Understanding brain organisation in the face of functional heterogeneity and functional multiplicity

Koen V. Haak a,*, Christian F. Beckmann a,b

a Donders Institute for Brain Cognition and Behaviour, Radboud University Medical Centre, Nijmegen, the Netherlands
b Wellcome Centre for Integrative Neuroimaging (WIN FMRIB), University of Oxford, Oxford, United Kingdom

ABSTRACT

Understanding the fundamental organisation of the brain in terms of functional specialisation and integration is one of the principal aims of imaging neuroscience. Many investigations into the functional organisation of the brain are predicated on parcellating the brain into patches of assumed piece-wise constant connectivity. There are, however, many brain areas where the assumption of piece-wise constant organisation is violated. Connectivity, and by extension function, often varies continuously across the grey matter according to multiple overlapping modes of change. The organisation is governed by functional heterogeneity (continuous change) as well as functional multiplicity (overlapping modes). Functional heterogeneity and multiplicity have important implications for how we can and should analyse our data and how we ought to interpret the results, both in the classical context of parcellated modes and under models that allow for overlapping modes of continuous change. The goal of this opinion paper is to raise awareness of these issues and highlight recent methodological developments toward accounting for these important fundamental features of brain organisation.

It is human nature to think in terms of categories and oppositions. Systems neuroscience, being a human endeavour, has been similarly partial to characterising the brain in terms of distinct areas and pathways of assumed homogeneous structure and function. The approach greatly simplifies inference because it enables simple A-versus-B comparisons and permits outer-product models of brain connectivity. However, whenever the assumption of piece-wise homogeneity fails, these approaches inadvertently yield inaccurate models of brain organisation and conceal locally distributed structure in function that may be crucial to performing certain neural computations. As we set out to use neuroimaging to understand how the brain underpins behaviour, we ought to employ analytical approaches that account for these fundamental features of brain organisation.

1. Functional heterogeneity and functional multiplicity

The idea that the brain is comprised of a mosaic of patches with piece-wise constant function has existed throughout the history of neuroanatomy and originates from the seminal work of Korbinian Brodmann (1909) and Corille and Oskar Vogt (Vogt and Vogt, 1926, 1919). These pioneers were the first to divide the brain into areas based on common cytoarchitectonic features to understand cortical organisation and obtain a reference for functional localisation. Ever since, the concept of a brain area has been central to brain mapping, though its definition has broadened to additionally include parcellations based on various other anatomical features, function, connectivity, or a combination of these. In this essay, we adopt this broad definition of brain area, because the parcellation method (i.e. how we identify the borders of a brain area) does not matter in the present discussion. What matters is that—contrary to implicit assumption—brain areas have functionally important internal structure that we contend should be studied and accounted for.

Indeed, there is no doubt that the concept of a brain area is fundamental to understanding brain structure and function. However, it is also important to keep in mind that its definition is relative: we consider a patch of cortex an area if the structural changes within that area are more homogeneous than the structural changes across its borders. Thus, although the internal structure within an area can be considered homogeneous when compared against its neighbours, it may still be meaningfully heterogeneous when comparisons are made, e.g. within a particular patch of cortex. In addition, while the internal structure of an anatomical area may be (relatively) homogeneous in terms of cytoarchitecture, it can be heterogeneous in terms of function and neuronal connectivity. Further, advances across different scientific disciplines can now give rise to several quantitative biological measures, each giving rise to distinct definitions of brain areas, e.g. when considering gene expression profiles or fibre orientations. Either way, one should ask whether it makes sense to average signals or features from across the entire area of interest. For instance, while primary motor cortex...
(Brodman area 4) may be comprised of homogeneous cytoarchitecture, it exhibits a rather heterogeneous composition in terms of axonal projections down the spinal cord (Purves et al., 2008). Thus, averaging signals across the entire primary motor cortex would conceal the highly relevant functional heterogeneity of that area, and one may reasonably ask whether it is meaningful to average signals related to foot, hand and tongue movements.

The problem of functional heterogeneity within one measurement modality—but also the problem of multiple, competing models of piece-wise constancy—can in principle be overcome by performing parcellations at finer scales of functional differentiation, though such super-positioning of parcellations will quickly annul the conceptual benefits of area definition. Further, function within areas often does not exhibit sharp boundaries, and the functional heterogeneity within anatomical areas is often not limited to just one mode of variation. As such, besides functional heterogeneity, brain areas can also exhibit functional multiplicity. Areas in visual cortex, for instance, exhibit at a minimum two gradual modes (i.e. trajectories across the cortical surface) of functional change that simultaneously coexist to keep track of spatial locations within the retinal image (Wandell et al., 2007). In primary visual cortex, distance from the centre of the image (eccentricity) is represented along the calcarine sulcus, whereas the angle about the centre of the image is represented perpendicular to the calcarine sulcus. These representations arise from two gradual modes of connectivity change that are preserved throughout the visual system, through topographically organised retino-fugal, thalamo-cortical and cortico-cortical connections (Udin, 1998). Importantly, the existence of these overlapping modes of organisation implies that we need to think hard about how to analyse the signals within these areas and whether it makes sense to apply parcellation at these scales of brain organisation at all.

Functional heterogeneity and functional multiplicity are features of anatomical and functional connections that likely apply across the brain. Indeed, beyond the visual and motor cortices, functional heterogeneity and multiplicity have been observed in the connectivity patterns of the somatosensory, auditory and olfactory systems, hippocampal, perihinal and entorhinal cortex, striatum, anterior temporal cortex, and frontal cortex (Jbabdi et al., 2013). That is, all of these brain regions exhibit orderly connection topographies (i.e. pointwise mappings such that nearby locations in one region connect to nearby locations in another region) that imply both functional heterogeneity and functional multiplicity. Why these topographic mappings are so ubiquitous is still unclear, but there are several reasons to believe that they are of fundamental importance to brain function, cognition and behaviour. For instance, they permit highly efficient wiring and enable context dependent comparisons, while overlapping connection topographies can efficiently implement transformations between different spatial representations of both sensory as well as abstract non-sensory information (Jbabdi et al., 2013; Kaas, 1997; Thivierge and Marcus, 2007; Tinsley, 2009). Regardless of their exact function, it is important to take note of the fact that the considerations put forward here are not limited to sensory-motor cortex. They appear to be a general feature of brain organisation that has—up to recently—been largely glossed over by the field (Cavada and Reinoso-Suárez, 1985; Draganski et al., 2008; Goldman-Rakic, 1988; Haber et al., 2000; Harvey et al., 2020, 2015; Harvey and Dumoulin, 2017; Hirose et al., 1992; Mackey et al., 2017; Olson and Musil, 1992; Silver and Kastner, 2009; Szinte and Knappen, 2019; Tamamaki and Nojyo, 1995; van Es et al., 2019).

### 2. Modes or models: implications for signal estimation

A large fraction of present-day neuroimaging studies involve mapping a set of brain networks from resting-state functional MRI data in order to characterise the functional connectome. The procedure typically starts with the identification of a set of brain regions (nodes) and the aim is to estimate the connections (edges) between them. Previous work has shown that the exact placement of the network nodes can have strong influence on the connection estimates independent of the method that is used for connectivity estimation (Smith et al., 2011). Moreover, bottom-up simulations demonstrate that even small inaccuracies in region definition are highly detrimental to the network estimation (Smith et al., 2011), and as a result, inter-individual anatomical variations interact strongly with the network estimates (Bijsterbosch et al., 2019, 2018; Lera et al., 2019). These issues already demand extreme caution in the interpretation of network modelling results and are compounded by the possible presence of regional variation within the network nodes. Indeed, this is amongst the reasons that most graph-theoretical summary measures of brain network topology, such as rich clubs and eigenvector centrality (Bullmore and Sporns, 2009; Rubinov and Sporns, 2010; Sporns et al., 2005) are rather distant from biologically meaningful characterisations of the brain.

The problem is one of signal mixing, in that inaccurate node definitions lead to mixed signals from two or more brain areas. Current network estimation approaches summarise a node’s component signals into a single estimate (e.g. the average time-series) and regional variation already presents us with the problem of mixed signals, detrimental to any further downstream inference. This issue is significantly amplified in the case of functional heterogeneity and multiplicity. The former already implies that there is no single time-series characteristic that suitably summarises the signal dynamics within the area, even if the cross-subject or subject-to-area (e.g. via an atlas) regional mapping is refined, for instance through subject-specific parcellation or subject-refined warping procedures. In principle, this problem—and also the problem of inaccurate node definition—may be addressed by characterising the connectivity patterns at finer spatial scales in order to increase within-parcel functional homogeneity. In the presence of functional multiplicity, however, not even finer-grained parcel definition and refined warping procedures can ensure suitable or even sensible characterisation of the functional dynamics, because even at the finest scales (e.g. a single voxel) the signal will be composed of multiple components.

The presence of functional multiplicity also has profound implications for the broader set of inferential techniques applied to fMRI data beyond the simplest forms of connectivity modelling. For instance, because underlying model assumptions are violated, estimating effective connectivity through the use of e.g. Dynamic Causal Models (Friston et al., 2003) will similarly lead to inaccurate characterisations of region-specific functional signal dynamics and henceforth result in erroneous inferences on the causal structure. That is, estimates of effective connectivity rely on generative models that convert neuronal dynamics and relatively simplistic network architectures into region-specific hemodynamic signal predictions that are compared against the data, and these models currently do not account for functional heterogeneity or multiplicity. Similarly, brain reading approaches that rely on multivariate patterns or distributed representations (e.g. as in Representational Similarity Analysis (Kriegeskorte, 2008) – RSA) will attempt to characterise a single spatially distributed signal source and ignore the possible presence of overlapping modes that may confound the ability to predict stimulus status. At the most basic level, even classical General Linear Model-based approaches cannot cope gracefully with the presence of overlapping modes. In all cases, the associated signal dynamics will enter as additional structured noise in the data, harming all down-stream inferential processes. In the presence of multiple functional modes across the same patch, the task of defining a single representative temporal signal profile is entirely futile. Instead, it becomes necessary to model the complex hierarchical covariance structure (i.e. the covariance

---

1 Actually, even the borders of cytoarchitectonic areas can be gradual, as already noted by Brodmann himself: “the changes in the cortical layers are frequently not manifested suddenly and abruptly in a single region, but arise gradually over a broad area” (Brodman, 1909/1999, pp. 46). See also (Amunts and Zilles, 2015).
patterns at multiple levels of description) of the local connectivity patterns within and across brain areas. A first step toward achieving such characterisations is therefore to develop methods that can simultaneously characterise the local functional heterogeneity and multiplicity within brain areas. Data-driven tools such as Principal Component Analysis (PCA) or Independent Component Analysis (ICA) can already provide such descriptions through their ability to generate multiple spatial modes with the possibility of significant overlap. Here, however, we will motivate such multivariate techniques with the specific goal of characterising topographically ordered signals in order to model smooth and gradual variations across a spatial area.

3. Characterising functional heterogeneity and multiplicity through gradients

Over recent years, we (Blazquez Freches et al., 2020; Faber et al., 2020; Haak et al., 2018; Marquand et al., 2017; Navarro Schröder et al., 2015; Przedzicki et al., 2019), and others (Bajada et al., 2017; Cerliani et al., 2012; O’Rawe et al., 2019; Tian et al., 2019; Tian and Zalesky, 2018; Vos de Wael et al., 2018) have started to develop new analysis approaches that enable characterisations of, and statistical inference over, an area’s internal organisation using a novel class of methods that attempt to characterise an area’s internal organisation in terms of multiple overlapping modes of variation. In popular terms, they provide characterisations in terms of ‘gradients’ instead of ‘parcels’, where the gradients characterise the internal structure of the brain area of interest. Thus, these methods enable simultaneous characterisation of both the functional heterogeneity and multiplicity within a brain area. Importantly, the term gradient does not imply that the ensuing characterisations are necessarily gradual: they can—but do not have to—exhibit very rapid transitions (suggesting a boundary within the limits of the inherent smoothness of the data). They are referred to as gradients because their estimation is based on dissimilarity measures. As our work focuses on characterising the topographic organisation of functional connectivity, we try to avoid this possible confusion by referring to the mapping of the topographic organisation of connectivity: ‘connectopic mapping’.

The key idea behind these new ‘gradient’ or ‘connectopic mapping’ approaches is that, in the face of functional multiplicity, the problem of topographical localisation can be understood as a spatial variance decomposition problem. Just as music can be represented as a superposition of sine waves, spatial variance in brain structure, function and connectivity can be represented as a superposition of multiple modes of spatial change. Variance decompositions can be achieved in many different ways. This is because the decomposition requires constraints in order to yield a unique solution, and different approaches typically use different constraints (and thus yield different solutions). Factorising the number 9, for instance, has multiple possible solutions (e.g. 3 times 3, 9 times 1, and 2 times 4.5) unless the solution is constrained to consist of primes or e.g. multiples of 2. In similar vein, the decomposition can be constrained to involve an orthogonal linear transformation such as in principal component analysis (PCA). It turns out that linear decompositions of connectivity data do not work well in areas with known topographic maps such as primary motor and visual cortex, indicating that these maps are non-linearly embedded in the higher-dimensional space of spatially varying connectivity patterns (Haak et al., 2018). Hence, the gradient or connectopic mapping methods of choice involve non-linear variance decompositions (i.e. non-linear manifold learning) that can deal with both linearly and nonlinearly embedded data (Lee and Verleysen, 2007).

Most non-linear decomposition methods can be summarised as follows: first, data points (e.g. connectivity patterns) are represented as a graph (Fig. 1). The edge length between the graph nodes correspond to the distance between the data points and edges are deleted if they exceed some threshold or if they do not involve one of that node’s k-th nearest neighbours. Next, the variance decomposition is performed on the graph structure instead of the data (e.g. by performing an eigen-decomposition of the graph Laplacian instead of the data covariance matrix such as in PCA). By operating on a graph representation of the data rather than the data itself, these methods can capture the local instead of the global geometry of the data cloud. Conceptually, this can be understood as if these methods operate on the surface of a crumbled piece of paper: linear decompositions would find the main axes of the ball of paper, whereas the nonlinear approaches first unfold it to then find its length and width.

Importantly, all non-linear methods have a number of free parameters that can influence the results, and so can the selected decomposition algorithm (e.g. Belkin and Niyogi, 2003; Coifman et al., 2005; Donoho and Grimes, 2003; Ham et al., 2004; Roweis and Saul, 2000; Tenenbaum et al., 2000). In addition, one would need to make a choice as to how many overlapping modes will be considered further. Here it is also of note that the results are often subject to an orthogonality constraint, which renders the biologically validity of higher-order gradients increasingly unlikely. Therefore, it is crucial that the results are thoroughly validated, either by comparing them against ground-truth data (e.g. known retinotopic organisation in visual cortex) and/or by showing that individual variations in the ensuing gradients co-vary with another variable such as behaviour or genetic relatedness. Without such validation—which requires statistical inference—one cannot make strong statements about the biological validity of the results.

4. Statistical (spatial) inference over topographical maps

Much of what we know about the topographical organisation of brain...
areas is based on experiments that carefully map neural responses onto
systematic changes in some relevant stimulus dimension. For instance,
we can characterise the somatotopic organisation of sensorimotor cortex
by systematically stimulating individual body-parts and monitoring the
spatial location of the neural responses (Barch et al., 2013). Similarly, we
can monitor the spatial location of neural responses to visual expanding
ring and rotating wedge stimuli to reveal the retinotopic organisation of
visual cortex (Wandell and Winawer, 2011). The statistical significance
of the individual responses (i.e. our confidence that the neural signal
fluctuations are truly related to the experimental stimulus) can be char-
acterised using standard techniques. This type of statistical inference will
tell us whether the maps are meaningful. However, if we do not know
how to systematically stimulate a brain area of interest, but have instead
derived gradients using the decomposition approaches described above
based on some feature other than responses to systematic stimulus var-
iations (e.g. resting-state functional connectivity), we need to adopt a
different strategy to ascertain that the gradients are a meaningful char-
acterisation of brain organisation. This can be done by relating individual
variations in behaviour or other meaningful traits to individual variations
in the spatial layout of the gradients.

In principle, we could simply take each voxel’s or vertex’ gradient
value (i.e. eigenvector component after proper normalisation) and
correlate it against behaviour, but this would imply that we need to
correct for performing multiple comparisons. This issue—and the related
issue that the gradient values are not at all independent—can be avoided by
instead parameterising the spatial layout the gradient using an explicit
spatial model. That is, after the connectivity graph decomposition, the
ensuing modes or ‘gradients’ can be summarised in terms of a limited
number of parameters using spatial modelling methods to enable statisti-
cal inference over the area’s internal heterogeneity.

In our own work (Haak et al., 2018) we proposed an approach called
trend surface modelling (TSM) that is widely used in the geosciences
(Gelfand et al., 2010; Haas and Wackernagel, 1996). It is based on fitting
a polynomial basis set to the gradient values. Depending on whether we
are interested in inherently three-dimensional brain structures like
subcortical nuclei, or inherently two-dimensional surfaces like neocor-
tical areas, the polynomial basis set is defined in either volumetric (e.g.
MNI) or spherical coordinates, respectively. Note however, that the
basis set does not have to consist of polynomials, as it can also be formed
by an e.g. Fourier or Gaussian basis, or more theory-driven spatial bases
such as allowing for a logarithmic compression to model cortical
magnification in visual cortex (Schira et al., 2007; Schwartz, 1977). It
would even be possible to fit a parcellated atlas to the gradient and return
to a multiple-piece-wise constant area definition of each gradient while
still keeping the benefits of being able to model multiple overlapping
modes of organisation (Huertas et al., 2017). Alternative basis sets might
be considered if the spatial patterns cannot be captured adequately (e.g.
polynomial basis functions require relatively many polynomial terms for
modelling data with complex structure, leading to instability), if there are
specific modelling requirements (e.g. for extrapolation), and if there are
biological reasons to assume specific basis sets. Regardless of the basis
set, the benefit of fitting a spatial model to the gradient is that it allows us
to perform classical statistical testing techniques on a limited number of
model parameters. If suitably chosen, these parameters would also be
independent and hence multiple comparison correction for inference on
gradients would simply both in terms of the number of correction and in
terms of accounting for covariations. It is also of note that the spatial
modelling enables us to analyse just the larger (or finer) spatial trends by
selectively omitting the parameter estimates of the higher (or lower)
polynomial terms.

An important issue that is inherent to the decomposition methods, is
that the sign of the gradients is arbitrary and that small amounts of noise
can cause it to flip. It is also possible that the ordering of corresponding
gradients can be different between subjects. Thus, before comparing
gradients across subjects using spatial modelling we must ensure that the
gradients are ‘aligned’ such that they all have the same sign and that each
subject’s gradients are arranged in comparable order. This issue is
entirely analogous to single-subject ICA decompositions of fMRI data and
and can be achieved in two ways. First, it is possible to correlate each sub-
ject’s gradients against a set of reference gradients and then flip signs and
reorder until the sum of correlations is maximised. Alternatively, we can
estimate the gradients first at the group level (or in another dataset) and
then dual-regress (Nickerson et al., 2017) them into the individual
datasets (i.e. use the group or template maps as spatial regressors and
regress them onto the individual connectivity or time-series data, then
regrid the ensuing temporal dynamics to find the subject-specific spatial
maps). The first approach enables statistical inference on gradients that
are estimated directly in each individual, whereas the second implies
statistical inference over individual manifestations of the group-level
gradients. Note that the latter approach can also be used to map gradi-
ents in noisier datasets where individual-level gradient estimation fails
but group-level estimation does not.

One crucial advantage of explicit spatial statistical modelling is that it
enables classical statistical testing such as linear regression analysis,
tests and ANOVAs to be conducted on the gradients. In this way, it is
possible to establish that the spatial layout of the gradients expresses
meaningful variance related to behaviour, experimental conditions or
group differences. Besides these applications, however, the spatial sta-
tistical modelling also enables interesting model comparisons. In
particular, it enables statistically principled comparisons between
different models of spatial organisation such as smooth versus sharp
transitions. This means that it can be used to determine whether the data
provide evidence of area boundaries or not. For example, say we are
given an estimated gradient in some region-of-interest encompassing two
biological areas. If the two areas have clearly distinct features (e.g.
connectivity fingerprints), and if the imaging data is sufficiently precise,
the estimated gradient will exhibit a rapid transition at the area bound-
ary. We can now use spatial modelling and alternative hypothesis testing
to determine whether the gradient can be taken as evidence of an area
boundary by comparing the fits of models with smooth basis functions
against models with sharp (step-wise) basis functions. Smooth models
such as those based on polynomial or Fourier basis expansions will
require a much greater model complexity (i.e., many bases with non-zero
weights) than models that allow for step functions to model rapid tran-
sition accurately.

5. An example study

Numerous studies have applied the gradient estimation procedures in
a number of cortical and subcortical areas using a variety of imaging
modalities (Bajada et al., 2017; Blazquez Freches et al., 2020; Cerliani
et al., 2012; Faber et al., 2020; Marquand et al., 2017; Navarro Schröder
et al., 2015; O’Rawe et al., 2019; Przedziski et al., 2019, 2012; Tian et al.,
2019; Tian and Zalesky, 2018; Vos de Wael et al., 2018). One example
(Marquand et al., 2017) is particularly worth highlighting because it
underscores many of the points that were made here so far, and follows
the recommended procedural steps outlined above (Fig. 2). The work
considered the functional organisation of human striatum and we applied
connectomic mapping and spatial statistical modelling to this subcortical
nucleus in the basal ganglia. The striatum is a critical component of the
brain’s motor and reward system (Purves et al., 2008), and therefore of
great interest in investigations into neurological and psychiatric condi-
tions such as Parkinson’s disease and addiction. Our work was motivated
by the observation that existing functional connectivity studies into
human striatum often ended up with a sub-parcellation that did not
 correspond very well to characterisations of striatal organisation in

---

3 Each cortical surface has a sphere topology, so after ‘inflating’ the cortical
surface to a sphere each voxel or vertex can be identified by spherical co-
ordinates $(r, \theta, \phi)$. The radius $r$ can be omitted to indicate location on the surface of
the sphere, leaving just the polar angle $\theta$ and azimuthal angle $\phi$. 

---

K.V. Haak, C.F. Beckmann
NeuroImage 220 (2020) 117061
experimental animals. Indeed, the animal studies consistently described the presence of a rostro-caudal gradient relating to both nigrostriatal and cortico-striatal connectivity (Haber, 2003; Haber et al., 2000; Haber and Knutson, 2010). The human work, on the other hand, described the striatum in terms of subdivisions that appeared to follow a combination of the striatum’s anatomical constituents (the caudate nucleus, putamen, and nucleus accumbens) and a discretised version of the connectivity gradient that has been described in the animal literature (Choi et al., 2012; Cohen et al., 2009; Di Martino et al., 2008; Draganski et al., 2008; Haber, 2003; Leh et al., 2007). In the context of multiple overlapping gradients this suggested to us that the human functional connectivity studies had failed to account for the functional heterogeneity and multiplicity of the striatum.

We thus applied connectopic mapping to the striatum, using the resting-state functional MRI data of the Human Connectome Project. The ensuing dominant gradients confirmed our expectation that the striatal connectivity patterns observed in previous work were actually manifestations of at least two superimposed patterns of connectivity: one with rather sharp transitions that followed the distinction between the caudate nucleus, putamen and nucleus accumbens, and another gradual gradient that followed the expected rostro-caudal trajectory. By further projecting the rostro-caudal gradient—at this point isolated from the dominant anatomical gradient—onto cortex we were able to reproduce the cortico-striatal connectivity maps predicted in the animal literature. Encouraged by these results, we went on to demonstrate that the rostro-caudal connectivity gradient was meaningfully related to striatum-related behaviours. We employed spatial statistical modelling to summarise the gradient in terms of nine model parameters (i.e., a cubic model in three spatial dimensions) and submitted these to a canonical correlation analysis (CCA) to establish possible associations with an enormous battery of diverse behavioural measures. This analysis revealed that individual variations of the rostro-caudal connectivity gradient are related to individual variations in goal-directed behaviour such as delay discounting.

Thus, the striatum is an example of a brain area that exhibits meaningful functional heterogeneity and multiplicity that should be considered before averaging signal from across the striatum or in the context of network modelling studies. In addition, we believe that by decomposing striatal organisation into biologically meaningful components, we improved our ability to probing specific aspects of striatal functioning. This last point is also supported by ongoing work suggesting that yet higher order gradients in striatum can be linked to dopaminergic projections that would otherwise be obscured by the presence of the more dominant anatomical and rostro-caudal connectivity differences.
6. Relation with global gradient estimation

As our own work focused on mapping connection topographies within predefined brain areas in an attempt to learn about and account for the functional heterogeneity and multiplicity that can be present at these levels of description, others have started to apply similar non-linear variance decompositions of functional connectivity data to understand global brain organisation (Buckner and Margulies, 2019; Goulias et al., 2019; Guell et al., 2019, 2018; Hong et al., 2019; Huntenburg et al., 2018; Margulies et al., 2016; Paquola et al., 2019b, 2019a; Vos de Wael et al., 2019). Indeed, the methods can also be applied at the whole brain level to describe the overall constellation of brain areas in terms of their location in overlapping gradient dimensions. As such, the ensuing ‘gradients’ effectively provide a new basis set (i.e. coordinate system) for the topographical localisation of brain areas—the dominant gradients will be driven by factors that explain most of the variance, which at the global brain level is associated with inter-areal differences.

Perhaps the most prominent example of this line of work is a study by Margulies et al. (2016) who applied diffusion embedding to resting-state fMRI data to learn a coordinate system wherein the first dimension (i.e. gradient) puts brain areas within the brain’s default mode network far apart from primary sensory and motor cortex, and the second distinguishes between visual then default mode then motor cortex. Note that the large-scale gradient characterisations by Margulies et al. (2016) provide an alternative to the coordinate basis provided by whole-brain ICA. That is, the component maps (i.e., resting state networks) provided by ICA can also be viewed as forming a coordinate basis for the topographical location of brain areas. Thus, the global gradient maps are conceptually the same as the component maps derived with ICA, but differ in their appearance due to different algorithmic approaches optimising different cost functions. Which of these bases best reflects the intrinsic organisation of the human brain at the global level is an open question that may be answered comparing the different decompositions in terms of their ability to predict meaningful behavioural variables. Though ICA has been shown to yield similar decompositions when the input data are resting-state fMRI time-series and when derived from thousands of task-activation maps in the BrainMap database (Smith et al., 2009), and the global connectivity gradients based on diffusion embedding have been shown to reflect a functional hierarchy of increasing abstraction in meta-analytic topic terms (Margulies et al., 2016), neither have so far been validated explicitly (and against each other) in terms of their ability to predict inter-individual differences in behaviour.

7. Future directions

As of yet, we have only just begun exploring the possible functional heterogeneity and multiplicity of brain areas and further methodological developments are necessary to deal with a number of loose ends. First is the question of how many higher order gradients are meaningful. This, as mentioned above, is in part an empirical question that can be answered on a case by case basis by assessing a gradient that puts brain areas in an attempt to learn about and account for the functional heterogeneity and multiplicity that can be present at these levels of description. Generalising, however, it is a problem of intrinsic dimensionality estimation, which currently remains unsolved in the manifold learning literature. One possible approach is to compare the empirical eigenspectrum against the eigenspectrum of random matrices (e.g. as in MELODIC in FSL, Beckmann and Smith, 2004), but this approach first needs to be extended to estimating the intrinsic dimensionality of manifolds that are nonlinearly embedded in a higher dimensional space. A related open problem pertains to the fact that most decompositions lead to orthogonal representations, leading to decreased degrees of freedom for the estimation of higher order gradients and thereby decreased likelihood that they too are meaningful characterisations of brain organisation. It is therefore worth exploring the utility and validity of methods without orthogonality constraints.

Secondly, while connectopic mapping allows for eliciting and modelling the spatial pattern functional heterogeneity, it does not quantify the amount of heterogeneity. Developing quantitative measures of the amount of functional heterogeneity could be important to determine whether it be reasonable to summarise signals into a single estimate, or to determine if modelling the internal structure of a region is worthwhile at all (i.e. as an alternative to post-hoc validation). Previous work (Zang et al., 2004) proposed methods to quantify regional homogeneity (i.e., the complement of heterogeneity) using Kendall’s coefficient of concordance, but this approach does not account for the possible presence of overlapping modes of organisation. A more principled approach would be one that leverages the fact that connectopic mapping involves a variance decomposition, and that it is thus possible to estimate the relative amount of variance explained by each gradient based on the corresponding eigenvalues (Bajada et al., 2020). However, the key issue is that the notion of functional heterogeneity is inherently relative, and its quantification therefore requires baseline comparisons. From this perspective it becomes clear that the problem is related to the question of how to determine the number of gradients that are potentially meaningful: it requires understanding the spectral distributions of random data that are nonlinearly embedded in higher dimensional space. While we can probably rely on the results of random graph theory (i.e., for nonlinear decompositions that operate on graph representations), these ideas are still in their infancy and require further study and validation before they can be applied.

Thirdly, it will be important to integrate gradient estimation and region parcellation into a single framework. Although the gradient estimation procedures are relatively robust to small differences in region definition (Haak et al., 2018), inaccuracies in region definition can affect the results. The most pronounced differences in results can be expected when one region definition does and another does not include patches of cortex with very different connectivity patterns. If included, most of the variance can be explained by the very different connectivity patterns yielding gradients with very rapid transitions, and if not, most of the variance can be explained by more subtle and gradual modes of variation. These features can be quantified using spatial statistical modelling, and it is therefore conceivable that region definition and gradient estimation are iteratively refined in a single automated estimation procedure. However, this procedure has yet to be developed and validated.

A fourth important direction for future work is to integrate the gradient characterisations in brain network estimation. As mentioned, in the face of functional heterogeneity and multiplicity, it is often nonsensical to average signals and pretend they can offer a meaningful approximation of the interactions between brain areas. Rather, accurate network modelling requires accounting for the complex hierarchical covariance structure governed by both the multidimensional topographical organisation within and between areas. Gradient estimation and spatial modelling methods provide a means to understanding this structure, but have yet to be incorporated into a framework for network estimation.

Finally, it is worth noting that gradient mapping may lead to new methods for learning in what way the spatial organisation of specific brain areas support specific cognitive functions in the same way as retinotopic organisation in visual cortex supports spatial vision. For many brain areas, it is still unclear what functions they support, let alone how we should vary an experimental stimulus or task to evoke responses at specific topographical locations. Beyond the sensorimotor cortices, functional differences are likely very abstract but could still be organised along orderly topographic representations (Behrens et al., 2018; Garvert et al., 2017; Thivierge-Dugourd et al., 2007). Rather than employing gradient mapping, we can derive possible spatial dimensions along which we can expect functional differentiation—this in turn can be exploited to reverse engineer that area’s functional properties, e.g. through exploring the association of inter-individual variations on gradient configuration vs inter-individual variations across behavioural or demographic variations. Big Data resources such as the Human Connectome Project (Van Essen et al., 2013) or the imaging extension to UK Biobank (Miller et al., 2016) along with extensive behavioural and demographic measures have a
crucial role to play and enabling such validation work.

8. Conclusions

We have argued that the assumption of piece-wise constant brain areas is generally violated and that the failure of this assumption has profound implications for how we can go about analysing neuroimaging data if we wish to understand brain connectivity and function. Indeed, we believe that a substantial fraction of inconsistent findings in the brain connectomics literature can be attributed to the failure of this assumption (see e.g. Fig. 5 of Haak et al., 2018). We have further argued that the only way around this problem is to start characterising the functional heterogeneity and functional multiplicity of brain areas in an attempt to account for the complex hierarchical covariance structure—both within and across brain areas. Recent gradient methods such as connectomic mapping and spatial statistical modelling represent a step in this direction. These methods have already yielded important new insights, in part because they are employed in the context of very large datasets with enormous batches of meaningful behavioural measures. While many methodological issues remain to be resolved, it is crucial to keep in mind that the results of these methods must be thoroughly validated to ascertain that they reflect biologically meaningful characterisations for a better understanding of brain organisation in the face of functional heterogeneity and functional multiplicity.

CRediT authorship contribution statement

Koen V. Haak: Writing - original draft, Writing - review & editing. Christian F. Beckmann: Writing - review & editing.

Acknowledgments

The authors gratefully acknowledge funding from the Netherlands Organisation for Scientific Research (NWO-Veni 016.171.068 to K.V.H. and NWO-Vidi 864.12.003 to C.F.B.) and the Wellcome Trust UK Strategic Award [098369/Z/12/Z]. C.F.B. further gratefully acknowledges support from the Netherlands Organisation for Scientific Research under the Gravitation Programme Language in Interaction (grant 024.001.006).

References

Amunts, K., Zilles, K., 2015. Architectonic mapping of the human brain beyond Brodmann. Neuron 88, 1086–1107. https://doi.org/10.1016/j.neuron.2015.12.001.
Bajada, C., Campo, L.Q.C., Caspers, S., Muscat, B., Lambon, R., Parker, G., Cloutman, L.L., Trujillo-Barreto, N., 2020. A tutorial and tool for exploring feature similarity gradients with MRI data. Preprint. https://doi.org/10.1101/7v9zk.
Bajada, C.J., Jackson, R.L., Haroon, H.A., Azadbakht, H., Parker, G.J.M., Lambon Ralph, M.A., Cloutman, L.L., 2017. A graded tractographic parcellation of the temporal lobe. Neuroimage 155, 503–512. https://doi.org/10.1016/j.neuroimage.2017.04.016.
Barch, D.M., Burgess, G.C., Hamrs, M.P., Petersen, S.E., Petersen, S.E., Schlaggar, B.L., Corbetta, M., 2003. Geometric diffusions as a tool for harmonic analysis and structure determination of data: diffusion maps. Proc. Natl. Acad. Sci. Unit. States Am. 100, 7152–7157. https://doi.org/10.1073/pnas.1220731100.
Beckmann, C.F., Smith, S.M., 2004. Probabilistic independent component analysis for high-dimensional data. Proc. Natl. Acad. Sci. Unit. States Am. 101, 3103–3108. https://doi.org/10.1073/pnas.0400375101.
Buckner, R.L., Margulies, D.S., 2019. Macroscale cortical organization and a default-like apex transmodal network in the mammal brain. Nat. Commun. 10 https://doi.org/10.1038/s41467-019-09812-8, 1976.
Buhmann, J., 1999. Vergleichende Lokalisationselehre der Grosshirnrinde. Johann Ambrosius Barth, Leipzig.
Blazquez Freches, G., Haak, K.V., Bryant, K.L., Schurz, M., Beckmann, C.F., Mars, R.B., 2020. Principles of temporal association cortex organisation as revealed by connectivity gradients. Brain Struct. Funct. https://doi.org/10.1007/s00429-020-02047-0.
Brockhaus-Kram, E., 1909. Vergleichende Lokalisationselehre der Grosshirnrinde. Johann Ambrosius Barth, Leipzig.
Paquola, C., Bethlehem, R.A., Seidlitz, J., Wagstyl, K., Romero-Garcia, R., Whitaker, K.J., Olson, C.R., Musil, S.Y., 1992. Topographic organization of cortical and subcortical

Lee, J.A., Verleysen, M. (Eds.), 2007. Nonlinear Dimensionality Reduction, Nonlinear

Nickerson, L.D., Smith, S.M.,

Huertas, I., Oldehinkel, M., van Oort, E.S.B., Garcia-Solis, D., Mir, P., Beckmann, C.F.,

Harvey, B.M., Klein, B.P., Petridou, N., Dumoulin, S.O., 2013. Topographic representation of the human entorhinal cortex. Elife 4.https://doi.org/10.7554/eLife.06032.

Tinsley, C.J., 2009. Creating abstract topographic representations: implications for coding, learning and reasoning. Biosystems 96, 251 – 268. https://doi.org/10.1016/j.biosystems.2008.09.002.

Tian, Y., Zalesky, A., 2017. Characterizing the functional connectivity diversity of the insula cortex: subregions, diversity curves and behavior. Neuroimage 183, 716 – 733. https://doi.org/10.1016/j.neuroimage.2018.08.055.

Tian, Y., Zalesky, A., Bousman, C., Everall, I., Pantelis, C., 2019. Insula functional connectivity in schizophrenia: subregions, gradients, and symptoms. Biol. Psychiatr Neuroimage 183, 716 – 733. https://doi.org/10.1016/j.neuroimage.2018.08.055.

Tian, Y., Zalesky, A., Bousman, C., Everall, I., Pantelis, C., 2019. Insula functional connectivity in schizophrenia: subregions, gradients, and symptoms. Biol. Psychiatr Neuroimage 183, 716 – 733. https://doi.org/10.1016/j.neuroimage.2018.08.055.

Tian, Y., Zalesky, A., Bousman, C., Everall, I., Pantelis, C., 2019. Insula functional connectivity in schizophrenia: subregions, gradients, and symptoms. Biol. Psychiatr Neuroimage 183, 716 – 733. https://doi.org/10.1016/j.neuroimage.2018.08.055.

Tian, Y., Zalesky, A., Bousman, C., Everall, I., Pantelis, C., 2019. Insula functional connectivity in schizophrenia: subregions, gradients, and symptoms. Biol. Psychiatr Neuroimage 183, 716 – 733. https://doi.org/10.1016/j.neuroimage.2018.08.055.

Tian, Y., Zalesky, A., Bousman, C., Everall, I., Pantelis, C., 2019. Insula functional connectivity in schizophrenia: subregions, gradients, and symptoms. Biol. Psychiatr Neuroimage 183, 716 – 733. https://doi.org/10.1016/j.neuroimage.2018.08.055.

Tian, Y., Zalesky, A., Bousman, C., Everall, I., Pantelis, C., 2019. Insula functional connectivity in schizophrenia: subregions, gradients, and symptoms. Biol. Psychiatr Neuroimage 183, 716 – 733. https://doi.org/10.1016/j.neuroimage.2018.08.055.

Tian, Y., Zalesky, A., Bousman, C., Everall, I., Pantelis, C., 2019. Insula functional connectivity in schizophrenia: subregions, gradients, and symptoms. Biol. Psychiatr Neuroimage 183, 716 – 733. https://doi.org/10.1016/j.neuroimage.2018.08.055.

Tian, Y., Zalesky, A., Bousman, C., Everall, I., Pantelis, C., 2019. Insula functional connectivity in schizophrenia: subregions, gradients, and symptoms. Biol. Psychiatr Neuroimage 183, 716 – 733. https://doi.org/10.1016/j.neuroimage.2018.08.055.