences).

acquisition[52]. Therefore, data obtained with rfMRI provide a crude measurement that collapses information both spatially and temporally. As a result, weighted parcellations that allow for overlapping organization and fuzzy boundaries may provide a better model of the data measured from rfMRI. Indeed, previous comparisons have suggested that weighted brain representations may perform better in terms of predicting behavioral traits than binary parcellations[17,54–57]. However, care is required when interpreting such weighted parcellations. For example, spatial overlap between brain units may be an important summary measure to take into consideration (see section "Representational ambiguity")[40].

Dealing with variability. Brain representations are often defined based on a large number of participants to achieve correspondence across individuals for group comparisons and to overcome the limited signal-to-noise ratio in single-participant data. However, variability across individuals in measured functional brain organization may result from spatial misalignment across individuals[58,59] or 'true' individual differences in brain structure[60,61] and/or function[62,63]. Despite advances in surface-based alignment methods designed to address this variability[64,65], recent studies of extensively scanned individuals point to detailed individualized features of organization that are misrepresented in group-derived brain representations[66–68].

A number of recent approaches aim to address these issues of between-participant variability by estimating individualized parcel boundaries[32,41,69,70], integrating group and participant estimates in the same Bayesian framework[24,71], adopting naturalistic movie-watching paradigms to control variability during data acquisition[72] or moving toward connectivity-based hyperalignment across participants[73]. It is yet to be seen how successful these methods will be in generating brain representations that are robust to variability, yet sensitive to valuable inter-individual information.

In addition to between-participant variance, instability of brain representations within an individual over time (for example, between-session variance or even within-session dynamics within an individual) is a further source of variance for brain representations. While some studies have reported stable trait-like characteristics of brain representations[74–76], other work has shown state-dependent changes based on task demands[77], as well as fluctuations associated with arousal states[39,67] and physiology[78]. Within-participant longitudinal changes that occur as a function of development, aging or disease progression have yet to be characterized. Together, these potential sources of within-participant variability point to the importance of disambiguating trait and state effects in brain representations.

Complex interactions between within-participant variability, between-participant variability and dimensionality form a major challenge for the definition and interpretation of brain representations. Applications of brain representations mostly aim to investigate between-participant effects (for example, patient–control comparisons; fingerprinting identification of individuals; prediction or regression with respect to behavior, cognition or diagnosis). Determining which brain representations are most sensitive to between-participant effects is therefore critically important. For example, estimating individualized brain unit boundaries to remove misalignment as a source of between-participant effects will improve interpretability. Furthermore, comparisons to empirically inform the best dimensionality and summary measure for specific between-participant questions are of great importance. A recent example of such a benchmarking study performed by Dadi and

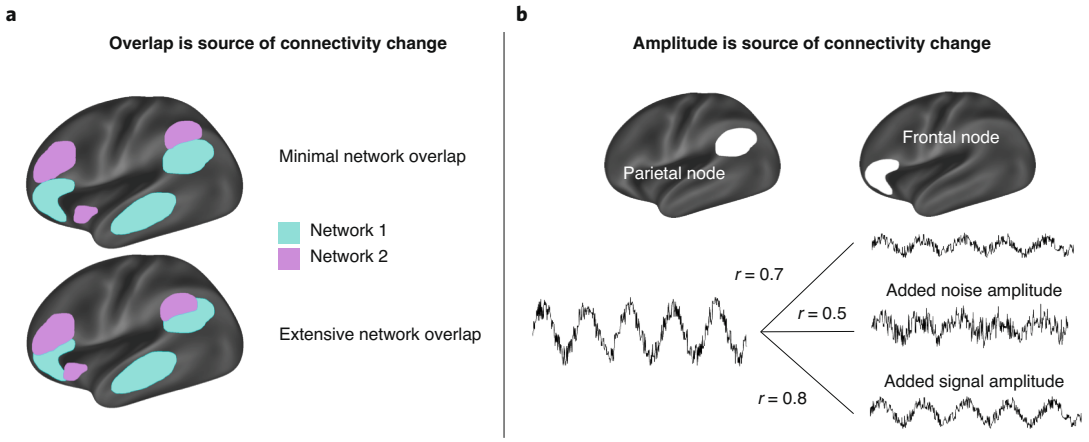

Fig. 3 | Toy examples of representational ambiguity. a, Individual differences in the degree of spatial overlap between two separate networks can be the underlying source of apparent changes in functional connectivity. Depending on the type of representation, network overlap may result in apparent connectivity changes through averaging mixed signals for hard parcellations or subtle biases in dual regression (weighted networks)[40]. **b**, The amplitude of the measured BOLD signal may increase in one region due to either added noise or added signal, which can result in apparent connectivity changes between this node and one or more other nodes[42].

colleagues provided specific recommendations, showing that weighted brain units and tangent space embedding of edges outperformed alternative brain representations[17]. Additional future comparative work will be of great value.

Non-neural confounds in functional MRI (fMRI) data add yet another source of unwanted variance. Structured artifacts can be caused by participant head motion, cardiac and respiratory cycles, and interactions of these participant factors with the magnetic field (in-)homogeneity, excitation pulses and image readout. Current preprocessing methods for removing non-neural confounds are imperfect[79], and their application to rfMRI data can have unwanted side-effects[80–82]. The summary measures of rfMRI brain representations are mostly based on similarities in the measured BOLD signal, and therefore require one or more source(s) of stochastic variation (for example, spontaneous neural activity fluctuations). The uncertain nature of these sources increases the impact of confounds for rfMRI (in general more than with task fMRI). The development and comparison of improved data preprocessing strategies is therefore an active area of research and discussion.

Representational ambiguity. The importance of comparing and consolidating across available brain representations is underscored by their diverging interpretations. To demonstrate this, we take the example of a group of patients and a group of healthy controls, where the patients have relatively weaker functional connectivity between the frontal and parietal cortex. When these data are summarized using different brain representations, all will likely capture the group differences in some way. However, depending on the methodology, the results will potentially be represented very differently, leading to conflicting interpretations.

In a brain representation that defines a low-dimensional set of non-contiguous weighted brain units and allows the spatial voxel-wise weights to change across participants as the features (for example, as estimated from dual regression), the weaker frontoparietal functional connectivity in patients will be captured as lower weights in the frontoparietal brain unit map in patients. The interpretation of this finding would be described as a between-group difference in the spatial topography of the frontoparietal network. Alternatively, in a brain representation that defines a high-dimensional set of contiguous binary brain units and investigates between-unit edge connectivity as the summary measures, the above example will be captured as weaker pairwise functional connectivity estimates (i.e., edges) between frontal and parietal brain units in patients. Here the interpretation would be explained as a between-group difference in the amount of integration (or coupling) between frontal and parietal brain regions. Furthermore, in a connectivity gradient brain representation, the frontal and parietal regions may end up further apart in the embedding space in the patients, as a result of their connectivity patterns with the rest of the brain. This could result in a shift of one region relative to the other along the principal gradient, which would be interpreted as a difference in the organizational hierarchy of processing streams for multimodal integration in patients compared with healthy controls. In a final alternative example of a brain representation in which the signal amplitudes of brain units are considered as the features, the above group difference may be observed as a reduction in the signal strength in the frontal and/or parietal brain unit(s) in patients, leading to an interpretation in terms of between-group differences in levels of activation.

Depending on the research goal of a given study, these representational ambiguities may be more or less critical to the resulting conclusions. More general insights into functional brain organization (such as the similarity between resting state and task organization[83]) should be relatively independent of the specifics of the chosen brain representation. Similarly, the choice of brain representation may not be critical if the goal is to achieve accurate clinical or behavioral predictions (for example, the clinical prediction accuracy is potentially similar for all options described in the example above, because each brain representation captures the patient–control differences). However, the different brain representations have strongly diverging implications with respect to the hypothesized mechanism of psychopathology, with potentially conflicting treatment suggestions. Therefore, the 'best' brain representation should inform theories on the origin (as opposed to downstream effects) of disease mechanisms and generate testable hypotheses for follow-up research. The absence of such biological interpretability may be acceptable if the goal is to adopt rfMRI purely for clinical or behavioral prediction. Unfortunately, however, ill-founded explanations often end up being given for such predictions.

The examples and studies discussed in this section illustrate two possible scenarios of representational ambiguity. First, we observe

Box 3 | Best practice guidelines

1. **Avoid the use of outdated anatomic definitions of brain units**
   There is robust evidence that parcellations derived purely from gyral and sulcal anatomy, such as the automated anatomical labeling (AAL), Harvard–Oxford atlas or FreeSurfer folding-based parcellations, are not suitable brain units for use in functional rfMRI brain representations[17,101], particularly when they are based on only small numbers of participants. Instead, definitions of brain units that take into account functional data (potentially alongside other imaging modalities) should be used.
2. **Include a justification for the chosen brain representation**
   We recommend that all articles report a justification for the chosen brain representation and a clear description of the associated assumptions in relation to challenges raised in the section "Challenges for brain representations" (see checklist in Box 4). These assumptions should be explicitly acknowledged and considered when the results are discussed (see guideline #10). This level of transparency will help with the interpretation of findings within the broader literature and will inform follow-up work.
3. **Adopt established, state-of-the-art data acquisition standards and appropriate confound monitoring**
   Recent findings clearly indicate advantages of eyes-open instructions (ideally with fixation cross), longer data acquisition and/or multiple sessions per participant, faster repetition time (potentially using accelerated imaging methods) and adequate spatial resolution[9,10,102]. In addition, concurrent tracking of physiology and arousal states are beneficial[78]. If these recommendations are not possible to implement (for example in legacy data), then be aware of the limitations this imposes on downstream analyses (in terms of, for example, the resolution of brain units it is possible to utilize).
4. **Implement optimal data preprocessing, alignment and harmonization methods**
   While consensus on optimal preprocessing pipelines is lacking (see guideline #5), it is clear that nuisance regression of head motion, white matter and cerebrospinal fluid regressors, combined with low-pass filtering, is insufficient for denoising[79,103]. Instead, spatial ICA-based cleanup methods should be used where possible (and/or volume censoring where not). In addition, when possible, studies should use multimodal surface-based alignment to reduce between-participant variability[64]. For transparency, the use of validated and published methods and pipelines (for example, refs. [104,105]) is preferable over preprocessing scripts developed in-house. Lastly, when collating data from different studies or centers, data harmonization techniques (beyond a simple site regressor) are required to appropriately account for site effects[106].
5. **If global signal regression changes your findings, also present results without it**
   Aside from the recommendations in #4, some studies remove (i.e., regress out) any variance that can be explained by the whole-brain (globally averaged) BOLD time series in an attempt to remove spatially diffuse confounds[107]. However, this practice (termed 'global signal regression' or GSR) has been shown to affect the definition of brain units by shifting boundaries[10] and to affect summary measures by shifting the edge distribution[108]. Therefore, if authors wish to publish findings with the use of GSR, these should be accompanied by comparable results without GSR (potentially placing one in the supplementary files). This will facilitate comparisons across studies and may inform the discussion. The use of alternative approaches that avoid some of these biases, such as physiological models[109] and/or temporal ICA cleanup[110], should also be considered. It should also be noted that the use of partial correlation instead of full correlation to estimate edges removes the need for GSR[28].
6. **Use well-powered samples and out-of-sample validation**
   A wide range of freely shared large-scale data are now available from healthy participants and various clinical cohorts. To overcome challenges with replicability, studies should make extensive use of these resources to achieve robust sample sizes and to validate results obtained from one dataset by including external replications in other datasets. Funding bodies play a role in recognizing such efforts by supporting ongoing analysis of existing data resources.
7. **Adopt robust statistical approaches**
   Quantitative statistical tests should be favored over qualitative descriptive results where possible. Many common methods such as clustering and canonical correlation analyses will by definition return results, and appropriate significance testing and stringent cross-validation are essential for robust inference[111,112]. Robust statistical methods should include correction for multiple comparisons with the use of appropriate statistical thresholds[113].
8. **Compare different brain representation where possible**
   To aid the generalizability of results, we recommend that studies show how their findings vary as a function of brain representation. While articles are starting to include results across a variety of dimensionalities and/or using a number of different binary parcellations, broader comparisons are rare (although see refs. [17,114]). Providing directly comparable findings will be of great value in building an understanding of how different representations of the same data relate to each other. Additionally, comparative testing of different types of summary measures from the same brain units (such as signal amplitudes, network matrices and individualized brain unit shape and/or size) can reduce ambiguity and inform mechanistic interpretability[39,42,62].
9. **Publicly share all code, data and results**
   It is essential for transparency and replicability to make research outputs publicly available[115]. Many excellent platforms are now available to share code (https://github.com/git), data alongside analysis pipelines (https://openneuro.org/) and results (https://balsa.wustl.edu/). Importantly, sharing data output facilitates meta-analytical approaches and allows the academic community as a whole to build on the research of others, as required for productive and cumulative scientific progress.
10. **Avoid over-interpretation of findings and clearly state caveats**
    When describing findings, we recommend that authors stay close to the empirical results and avoid terminology that is suggestive of biological mechanisms (unless directly warranted by the results). Explicitly discussing potential interpretational caveats (for example, in relation to the brain representation used, representational ambiguity or BOLD limitations) is encouraged (see the checklist in Box 4).

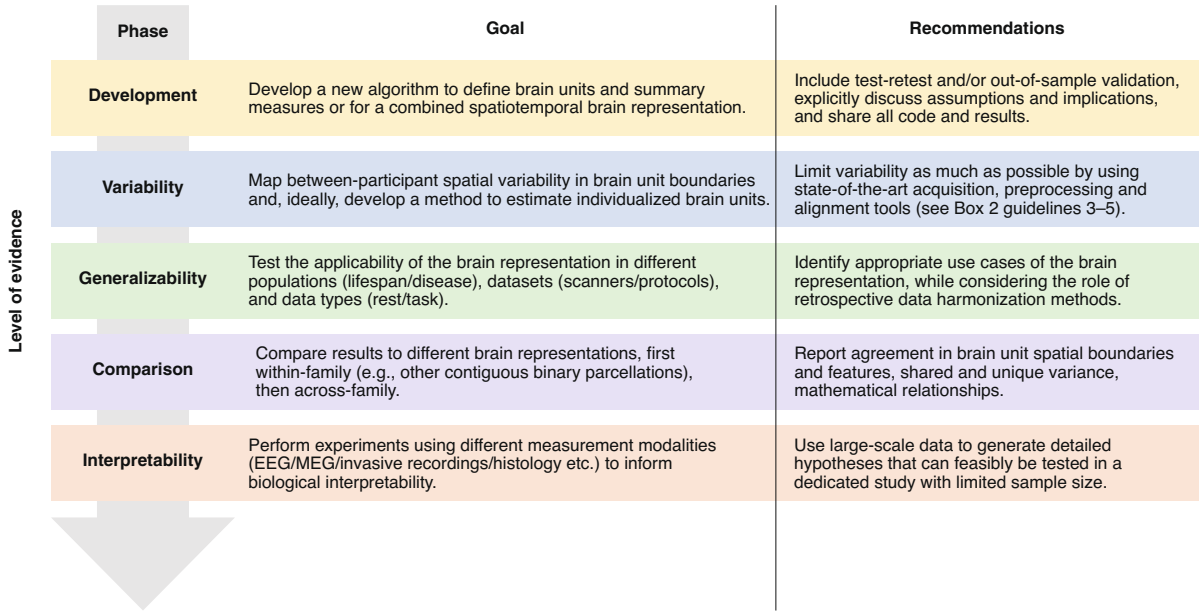

Fig. 4 | Proposal for different phases of validation to provide increasing levels of evidence for brain representations.

representational contamination, such that apparent changes in one summary measure (for example, connectivity) are driven by underlying changes in a different summary measure (for example, signal amplitude). This type of contamination is problematic because if researchers only test the contaminated summary measure (for example, connectivity), this would lead to an incorrect interpretation. Second, we observe the ill-posed nature of brain representations, because the same underlying effect may be represented either spatially or temporally depending on the dimensionality and characteristics of the brain representations. As a result, there is no single unique solution, and different internally consistent and potentially equally valid interpretations may be invoked to explain the same aspect of the underlying data (for example, change in spatial network shape or between-unit edge connectivity). Importantly, in the absence of ground-truth knowledge it can be difficult to distinguish whether a specific example constitutes an incorrect interpretation or one of several possible valid interpretations.

Studies reporting the types of representational ambiguity illustrated above have started to emerge. For example, data-driven simulations revealed that a large proportion of between-participant variance in network matrices was driven by inter-individual differences in spatial network topography, shifting the interpretation from coupling to spatial organization (Fig. 3a)[62]. Furthermore, the additive signal change approach[42] explicitly links different types of summary measures, by determining whether observed changes in functional connectivity network matrices are downstream effects of changes in signal amplitude (Fig. 3b)[39].

Recommendations and future directions for brain representations

The previous sections have highlighted tensions related to the question of brain representation. While a single gold standard representation is attractive in terms of generalizability and replicability, it may lack the flexibility to appropriately capture and account for variability, and it may curtail complementary insights that could be derived from alternative representations. Instead of championing one specific brain representation, we therefore propose a set of best practice guidelines intended to improve the transparency and robustness of research.

Box 3 provides a number of recommendations aimed at applied researchers who wish to adopt existing brain representations in their basic or translational research. The complexity of MRI acquisition and analysis approaches means that outdated research practices can persist, exacerbating the fractionation of the field. Therefore, several items in Box 3 (for example, #1, #3–7) are intended to encourage researchers to review their analysis pipelines and identify state-of-the-art approaches. Important building blocks for unifying the field include ensuring that the challenges of brain representation are widely discussed and increasing validation and replication. Therefore, several of the recommendations in Box 3 (#2, #9, #10) touch on reporting and open science practices (see also the reporting checklist in Box 4), which will inform discussion and enable replication.

In Box 3 we recommend that authors should justify the choice of brain representation based on the research question and hypotheses (#2) and encourage comparisons across multiple brain representations (#8). To differentiate between brain representations and determine the most appropriate ones to test, there are a number of considerations to take into account.

- Brain representations that address some of the challenges laid out in "Challenges for brain representations" should be preferred over and above representations that do not. For example, cortical brain units obtained from surface-based data with known improvements in between-participant alignment[64] are preferable to volumetrically defined cortical brain units. Additionally, brain units that estimate individualized boundaries for each participant (see section "Dealing with variability") are preferable over group-averaged brain units.
- It is important that the scale of the brain units is matched to the hypothesis. For example, if one expects lateralized effects, then an atlas or parcellation that defines left and right regions as separate brain units is appropriate. Similarly, if the hypothesis is specific to the posterior cingulate cortex, then the scale of the brain units should isolate this region instead of incorporating it within the larger default mode network.
- The population used for the definition of brain units should be matched to the population of interest. The majority of available functional atlases are derived from healthy young adult

Box 4 | Reporting checklist

- Have the brain units and summary measures been publicly shared?
- Does the article include a clear motivation for the following elements?
  - Preprocessing choices
  - Dimensionality
  - Choice of brain representation
  - Assumptions associated with the chosen brain representation
  - Implications of the chosen brain representation for interpretation
- Is the algorithm to produce the brain representation from (new) data freely available?
- Have alternate representations been tested?
  - Does using different brain units give concordant results?
  - Do different types of summary measures give concordant or illuminating results?

participants. These brain units may be less appropriate to study developmental, aging or disease cohorts because the brain representation is biased toward a subset of (control) participants. Related to this point, we recommend that existing and new brain representations should explicitly be tested for their generalizability and applicability across different cohorts (Fig. 4).

- Interactions between the brain units and the summary measure need to be considered. For example, if the intention is to use partial correlation as the summary measure to improve sensitivity to direct (rather than indirect) connections, then it is likely advantageous to combine left and right homologs together into a single brain unit to improve numerical stability, even with explicit regularization of the model fitting.

In addition to the best practice guidelines for adopting brain representations (Box 3), we also propose a series of validation phases that should be applied to existing and future brain representations (Fig. 4). Following the development of new brain representations, we propose explicit steps to (i) measure between-participant variance and ideally estimate individualized brain units, (ii) explicitly test the generalizability across different disease and lifespan populations and across scanners, (iii) perform systematic and extensive comparisons against multiple existing types of brain representations and, ultimately, (iv) inform the interpretation of a brain representation based on multimodal experiments.

To aid our ability to synthesize and compare different representations, targeted theoretical and empirical work to establish direct relationships between different brain representations would constitute a major contribution. For example, bottom-up biophysical simulations of neural firing (in which the ground truth is known) may provide a useful test environment to link different representational definitions of brain units and summary measures to each other and to elements of the underlying biophysiology[58,84,85]. In addition, mathematical efforts to establish known mapping functions between different brain representations, and/or between a common atlas and customized brain units of individuals or sub-groups, would greatly facilitate cross-study generalizability and replicability. These suggestions for future work will be aided by continued efforts to characterize and disentangle different sources of variance (confounds as well as between- and/or within-participant variation).

Establishing biological interpretability of fMRI-derived brain representations by studying the underlying neural circuitry and cellular mechanisms is critically important for the next generation of rfMRI research (Fig. 4). Specifically, gaining neurophysiological insights into different summary measures (such as between-region connectivity and signal amplitudes) is expected to reduce representational ambiguity and inform appropriate interpretations. Integrating information across different (potentially invasive) measurement modalities is needed to determine how different brain representations derived from fMRI relate to neural firing and cellular processes[86–88]. Combining fMRI with complementary non-invasive electrophysiology, for example, electro- and/or magnetoencephalography (EEG/MEG), offers the opportunity to noninvasively probe large-scale brain networks with more direct neuronal measures, albeit with coarser spatial detail, but free from vascular confounds and at faster time-scales, thereby aiding in the validation of dynamic representations[36,89]. Additionally, electrode recordings in animals measuring local field potentials, neural spiking and/or oxygen polarography, as well as tracer studies measuring structural connectivity, can offer more direct biological validation[90,91]. Taken together, multimodal research efforts are expected to provide insights into the biological basis of rfMRI-derived brain representations, which will be critical for determining the most appropriate representation and to inform a biological grounding to inferences.

Conclusion

The rfMRI field is fractionated by the divergence of brain representations. Despite the importance of the chosen brain representation for study outcomes and interpretation, articles rarely include a clear justification for adopting a specific representation. Instead, it is common for laboratories to subscribe to a specific approach and apply this across all research projects with relatively little consideration for the implicit assumptions of their chosen brain representation. This tendency can produce research silos of segregated reasoning and assumptions that are at odds with the fundamental principles of cumulative science. Advancing the field beyond these research silos and toward successful collaborative brain mapping and interpretable biomarker discovery critically requires a better understanding of the relationships between different representations of the same data. Once we gain a clearer understanding of the relationships between different brain representations and between representations of rfMRI data and the underlying neurophysiology, some of the key concepts, interpretations, definitions and nomenclature in the field may need to be redefined or updated. This will require a commitment and willingness from members in the field to test, challenge and revise our assumptions and core principles. Improving the interpretability of rfMRI brain representations will increase the replicability of results across different research laboratories and improve the real-life clinical impact of rfMRI to inform diagnosis and treatment. The guidelines and suggestions proposed in this article are intended to bring the broader community together in setting new standards for the field.

Received: 4 May 2020; Accepted: 16 September 2020;
Published online: 26 October 2020

References

1. Van Essen, D. C. & Glasser, M. F. Parcellating cerebral cortex: how invasive animal studies inform noninvasive mapmaking in humans. *Neuron* **99**, 640–663 (2018).
This review of brain mapping (i.e., parcellating the brain into units) summarizes extensive research performed in non-human primates and discusses how these results inform cortical brain parcellation efforts in humans using noninvasive imaging techniques including rfMRI. This article provides both a historical and a future perspective on cortical brain parcellation, which is an important aspect of brain representations (related to our description defining the brain units in section "Defining a brain unit").
2. Bijsterbosch, J., Smith, S.M. & Beckmann, C.F. *Introduction to Resting State fMRI Functional Connectivity* (Oxford Univ. Press, 2017).
3. Reid, A. T. et al. Advancing functional connectivity research from association to causation. *Nat. Neurosci.* **22**, 1751–1760 (2019).

4. Botvinik-Nezer, R. et al. Variability in the analysis of a single neuroimaging dataset by many teams. *Nature* **582**, 84–88 (2020).
5. Eickhoff, S. B., Yeo, B. T. T. & Genon, S. Imaging-based parcellations of the human brain. *Nat. Rev. Neurosci.* **19**, 672–686 (2018).
6. Ferrante, M. et al. Computational psychiatry: a report from the 2017 NIMH workshop on opportunities and challenges. *Mol. Psychiatry* **24**, 479–483 (2019).
7. Poldrack, R. A., Huckins, G. & Varoquaux, G. Establishment of best practices for evidence for prediction: a review. *JAMA Psychiatry* **77**, 534–540 (2020).
8. Lopes-dos-Santos, V., Ribeiro, S. & Tort, A. B. L. Detecting cell assemblies in large neuronal populations. *J. Neurosci. Methods* **220**, 149–166 (2013).
9. Smith, S. M. et al. Resting-state fMRI in the Human Connectome Project. *Neuroimage* **80**, 144–168 (2013).
10. Glasser, M. F. et al. A multi-modal parcellation of human cerebral cortex. *Nature* **536**, 171–178 (2016).
11. Eickhoff, S. B., Thirion, B., Varoquaux, G. & Bzdok, D. Connectivity-based parcellation: critique and implications. *Hum. Brain Mapp.* **36**, 4771–4792 (2015).
12. Uddin, L. Q., Yeo, B. T. T. & Spreng, R. N. Towards a universal taxonomy of macro-scale functional human brain networks. *Brain Topogr.* **32**, 926–942 (2019).
    This article discusses the challenges of scale and dimensionality of brain representations (related to our discussion of dimensionality in the section "Heterogeneity and dimensionality of brain units"). Focusing on the level of large-scale brain networks, the authors propose a consistent nomenclature for the naming of networks grounded in anatomy to address widespread problems with inconsistent terminology in the literature.
13. Eickhoff, S. B., Constable, R. T. & Yeo, B. T. T. Topographic organization of the cerebral cortex and brain cartography. *Neuroimage* **170**, 332–347 (2018).
14. Tzourio-Mazoyer, N. et al. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. *Neuroimage* **15**, 273–289 (2002).
15. Desikan, R. S. et al. An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage* **31**, 968–980 (2006).
16. Eickhoff, S. B. et al. A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* **25**, 1325–1335 (2005).
17. Dadi, K. et al. Benchmarking functional connectome-based predictive models for resting-state fMRI. *Neuroimage* **192**, 115–134 (2019).
    This paper is a great example of the benefit of comparing different brain representations to determine their relative value in addressing a specific research question. The results offer clear-cut insights that are important for brain representations. For example, functionally defined brain units clearly outperform anatomically defined brain units, which has informed our guideline #1 in Box 3.
18. Schaefer, A. et al. Local-global parcellation of the human cerebral cortex from intrinsic functional connectivity MRI. *Cereb. Cortex* **28**, 3095–3114 (2018).
19. Shen, X., Tokoglu, F., Papademetris, X. & Constable, R. T. Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. *Neuroimage* **82**, 403–415 (2013).
20. Craddock, R. C., James, G. A., Holtzheimer, P. E. III, Hu, X. P. & Mayberg, H. S. A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Hum. Brain Mapp.* **33**, 1914–1928 (2012).
21. Gordon, E. M. et al. Generation and evaluation of a cortical area parcellation from resting-state correlations. *Cereb. Cortex* **26**, 288–303 (2016).
22. Yeo, B. T. T. et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* **106**, 1125–1165 (2011).
23. Beckmann, C. F., DeLuca, M., Devlin, J. T. & Smith, S. M. Investigations into resting-state connectivity using independent component analysis. *Phil. Trans. R. Soc. Lond. B* **360**, 1001–1013 (2005).
24. Harrison, S. J. et al. Large-scale probabilistic functional modes from resting state fMRI. *Neuroimage* **109**, 217–231 (2015).
25. Varoquaux, G., Gramfort, A., Pedregosa, F., Michel, V. & Thirion, B. Multi-subject dictionary learning to segment an atlas of brain spontaneous activity. *Inf. Process. Med. Imaging* **22**, 562–573 (2011).
26. Eavani, H. et al. Identifying sparse connectivity patterns in the brain using resting-state fMRI. *Neuroimage* **105**, 286–299 (2015).
27. Friston, K. J. Functional and effective connectivity in neuroimaging: a synthesis. *Hum. Brain Mapp.* **2**, 56–78 (1994).
28. Smith, S. M. et al. Functional connectomics from resting-state fMRI. *Trends Cogn. Sci.* **17**, 666–682 (2013).
29. Betzel, R. F. & Bassett, D. S. Multi-scale brain networks. *Neuroimage* **160**, 73–83 (2017).
30. Finn, E. S. et al. Functional connectome fingerprinting: identifying individuals using patterns of brain connectivity. *Nat. Neurosci.* **18**, 1664–1671 (2015).
31. Smith, S. M. et al. A positive-negative mode of population covariation links brain connectivity, demographics and behavior. *Nat. Neurosci.* **18**, 1565–1567 (2015).
32. Nickerson, L. D., Smith, S. M., Öngür, D. & Beckmann, C. F. Using dual regression to investigate network shape and amplitude in functional connectivity analyses. *Front. Neurosci.* **11**, 115 (2017).
33. Hutchison, R. M. et al. Dynamic functional connectivity: promise, issues, and interpretations. *Neuroimage* **80**, 360–378 (2013).
34. Preti, M. G., Bolton, T. A. & Van De Ville, D. The dynamic functional connectome: state-of-the-art and perspectives. *Neuroimage* **160**, 41–54 (2017).
35. Allen, E. A. et al. Tracking whole-brain connectivity dynamics in the resting state. *Cereb. Cortex* **24**, 663–676 (2014).
36. Vidaurre, D. et al. Spontaneous cortical activity transiently organises into frequency specific phase-coupling networks. *Nat. Commun.* **9**, 2987 (2018).
37. Zou, Q.-H. et al. An improved approach to detection of amplitude of low-frequency fluctuation (ALFF) for resting-state fMRI: fractional ALFF. *J. Neurosci. Methods* **172**, 137–141 (2008).
38. Miller, K. L. et al. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. *Nat. Neurosci.* **19**, 1523–1536 (2016).
39. Bijsterbosch, J. et al. Investigations into within- and between-subject resting-state amplitude variations. *Neuroimage* **159**, 57–69 (2017).
40. Bijsterbosch, J. D., Beckmann, C. F., Woolrich, M. W., Smith, S. M. & Harrison, S. J. The relationship between spatial configuration and functional connectivity of brain regions revisited. *eLife* **8**, e44890 (2019).
41. Kong, R. et al. Spatial topography of individual-specific cortical networks predicts human cognition, personality, and emotion. *Cereb. Cortex* **29**, 2533–2551 (2019).
42. Duff, E. P., Makin, T., Cottaar, M., Smith, S. M. & Woolrich, M. W. Disambiguating brain functional connectivity. *Neuroimage* **173**, 540–550 (2018).
43. Huntenburg, J. M., Bazin, P.-L. & Margulies, D. S. Large-scale gradients in human cortical organization. *Trends Cogn. Sci.* **22**, 21–31 (2018).
44. Marquand, A. F., Haak, K. V. & Beckmann, C. F. Functional corticostriatal connection topographies predict goal directed behaviour in humans. *Nat. Hum. Behav.* **1**, 0146 (2017).
45. Margulies, D. S. et al. Situating the default-mode network along a principal gradient of macroscale cortical organization. *Proc. Natl. Acad. Sci. USA* **113**, 12574–12579 (2016).
46. Majeed, W. et al. Spatiotemporal dynamics of low frequency BOLD fluctuations in rats and humans. *Neuroimage* **54**, 1140–1150 (2011).
47. Woo, C.-W., Chang, L. J., Lindquist, M. A. & Wager, T. D. Building better biomarkers: brain models in translational neuroimaging. *Nat. Neurosci.* **20**, 365–377 (2017).
48. Bargmann, C. I. & Marder, E. From the connectome to brain function. *Nat. Methods* **10**, 483–490 (2013).
49. Jbabdi, S., Sotiropoulos, S. N. & Behrens, T. E. The topographic connectome. *Curr. Opin. Neurobiol.* **23**, 207–215 (2013).
50. Paquola, C. et al. The cortical wiring scheme of hierarchical information processing. Preprint at *bioRxiv* https://doi.org/10.1101/2020.01.08.899583 (2020).
51. Betzel, R. F. et al. The community structure of functional brain networks exhibits scale-specific patterns of inter- and intra-subject variability. *Neuroimage* **202**, 115990 (2019).
52. Logothetis, N. K. What we can do and what we cannot do with fMRI. *Nature* **453**, 869–878 (2008).
53. Moeller, S. et al. Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magn. Reson. Med.* **63**, 1144–1153 (2010).
54. Yu, Q. et al. Comparing brain graphs in which nodes are regions of interest or independent components: A simulation study. *J. Neurosci. Methods* **291**, 61–68 (2017).
55. Abraham, A. et al. Deriving reproducible biomarkers from multi-site resting-state data: An Autism-based example. *Neuroimage* **147**, 736–745 (2017).
56. Duff, E. P. et al. Task-driven ICA feature generation for accurate and interpretable prediction using fMRI. *Neuroimage* **60**, 189–203 (2012).
57. Pervaiz, U., Vidaurre, D., Woolrich, M. W. & Smith, S. M. Optimising network modelling methods for fMRI. *NeuroImage* **211**, 116604 (2020).
58. Smith, S. M. et al. Network modelling methods for FMRI. *Neuroimage* **54**, 875–891 (2011).
59. Allen, E. A., Erhardt, E. B., Wei, Y., Eichele, T. & Calhoun, V. D. Capturing inter-subject variability with group independent component analysis of fMRI data: a simulation study. *Neuroimage* **59**, 4141–4159 (2012).

60. Llera, A., Wolfers, T., Mulders, P. & Beckmann, C. F. Inter-individual differences in human brain structure and morphology link to variation in demographics and behavior. *eLife* **8**, e44443 (2019).
61. Smith, S. et al. Structural variability in the human brain reflects fine-grained functional architecture at the population level. *J. Neurosci.* **39**, 6136–6149 (2019).
62. Bijsterbosch, J. D. et al. The relationship between spatial configuration and functional connectivity of brain regions. *eLife* **7**, e32992 (2018).
This article uncovers an important source of representational ambiguity by showing that between-participant variance in node-to-node correlations is partly driven by variability in spatial organization. Data-driven simulations are used to interrogate interactions between different elements of brain representations, and shared versus unique variance is used to compare different brain representations.
63. Tavor, I. et al. Task-free MRI predicts individual differences in brain activity during task performance. *Science* **352**, 216–220 (2016).
64. Coalson, T. S., Van Essen, D. C. & Glasser, M. F. The impact of traditional neuroimaging methods on the spatial localization of cortical areas. *Proc. Natl. Acad. Sci. USA* **115**, E6356–E6365 (2018).
65. Robinson, E. C. et al. MSM: a new flexible framework for multimodal surface matching. *Neuroimage* **100**, 414–426 (2014).
66. Braga, R. M. & Buckner, R. L. Parallel interdigitated distributed networks within the individual estimated by intrinsic functional connectivity. *Neuron* **95**, 457–471.e5 (2017).
67. Poldrack, R. A. et al. Long-term neural and physiological phenotyping of a single human. *Nat. Commun.* **6**, 8885 (2015).
68. Seitzman, B. A. et al. Trait-like variants in human functional brain networks. *Proc. Natl. Acad. Sci. USA* **116**, 22851–22861 (2019).
69. Hacker, C. D. et al. Resting state network estimation in individual subjects. *Neuroimage* **82**, 616–633 (2013).
70. Gordon, E. M. et al. Precision functional mapping of individual human brains. *Neuron* **95**, 791–807.e7 (2017).
71. Harrison, S. J. et al. Modelling subject variability in the spatial and temporal characteristics of functional modes. *NeuroImage* **222**, 117226 (2020).
72. Vanderwal, T. et al. Individual differences in functional connectivity during naturalistic viewing conditions. *Neuroimage* **157**, 521–530 (2017).
73. Guntupalli, J. S., Feilong, M. & Haxby, J. V. A computational model of shared fine-scale structure in the human connectome. *PLOS Comput. Biol.* **14**, e1006120 (2018).
74. Kieliba, P., Madugula, S., Filippini, N., Duff, E. P. & Makin, T. R. Large-scale intrinsic connectivity is consistent across varying task demands. *PLoS One* **14**, e0213861 (2019).
75. Horien, C., Shen, X., Scheinost, D. & Constable, R. T. The individual functional connectome is unique and stable over months to years. *Neuroimage* **189**, 676–687 (2019).
76. Gratton, C. et al. Functional brain networks are dominated by stable group and individual factors, not cognitive or daily variation. *Neuron* **98**, 439–452.e5 (2018).
77. Salehi, M. et al. There is no single functional atlas even for a single individual: Functional parcel definitions change with task. *Neuroimage* **208**, 116366 (2020).
78. Chen, J. E. et al. Resting-state "physiological networks". *NeuroImage* **213**, 116707 (2020).
79. Ciric, R. et al. Benchmarking of participant-level confound regression strategies for the control of motion artifact in studies of functional connectivity. *Neuroimage* **154**, 174–187 (2017).
80. Murphy, K. & Fox, M. D. Towards a consensus regarding global signal regression for resting state functional connectivity MRI. *Neuroimage* **154**, 169–173 (2017).
81. Power, J. D., Laumann, T. O., Plitt, M., Martin, A. & Petersen, S. E. On global fMRI signals and simulations. *Trends Cogn. Sci.* **21**, 911–913 (2017).
82. Hallquist, M. N., Hwang, K. & Luna, B. The nuisance of nuisance regression: spectral misspecification in a common approach to resting-state fMRI preprocessing reintroduces noise and obscures functional connectivity. *Neuroimage* **82**, 208–225 (2013).
83. Smith, S. M. et al. Correspondence of the brain's functional architecture during activation and rest. *Proc. Natl. Acad. Sci. USA* **106**, 13040–13045 (2009).
84. Deco, G., Jirsa, V. K., Robinson, P. A., Breakspear, M. & Friston, K. The dynamic brain: from spiking neurons to neural masses and cortical fields. *PLOS Comput. Biol.* **4**, e1000092 (2008).
85. Buibas, M. & Silva, G. A. A framework for simulating and estimating the state and functional topology of complex dynamic geometric networks. *Neural Comput.* **23**, 183–214 (2011).
86. Logothetis, N. K., Pauls, J., Augath, M., Trinath, T. & Oeltermann, A. Neurophysiological investigation of the basis of the fMRI signal. *Nature* **412**, 150–157 (2001).
87. Shmuel, A. & Leopold, D. A. Neuronal correlates of spontaneous fluctuations in fMRI signals in monkey visual cortex: Implications for functional connectivity at rest. *Hum. Brain Mapp.* **29**, 751–761 (2008).
This paper used simultaneous rfMRI and intracortical recordings to show that between-region correlations measured from resting state fMRI (i.e., functional connectivity) are linked to synchronization of neuronal signals. This study is one of the few examples that link rfMRI brain representations to the underlying neurophysiology. Similar research efforts are needed to explicitly differentiate between different brain representations, as we have proposed in section "Recommendations and future directions for brain representations" and Fig. 4.
88. Kucyi, A. et al. Intracranial electrophysiology reveals reproducible intrinsic functional connectivity within human brain networks. *J. Neurosci.* **38**, 4230–4242 (2018).
89. Baillet, S. Magnetoencephalography for brain electrophysiology and imaging. *Nat. Neurosci.* **20**, 327–339 (2017).
90. Bentley, W. J., Li, J. M., Snyder, A. Z., Raichle, M. E. & Snyder, L. H. Oxygen level and LFP in task-positive and task-negative areas: bridging BOLD fMRI and electrophysiology. *Cereb. Cortex* **26**, 346–357 (2016).
91. Schölvinck, M. L., Maier, A., Ye, F. Q., Duyn, J. H. & Leopold, D. A. Neural basis of global resting-state fMRI activity. *Proc. Natl. Acad. Sci. USA* **107**, 10238–10243 (2010).
92. Vidaurre, D. et al. Discovering dynamic brain networks from big data in rest and task. *Neuroimage* **180**, 646–656 (2018). Pt B.
93. Van Essen, D. C. & Maunsell, J. H. R. Hierarchical organization and functional streams in the visual cortex. *Trends Neurosci.* **6**, 370–375 (1983).
94. Biswal, B., Yetkin, F. Z., Haughton, V. M. & Hyde, J. S. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. *Magn. Reson. Med.* **34**, 537–541 (1995).
95. Medaglia, J. D., Lynall, M.-E. & Bassett, D. S. Cognitive network neuroscience. *J. Cogn. Neurosci.* **27**, 1471–1491 (2015).
96. Park, H.-J., Friston, K. J., Pae, C., Park, B. & Razi, A. Dynamic effective connectivity in resting state fMRI. *Neuroimage* **180**, 594–608 (2018). Pt B.
97. Patel, R. S., Bowman, F. D. & Rilling, J. K. A Bayesian approach to determining connectivity of the human brain. *Hum. Brain Mapp.* **27**, 267–276 (2006).
98. Hyvärinen, A. & Smith, S. M. Pairwise likelihood ratios for estimation of non-Gaussian structural equation models. *J. Mach. Learn. Res.* **14**, 111–152 (2013).
99. Mumford, J. A. & Ramsey, J. D. Bayesian networks for fMRI: a primer. *Neuroimage* **86**, 573–582 (2014).
100. Schwab, S. et al. Directed functional connectivity using dynamic graphical models. *Neuroimage* **175**, 340–353 (2018).
101. Sala-Llonch, R., Smith, S. M., Woolrich, M. & Duff, E. P. Spatial parcellations, spectral filtering, and connectivity measures in fMRI: Optimizing for discrimination. *Hum. Brain Mapp.* **40**, 407–419 (2019).
102. Noble, S., Scheinost, D. & Constable, R. T. A decade of test-retest reliability of functional connectivity: a systematic review and meta-analysis. *Neuroimage* **203**, 116157 (2019).
103. Burgess, G. C. et al. Evaluation of denoising strategies to address motion-correlated artifacts in resting-state functional magnetic resonance imaging data from the Human Connectome Project. *Brain Connect.* **6**, 669–680 (2016).
104. Glasser, M. F. et al. The minimal preprocessing pipelines for the Human Connectome Project. *Neuroimage* **80**, 105–124 (2013).
105. Esteban, O. et al. fMRIPrep: a robust preprocessing pipeline for functional MRI. *Nat. Methods* **16**, 111–116 (2019).
106. Yu, M. et al. Statistical harmonization corrects site effects in functional connectivity measurements from multi-site fMRI data. *Hum. Brain Mapp.* **39**, 4213–4227 (2018).
107. Power, J. D., Plitt, M., Laumann, T. O. & Martin, A. Sources and implications of whole-brain fMRI signals in humans. *Neuroimage* **146**, 609–625 (2017).
108. Saad, Z. S. et al. Trouble at rest: how correlation patterns and group differences become distorted after global signal regression. *Brain Connect.* **2**, 25–32 (2012).
109. Glover, G. H., Li, T. Q. & Ress, D. Image-based method for retrospective correction of physiological motion effects in fMRI: RETROICOR. *Magn. Reson. Med.* **44**, 162–167 (2000).
110. Glasser, M. F. et al. Using temporal ICA to selectively remove global noise while preserving global signal in functional MRI data. *Neuroimage* **181**, 692–717 (2018).
111. Dinga, R. et al. Evaluating the evidence for biotypes of depression: methodological replication and extension of Drysdale et al. (2017). *Neuroimage Clin.* **22**, 101796 (2019).
112. Varoquaux, G. et al. Assessing and tuning brain decoders: cross-validation, caveats, and guidelines. *Neuroimage* **145**, 166–179 (2017). Pt B.

113. Eklund, A., Nichols, T. E. & Knutsson, H. Cluster failure: why fMRI inferences for spatial extent have inflated false-positive rates. *Proc. Natl. Acad. Sci. USA* **113**, 7900–7905 (2016).
114. Arslan, S. et al. Human brain mapping: a systematic comparison of parcellation methods for the human cerebral cortex. *Neuroimage* **170**, 5–30 (2018).
115. Poldrack, R. A. et al. Scanning the horizon: towards transparent and reproducible neuroimaging research. *Nat. Rev. Neurosci.* **18**, 115–126 (2017).

Acknowledgements

E.P.D. was supported by the SSNAP "Support for Sick and Newborn Infants and their Parents" Medical Research Fund (University of Oxford Excellence Fellowship). S.J.H. was supported by grant #2017-403 of the Strategic Focal Area "Personalized Health and Related Technologies (PHRT)" of the ETH Domain. S.S. is supported by a Wellcome Trust Strategic Award 098369/Z/12/Z and a Wellcome Trust Collaborative Award 215573/Z/19/Z. M.W. is supported by the NIHR Oxford Health Biomedical Research Centre and by the Wellcome Trust (106183/Z/14/Z and 215573/Z/19/Z). We thank M. Glasser for his helpful comments and discussions in relation to this article.

Author contributions

J.B. and E.P.D. conceived of the topic and structure for this article. J.B. wrote the manuscript with input from E.P.D. All authors took part in extensive discussions to refine the arguments presented in this manuscript, and all authors commented on the final manuscript.

Competing interests

The authors declare that they have no conflicts of interest.

Additional information

Correspondence should be addressed to J.B. or E.P.D.

Peer review information *Nature Neuroscience* thanks Finnegan Calabro, Thomas Yeo and the other, anonymous, reviewer for their contribution to the peer review of this work.