Connectivity-based parcellation: critique & implications

Simon B. Eickhoff[1,2,*], Bertrand Thirion[3], Gaël Varoquaux[3], Danilo Bzdok[1,2,3,+,*]

Simon B. Eickhoff (s.eickhoff@fz-juelich.de)
Danilo Bzdok (danilo.bzdok@fz-juelich.de)
[1]Institut für Neurowissenschaften und Medizin (INM-1),
Forschungszentrum Jülich GmbH,
52425 Jülich, Germany
&
[2]Institut für klinische Neurowissenschaften und Medizinische Psychologie
Heinrich-Heine Universität Düsseldorf
40225 Düsseldorf, Germany

Bertrand Thirion (bertrand.thirion@inria.fr)
Gaël Varoquaux (gael.varoquaux@inria.fr)
Danilo Bzdok (danilo.bzdok@inria.fr)
[3]Parietal team, INRIA
Neurospin, bat 145
CEA Saclay
91191 Gif-sur-Yvette, France

[+]Corresponding author: Danilo Bzdok (danilo.bzdok@inria.fr)

[*]Authorship determined by coin flip

For Peer Review

Abstract

Regional specialization and functional integration are often viewed as two fundamental principles of human brain organization. They are closely intertwined because each functionally specialized brain region is probably characterized by a distinct set of long-range connections. This notion has prompted the quickly developing family of connectivity-based parcellation (CBP) methods in neuroimaging research. CBP assumes that there is a latent structure of parcels in a region of interest (ROI). First, connectivity strengths are computed to other parts of the brain for each voxel/vertex within the ROI. These features are then used to identify functionally distinct groups of ROI voxels/vertices. CBP enjoys increasing popularity for the in-vivo mapping of regional specialization in the human brain. Due to the requirements of different applications and datasets, CBP has diverged into a heterogeneous family of methods. This broad overview critically discusses the current state as well as the commonalities and idiosyncrasies of the main CBP methods. We target frequent concerns faced by novices and veterans to provide a reference for the investigation and review of CBP studies.

Keywords

Brain parcellation, clustering, resting-state correlations, diffusion MRI, data-driven, statistical learning, statistical inference, double dipping

Introduction

The human brain is commonly assumed to be organized in distinct modules (Brodmann, 1909; Vogt and Vogt, 1919). These could be described according to structure, connectivity, and function. *Cortical areas* can be conceptualized as patches of the brain that differ from their neighbors in terms of their microarchitecture (e.g., cyto-, myelo- and receptorarchitecture), connectivity (i.e., set of input and output connections), and function (e.g., lesion-induced behavior or electrophysiological responses) (Felleman et al., 1991; Van Essen, 1985). The conjunction of i) input and output *connectivity* of a cortical area and ii) its local infra*structure* is thought to crucially determine what classes of computational problems (i.e., *function*) it can solve (Scannel et al., 1995; Mesulam 1998; Passingham et al., 2002; Saygin et al., 2012).

The correspondence between a cortical area and its axonal connectivity fingerprint has prompted connectivity-based parcellation (CBP) approaches (Behrens and Johansen-Berg, 2005; Wiegell et al., 2003). Capitalizing on the distinct connections of each area (Passingham et al., 2002), CBP divides a region of interest (ROI, i.e., a volume- or surface-based topographical definition) into distinct subregions. The key idea is to first compute a connectivity profile for each individual voxel or vertex in the ROI. The voxel/vertex-wise connectivity profiles are then used to group the ROI voxels/vertices such that connectivity is similar for the voxels/vertex within a group and different between groups. That is, distinct clusters are identified in the ROI by differences between long-rang interaction patterns of the voxels/vertices in the ROI. Historically, CBP has first been performed based on whole-brain structural (fiber) connectivity profiles as derived from diffusion magnetic resonance imaging (dMRI) (Behrens et al., 2003; Wiegell et al., 2003). Later, analogous approaches based on resting-state functional connectivity (RSFC) (Kim et al., 2010) and, most recently, meta-analytic connectivity modeling (MACM) (Eickhoff et al., 2011) have been introduced. On the one hand, previous investigations have demonstrated that CBP can reveal clusters that recover known histological parcellations (e.g., Bzdok et al., 2012; Johansen-Berg et al., 2005). On the other hand, there are also reports showing that CBP may yield more fine-grained subdivisions than classical cytoarchitectonic mapping (e.g., Clos et al., 2013). Hence, CBP-derived modules may be viewed as 'functional areas', although these are outlined by connectivity differences rather than function.

The ability of CBP methods to map functional areas led to rapid adoption by neuroimaging investigators (cf. Smith et al., 2013a). Yet, several circumstances encourage heterogeneity in this nascent field. Methodologically, the CBP procedure is based on practical choices inconsistent across laboratories. Importantly, no single package permitting CBP is, to the best of our knowledge, openly distributed at the moment. Rather, it seems that different research groups perform CBP analyses based on their own script library, in-house databases, and laboratory setups. However, sharing of code implementations and international collaboration on its successive improvement will hopefully contribute towards a widely accepted software infrastructure (cf. Pradal et al., 2013).

Challenges that typically arise in CBP studies will be discussed in different sections. We will start out with the purpose of CBP and the neurobiological conclusions that can be drawn from it. The subsequent sections deal with the initial, interrelated decision on the ROI definition and the connectivity aspect to be investigated. We will then outline the main clustering approaches and corresponding cluster-selection criteria. The ensuing CBP results frequently raise questions around statistical inference and double dipping, discussed in later parts of the manuscript. We finally reflect possible ways to capitalize on CBP results as a starting point for multi-modal

For Peer Review

studies. All these issues are discussed below by providing overview, potential pitfalls, and possible solutions aimed at neuroimaging novices and experts. We hope that these considerations may ultimately help unify the dynamically developing *spectrum of CBP methods*.

For Peer Review

I Motivation

1.1 Aim

The principle of brain segregation guided by long-distance connections can be attractive from different perspectives, including the investigation of local functional differentiation, the creation of data-driven brain atlases, and catalyzing the inception of unprecedented hypotheses.

Location mapping. Comparing to other current neuroimaging approaches, CBP has the key strength to actually *map* distinct brain areas. This can be opposed to either localizing a particular (dys)function or characterizing a particular region. Most whole-brain association studies, be it functional MRI, voxel-based morphometry, lesion mapping, or most resting-state connectivity analyses, primarily elucidate *the location* in the brain of a particular effect such as the recruitment by a particular task, a differential response between two conditions, a difference between two groups, or the association with a particular phenotype. They are mapping cognitive, behavioral, and clinical aspects onto the brain. They do however not allow constructing a map from the brain itself (cf. Amunts et al., 2014 ~~for a more detailed discussion~~). Put different, most neuroscientific methods associating behavior with aspects of neurobiology are naïve to underlying neurobiological compartments. While observing mappings between behavior and the brain, these methods are not well suited to establish or question the architecture of the brain itself (Frackowiak and Markram, 2015). That is, rather than providing a map of the brain, they provide a map of a particular functional or structural feature (such as recruitment by a particular task or aberrations in a particular group of patients) in brain space. The potency of brain-behavior interpretations can however be increased when constrained by knowledge of brain organization units (Devlin and Poldrack, 2007). CBP can propose such organizational units.

Altas mapping. CBP methods are capable of automatically compartmentalizing the human brain into topographically delineated, functionally distinct regions (Behrens and Johansen-Berg, 2005). That is, 3D brain atlases can be obtained as quantitative models of brain segregation. In that context, an atlas represents a map of (parts of) the brain that assigns each location (voxel/vertex) to a particular structure and hence provides a segregation of the assessed volume into distinct modules. In *whole-brain CBP*, the ROI to be segregated covers the entire gray matter. By evaluating connectivity strengths from each gray-matter voxel/vertex to every other gray-matter voxel/vertex, a compartmental model of functional organization in the cerebral cortex can be derived. In *local CBP*, the ROI to be segregated covers a circumscribed part of gray matter. It can thus be evaluated whether that brain patch contains functionally distinct modules. As another important CBP variant, a-priori hypotheses can be introduced by measuring connectivity only to preselected brain regions, instead of the whole brain (e.g., Behrens et al., 2013; Bach et al., 2011; Sallet et al., 2013). In sum, CBP can readily propose 3D models of brain organization for use as reference atlases.

Hypothesis generation. Many currently employed neuroimaging methods test spatial hypotheses by either localizing effects or characterizing a region, instead of providing explicit hints that encourage novel research hypotheses (cf. Biswal et al., 2010). CBP may be seen as an approach towards the generation of novel hypotheses on regional differentiation. These can subsequently be tested in hypothesis-driven experimental studies. For instance, exploratory CBP evidence supported existence of distinct subregions in the right temporo-parietal junction (Mars et al, 2012). This was subsequently confirmed by targeted neuroimaging studies based on cognitive fMRI experiments (Silani et al., 2013), ICA-based experiments (Igelstrom et al., 2015), hypothesis-

For Peer Review

driven meta-analysis (Krall et al., 2015), quantitative reviews (Schurz et al., 2014), as well as multivariate pattern analysis in clinical populations (Koster-Hale et al., 2013).

1.2 Neurobiological meaning

What does it actually mean to divide the brain based on differences in connectivity profiles? CBP performs a systematic summary of the Cortex cerebri by combining same tissue and separating different tissue according to an organizational criterion, namely, brain connectivity. Analogous to cytoarchitectonic mapping by microanatomical criteria (cf. Brodmann, 1909, pages 5 and 288-290), functional mapping by connectional criteria critically depends on the certainty that we have about the divisive criterion.

Cortical areas. From an anatomical perspective of brain segregation (Amunts et al., 2013), cortical areas are believed to be distinguishable from their neighbors by featuring a distinct (micro)structure, distinct connectivity, and distinct function. In fact, function may follow naturally given that structure and connectivity are thought to conjointly enable *locally specific* neuronal computations (Passingham et al., 2002). As CBP is based on connectivity (true in a strict sense only for dMRI-CBP, cf. below), the defined clusters are not directly interpretable as cortical areas. Note that the current concepts of what constitutes a cortical area are mainly derived from studies of early sensory (Van Essen et al., 1992) and motor (Rizzolatti et al., 1988) brain systems. They may not be readily applicable to higher-level associative brain areas (Yeo et al., 2011), such as the dorsomedial prefrontal cortex (cf. Eickhoff et al., in press). Indeed, with increasing distance from sensory input processing, it is more and more difficult to relate the connectivity pattern of an area to its functional roles (Bzdok et al., 2013a; Bzdok et al., in press; Mesulam 1998; Yeo et al., 2011). Claims about cortical areas based on CBP results may therefore become more and more delicate with increasing level in the cerebral processing hierarchy.

One single versus multiple parallel subdivisions. From a more methodological perspective of brain segregation, CBP does not address the neurobiological question whether there is a 'true' parcellation. It is employed to identify the 'optimal' clustering solution, in the sense of best describing the data. It is about the question whether different parcellation results for the same ROI capture different resolutions or dimensions of an underlying neurobiological organization (cf. Kelly et al., 2012; Amunts et al., 2014; Eickhoff and Grefkes 2011). The answer depends on the region of interest. For instance, previous CBP work on the insula (Kelly et al., 2012; Nanetti et al., 2009) and the right temporo-parietal junction (Bzdok et al., 2013b; Mars et al, 2012) have indicated close agreement between the parcellations based on different connectivity modalities. In contrast, parcellations of the posteromedial cortex diverged more strongly between dMRI- and RSFC-CBP studies (Cauda et al., 2010; Zhang and Li, 2012; Zhang et al., 2014). These observations corroborate the relevance of the conceptual differences between aspects of connectivity, such as dMRI, RSFC, and structural covariance. Moreover, there is probably no such thing as '*the* connectivity' for a particular location. There may neither be '*the* CBP parcellation.' Rather, brain segregation across connectivity modalities can possibly feature both similarity and dissimilarity.

Multi-modal comparison. Unfortunately, there are yet very few studies that address this fundamental question of brain organization. Such a comparison across connectivity modalities is currently impeded by two key factors. On the one hand, results from previous CBP studies are rather inconsistently available to the community as

image files, which renders most attempts to compare findings purely qualitative (Gorgolewski et al., in press). On the other hand, there appears to be a sentiment that a new CBP study in an already analyzed brain region is primarily a replication and hence lacking novelty. This discourages additional work on previously parcellated brain regions.

Meaning of CBP clusters. Given the biological and methodological characteristics of the most frequently used anatomical and functional connectivity measures, we would suggest the following tentative distinction. For dMRI-CBP, the delineated clusters most likely reflect truly connectivity-defined modules, even in light of known artifacts (cf. below). In contrast, MACM-CBP reveals clusters that are probably functionally distinct modules even though the spectrum of brain functions is likely to be larger than what can be probed by neuroimaging techniques (cf. Mennes et al., 2013; Smith et al., 2009). That is, task-based functional connectivity might be limited by real-world behavior being richer than in-scanner behavior. The neurobiological nature of RSFC-CBP derived clusters might remain most uncertain. This is because the relation of resting-state correlations to anatomical connectivity, function, and the brain's housekeeping physiology is currently only incompletely understood (Biswal et al., 1995; Zhang and Raichle, 2010).

The insula. To take a concrete example, previous CBP results match some neurobiological dimensions of the insula, but certainly not all of them (Fig. 1). On a microanatomical scale, investigations classically divided the insula into a rostroventral (agranular), rostrodorsal (dysgranular), and caudal (granular) portions in macaque monkeys (Mesulam and Mufson, 1982a). On a macroanatomical scale, the more rostral insula is preferentially connected to frontal regions, whereas the more caudal insula is preferentially connected to primary and secondary sensory as well as motor regions (Mesulam and Mufson, 1982b; Mufson and Mesulam, 1982a). On a developmental scale, anterior-posterior segregation in the human insula becomes observable within the first two years of life, as indicated by RSFC-CBP in infants (Alcauter et al., 2013). From the perspective of sensory input channels, the insula contains primary gustatory, primary auditory, as well as associative somato- and viscerosensory cortices. On a functional scale, along the caudo-rostral insula, primary interoceptive representation gradually shifts over environmental input representation into highly abstract cognitive representations of self and time (Craig, 2009). These observations exemplify that an identical ROI may be segmented along diverging features and notions of brain organization.

More globally, it appears that agreement across connectivity modalities decreases when the parcellation becomes more fine-grained. In this context, it is important to appreciate that many brain regions may be described at multiple scales and by multiple notions. It is hence likely that there are several correct answers to the question of a neurobiologically valid parcellation, even when based on a single approach. This has probably best been demonstrated for the insula, subject to repeated CBP analyses (Nanetti et al., 2009, Chang et al., 2013, Kelly et al., 2012, Cauda et al., 2012, Deen et al., 2011, Jakab et al., 2011, cf. also Kurth et al., 2010). This previous work has shown that the insula may be described by a primary rostral-caudal distinction (cf. Alcauter et al., 2013) as well as a repeatedly reported a triplet of rostroventral, rostrodorsal, and caudal portions. Diverse functional recruitments and more fine-grained parcellation schemes, such as described by Kelly et al. (2012) and Nanetti et al. (2009), should then reflect additional differentiation within these.

Hierarchical level, functional gradients, and completeness. Three further aspects that need to be considered in any neurobiological interpretation of CBP results are hierarchical level, functional gradients, and completeness. i) Generally, boundaries between brain regions become less clear with increasing abstraction level in the neural processing hierarchy. This is reflected by the fact that the more similar the connectivity patterns of two areas are, the more difficult it is to demonstrate a functional double dissociation by lesion studies (Young et al., 2000). In a CBP context, particular care and modesty is therefore recommended when investigators interpret functional borders in highly associative brain regions. ii) Both high- and low-level processing regions in the brain may feature dedicated functional gradients. For instance, the left inferior parietal lobe (i.e., a high-level region) might contain a functional gradient from more person-state- to more person-trait-related processing facets in social judgments (Hensel et al., in press), while V1 (i.e., a low-level region) contains functional gradients related to retinotopy ~~and ocular dominance~~ (Wandell et al., 2007). Whenever there is previous evidence for functional gradients in a ROI, investigators should be careful not to overstretch the discovered functional mapping. This is because the commonly used clustering algorithms (e.g., k-means, spectral, and hierarchical clustering) will impose clear-cut boundaries somewhere along such gradual transition zones. iii) It is moreover noteworthy that the distinct functional modules identified in a locally circumscribed ROI might extend beyond the boundaries of that ROI (cf. next section).

Taken together, convergence and divergence across parcellation schemes should raise the attention of CBP investigators. It is conceivable that discussing CBP results exclusively by a single parcellation solution of the ROI might entail loss of neurobiological insight. Investigators should treat diverging parcellation solutions (at same and different cluster numbers) as potentially complementary rather than strictly exclusive. Indeed, the insula did feature several stable parcellation solutions in a dMRI-CBP study (Nanetti et al., 2009). To facilitate more comparison across modalities and CBP approaches, however, the neuroimaging field probably needs increased sharing of CBP parcellations and more complementary investigations of already examined regions using different connectivity measures and approaches. Put differently, connectivity-derived clusters are primarily descriptions of the data. Meaning of clusters can only arise in the adoption of a neurobiological viewpoint. They do not, however, simply represent 'cortical areas'. Hence, this term should probably be avoided until sufficient evidence has been gathered on distinctions from neighboring areas in terms of structural, functional and connectivity-related features. Moreover, clusters in the associative cortices and those identified in regions with evidence for functional or structural gradients need to be interpreted with particular caution.

For Peer Review

II Procedure

2.1 ROI definition

The first decision in CBP analyses is on the part of the brain to analyze. The ROI outlines the set of target gray-matter voxels/vertices that the investigator wishes to segregate into subregions. This important step operationalizes the investigation, which strongly impacts the overall outcome and interpretation of the study.

Whole-brain CBP. As the perhaps most intuitive choice, the ROI may comprise the entire gray matter in the aim of *whole-brain parcellation* (e.g., Craddock et al., 2012; Shen et al., 2013; Thirion et al., 2006b; Thirion et al., 2014; Yeo et al., 2011). A ROI with all voxels/vertices in gray matter can for instance be drawn from the ICBM tissue map (International Consortium on Brain Mapping) at a gray-matter probability of choice (e.g., Bzdok et al., 2013b), from gray-matter segmentation (Fischl et al., 2004), as well as from the MNI (Evans et al., 1992) or Talairach-Tournoux (1988) template spaces. Importantly, it might not always be the most attractive option to group the brain in voxel or vertex units. Single voxels/vertices can hardly be interpreted by themselves (cf. Chumbley & Friston, 2009). Additionally, operating in a voxel/vertex space can make clustering procedures computationally expensive. Rather than *voxel/vertex-level clustering*, whole-brain parcellation lends itself to *node-level clustering*. That is, not individual voxels/vertices but predefined groups of voxels/vertices (i.e., nodes) are the units that are grouped as a function of connectional similarity (Smith et al., 2013b). Note that the nodes represent voxel/vertex combinations based on previous knowledge that can, for instance, be derived from ICA or structural atlases. Constructing nodes as an alternative unit of observation is therefore not itself an instance of clustering. An advantage of performing a connectivity-based grouping of nodes covering the brain's gray matter relies in the increased neurobiologically interpretability. This is because these nodes are often created based on neurobiological features, whereas a single voxel/vertex does usually not allow a one-to-one mapping of salient neurobiological properties. For instance, using each region of the default-mode network as nodes allows finding node clusters that represent functionally distinct subnetworks (Andrews-Hanna et al., 2010). Such node definitions can be called *hard* (i.e., one single shape) or *soft* (i.e., several slightly different shapes dependent on occurrence likelihood) (Varoquaux et al., 2013). Concretely, subregions from hard ROI clustering are typically non-overlapping, whereas subregions from soft clusterings can typically be overlapping. Frequently used hard brain nodes include cytoarchitecture (Brodmann, 1909) and AAL (Tzourio-Mazoyer et al., 2002), while frequently used soft brain nodes include the probabilistic atlases from Jülich (microanatomical) and from Harvard-Oxford (macroanatomical). As a more data-driven variant of whole-brain parcellation, sets of coherent functional nodes can be obtained by (spatial) independent component analysis (ICA; Beckmann et al., 2005; Malherbe et al., 2014; cf. below). Yet, note that the optimal conceptualization of a ‘node’ is unclear and the practical choice is a matter of debate (Zalesky et al., 2010). A whole-brain atlas of functional nodes can also be learned directly from RSFC data of multiple subjects in a probabilistic hierarchical model (Varoquaux et al., 2011). Among whole-brain CBP approaches, one might further distinguish *3D-volume-based parcellation* (most studies cited in this paper) and *surface-based parcellation* (e.g., Yeo et al., 2011; Blumensath et al., 2013; Gordon et al., 2014). Further, whole-brain parcellations, be it based on individual gray-matter voxels/vertices or preset voxel/vertex groups as nodes, enjoy increasing popularity. Whole-brain CBP might be particularly important for current and future high-throughput projects in neuroscience (e.g., the European Human Brain

Project [HBP], the international Human Connectome Project [HCP], the American Brain Research through Advancing Innovative Neurotechnologies [BRAIN]) (a very good overview is given in Poldrack & Gorgolewski, 2014).

Regional CBP. In contrast to whole-brain CBP, the majority of existing CBP studies used a ROI that outlines a circumscribed part of the brain. The underlying motivation typically relates to a test of functional heterogeneity. Note however that whole-brain CBP provides individual clusters that can each be used as circumscribed ROI. Practically, cluster from whole-brain CBP studies can subsequently serve as targets for local regional CBP studies. One can distinguish between *anatomical* and *functional* ROIs. An anatomical ROI can be constructed in a straightforward fashion by manually outlining macroanatomical landmarks guided by gyri, sulci, ventricle borders, or white matter (e.g., Anwander et al., 2007; Beckmann et al., 2009; Mars et al., 2011; Solano-Castiella et al., 2010). One might note three relevant aspects. First, this type of ROI definition may be limited by the fact that sulcal/gyral boundaries do not always coincide with functional boundaries (Amunts et al., 1999; Zilles et al., 1997; in contrast to Weiner et al., 2014). Intentionally extending the ROI may therefore be an attractive option (cf. below). Second, these gross anatomical features can be subject to considerable inter-individual variability (Kochunov et al., 2010). This encourages region delineation on a single-subject basis. The presence or absence of the paracingulate sulcus may, for instance, not be captured by an automated group-level procedure (cf. Beckmann et al., 2009). Third, manually defined ROIs might be thought of as more subjective by some authors. While the ensuing CBP studies could suffer from poor reproducibility, a completely automatic method is however no guarantee for better results (cf. below). Note that all three presented caveats are controversial in the literature. As a frequently used alternative, macroanatomical ROIs may be based on probabilistic maps such as, e.g., provided by the Harvard-Oxford atlas (http://fmrib.ox.ac.uk/fsl/) or constructed by automatic segmentation (Fischl et al., 2004). Both strategies have frequently been employed in previous CBP studies (e.g., Bach et al., 2011). Such maps provide objective and reliable masks reflecting the location of a particular structure in a group of subjects that contrast investigator-guided or hand-drawn ROI definitions. Microanatomical ROIs, in turn, represent an attractive alternative because regional heterogeneity of histological features, such as cytoarchitecture (cf. Amunts and Zilles, 2010), is a likely indicator of regional specialization. Such probabilistic cytoarchitectonic ROIs have already been used for CBP studies (e.g., Bzdok et al., 2012; Johansen-Berg et al., 2005). It is however important to appreciate that cytoarchitectonic areas exhibit often marked inter-individual variability and hence may not map well to the corresponding regions of an individual subject following standard-space alignment. Consequently, the resolution of any CBP analysis can only be as accurate as permitted by the preceding realignment procedures. In summary, when using an anatomical ROI, the ensuing CBP analysis addresses the question whether a particular (macro- or microanatomical) structure contains subregions featuring distinct connectivity. More specifically, CBP performed on microstructural ROIs then yields modules that are defined by a particular structure (due to the ROI definition) and connectivity (due to the CBP logic) and may therefore be more likely to represent actual functional units in the brain.

Contrary to anatomical ROIs, neuroimaging researcher sometimes use alleged neuroanatomical terms to predominately denote a certain *function* rather than a certain *location*. Examples for such 'pseudo-anatomical' regions that often have a rather coarsely defined or even disputed relationship to structural anatomy would be the 'fusiform face area' (cf. Kanwisher et al., 1997), 'frontal eye field' (cf. Grosbras et al., 2002), and 'temporo-parietal junction' (cf. Mars et al., 2012). When interested in such region, a functionally defined ROI might be the

For Peer Review

preferred approach. It is possible to use the results of a single fMRI study. In its simplest form, it could be defined by voxels/vertices activated by one or more fMRI tasks. Yet, the functional definition would be highly specific to the respective experimental setup. A principled approach to create a functional ROI from experimental fMRI that acknowledges inter-subject variability is known as *functional localizers* (Friston et al., 2006; Fox et al., 2009; Saxe et al., 2006). Originally, this approach used a separate neuroimaging experiment performed to constrain both analysis (i.e., increasing sensitivity) and interpretation of the actual study. If inter-subject variability is not of interest, it can be an attractive option to consolidate the location of the functional process of interest by means of quantitative image-based (i.e., using whole-brain activation maps; Dehaene et al., 2003; Schilbach et al., 2008) or coordinate-based (i.e., using peak activation information; Eickhoff et al., 2009; Radua and Mataix-Cols, 2009; Wager et al., 2007) meta-analysis. The resulting ROI would then statistically constrain the most robust location of activation convergence underlying the process of interest across various subjects, study designs, and laboratories. In either case, CBP will address the heterogeneity of connectivity patterns within this functionally defined region. This becomes particularly interesting when ROIs are defined by (partially) overlapping activation blobs from experimental neuroimaging studies (Cieslik et al., 2013) or meta-analyses summarizing different functions (Bzdok et al., 2013b) that are located in closely neighboring yet potentially different locations. In cases where a 'composite' functional ROI is used, CBP allows answering a new type of question: 'Are the different cognitive processes reflected by the different activations or meta-analyses related to the same or different connectivity-defined modules in the human brain?' Note that this is not circular even if CBP always locates subregions in the ROI. This is because cluster validity criteria may provide evidence that all obtained cluster solutions are instable and therefore not neurobiologically meaningful. Such judgments should be weighed against external knowledge. CBP with composite ROIs has hence the potential to reconcile controversies in cognitive neuroscience. For instance, the one-module versus mosaic-modules debate for the temporo-parietal junction (Decety and Lamm, 2007; Mitchell, 2008) has probably been resolved by repeated demonstration of functionally distinct subregions using CBP (Bzdok et al., 2013b; Mars et al., 2012).

ROI borders. As an important consideration in CBP studies, the outside borders of the non-whole-brain ROI are not tested or validated. They will hence be taken as borders of the ensuing clusters. A thorough motivation of why and how these outside boundaries are defined is pertinent to any CBP study. In turn, if the localization of a particular module is the primary interest of an investigation, it is advisable to dilate the ROI to include sufficient coverage of the neighboring cortex, allowing for additional clusters around the volume of primary interest. That is, CBP may find all borders of the main regions of interest by including neighboring regions of no or limited interest (cf. Sallet et al., 2013; Muhle-Karbe et al., 2015). In principle, if the ROI extends only a little beyond the ground-truth area(s), only a neglectable amount of noise should be introduced into the cluster estimation. If the ROI extends beyond the ground-truth area(s) to extended parts of areas of no interest, then new clusters (of no interest) emerge that delineate the cluster(s) of interest. As an extreme scenario, if the ROI is incorrectly defined, interpretation of the clustering results becomes challenging to impossible.

In sum, different ways for anatomical or functional definitions of a ROI for CBP have been used and are legitimate. This choice and its operationalization should be well motivated. Generally, population or meta-analysis based ROIs are an alternative to hand-drawn ones or those that are based on a single-subject data (e.g., AAL). In principle, microanatomical ROI definitions can be preferable to macroanatomical ones. In case of

For Peer Review

lacking neuroanatomical consensus for the target region in the literature of interest, a functionally motivated ROI suggests itself. A functional ROI can be constructed from neuroimaging studies or activation convergence quantified by meta-analytic methods. Finally, 'composite' ROIs allow answering the specific question how a particular set of findings relate to regional specialization. Apart from that, if the aim is whole-brain parcellation, voxel/vertex-level CBP does less crucially depend on the ROI definition, while node-level CBP currently suffers from uncertainty about the most biologically valid reduction to network nodes and about methodological biases incurred by node choice (Zalesky et al., 2010). As a consequence, the motivation of the CBP study should be consolidated *before* selecting the preferred ROI. This is because the location and type of ROI explicitly frame the scientific question and motivation underlying a CBP study. Consequently, the initially selected ROI constraints the spectrum of permissible conclusions from the later CBP results. Critically, the decision on the ROI should be taken hand-in-hand with the connectivity data of interest. This is because the underlying neurobiological question should guide methodological choices.

2.2 Measures of brain connectivity

Note that the concepts underlying CBP are not bound to a particular connectivity approach. Any method can be employed that yields a connectivity profile for each voxel/vertex in the ROI. In general, the *anatomical connectivity* modality most frequently used for CBP-analyses is dMRI (e.g., Anwander et al., 2007; Behrens et al., 2003; Johansen-Berg et al., 2004). The *functional connectivity* modality in most frequent use is RSFC (e.g., Cauda et al., 2010; Kim et al., 2010; Zhang and Li, 2012). It has recently been complemented by MACM (e.g., Bzdok et al., 2012; Eickhoff et al., 2011), which is rapidly gaining usage in the field. As an alternative to anatomical and functional connectivity, structural covariance (Evans et al., 2013) has been used in a small number of CBP studies (e.g., Kelly et al., 2012; Wang et al., 2014). We do not cover the latter in the interest of simplicity and space. It is important to appreciate that all these measures of connectivity strength reflect drastically different ways to conceive and quantify interneuronal communication between brain regions. Choosing one of them is as important as the ROI selection and has far-reaching implications for the interpretation of the identified clusters. At this point it might be helpful to reiterate *that it is not the individual connectivity profiles of the voxels/vertices that drive the parcellation but only the differences between those*.

DMRI. *Anatomical connectivity* between brain regions can be measured (or rather approximated) using *diffusion magnetic resonance imaging*. It delineates the likelihood of white-matter fiber bundles traced to link brain regions (Johansen-Berg and Rushworth, 2009; Jones, 2008). The number of samples reaching any voxel/vertex in the gray matter or, more frequently, the likelihood of passing through brain white matter then provides the connectivity profile of a particular voxel/vertex or node in the ROI. In fact, in *whole-brain CBP* dMRI is seeded from every gray-matter voxel/vertex. In *local CBP* every voxel/vertex in the circumscribed ROI is a seed, while in *node-level CBP* voxel/vertex groups are seeds (goes for all three connectivity modalities). DMRI tractography is evidently closest to the notion of structural connections. Yet, it does not actually capture axonal connections as classically identified by axonal tracing studies in monkeys (cf. Mesulam, 1976). Caveats of tractography include i) the dominance of large fiber bundles, thus omitting sharply curved or very long fiber bundles, which precludes exhaustive assessment of all connections (Ng et al., 2013; Jbabdi and Behrens, 2013) as well as impaired

For Peer Review

detection of ii) poorly myelinated or iii) closely neighboring ('kissing') connections. Finally, *dMRI can neither precisely delineate cortical origin nor cortical termination of fiber bundles* (Petrides et al., 2012).

RSFC. Alternatively, *functional connectivity* can be measured by *resting-state correlations* under the assumption that the coupling strengths between distant brain regions is measurable by correlation between time series of BOLD signal fluctuations *outside of an experimental context* (Biswal et al., 1995; Buckner et al., 2013; Zhang and Raichle, 2010). It quantifies the correlative relationships between distant brain regions in subjects idling in the MRI scanner. This is possible because interneuronal communication continues and is reflected by ongoing physiological fluctuations in the absence of an experimentally imposed cognitive set, i.e., during natural mind wandering, which can be measured using fMRI (Bzdok & Eickhoff, 2015). While RSFC signals have been shown to recover well-documented axonal connections and functional networks, there is an increasing awareness that much of the observed signal may be influenced, if not distorted, by physiological sources (but see Hipp & Siegel). The ensuing conundrum may challenge the interpretation of brain-behavior relationships discovered by RSFC. Despite initial skepticism, the consistency of RSFC results has been demonstrated repeatedly across subjects, brain scans, time points, and other factors (Damoiseaux et al., 2006; Shehzad et al., 2009). RSFC thus provides proxies of dynamic neuronal interactions that might reflect mixtures of various cognitive processes and physiological factors (Smith et al., 2009; contrasted by Mennes et al., 2013).

MACM. *Meta-analytic connectivity-modeling*, in turn, operates under the assumption that *functional connectivity* between brain regions should entail reliable coactivation (Toro et al., 2008; Robinson et al., 2010). It quantifies correlative increase/decrease of neural activity in distant brain regions *throughout various experimental paradigms.* This connectivity modality capitalizes on the increasing trend for large-scale data aggregation, exemplified by Neurosynth (Yarkoni et al., 2011), BrainMap (Fox and Lancaster, 2002), and NeuroVault (Gorgolewski et al., in press). Caveats of MACM include i) reliance on very sparse activation information (i.e., peak coordinates of significant activation), which might entail missing information and biased sampling, ii) inability of subject-specific connectivity analysis, and iii) inheritance of the limitations from experimental neuroimaging studies. In spite of these limitations, the analysis of coactivation likelihoods represents a complementary approach by focusing on the interactions during the performance of externally purported tasks.

Commonalities and differences. Several aspects are of note when choosing between anatomical and functional connectivity modalities in a CBP study. None of the three introduced connectivity modalities provides axonal connectivity in stricto sensu (as gleaned from tracing studies in monkeys). dMRI and RSFC are *task-unconstrained* (i.e., *task-independent*) as opposed to *task-constrained* (i.e., *task-dependent)* MACM. While dMRI is a measure of *anatomical* or *structural connectivity* by assessing *white-matter trajectories*, RSFC and MACM identify *temporal coincidence* of *neural signals in gray matter*, that is, functional connectivity. MACM builds on experimental fMRI and PET studies motivated by cognitive theory (i.e., *interventional*, capturing metabolic changes in the brain caused by manipulation of environmental variables), whereas participants simply lie still during RSFC and dMRI measurements (i.e., *observational*, capturing baseline brain features without controlled environmental modulation). It may also be noted that none of these methods can distinguish between involved neurotransmitters (i.e., excitatory versus inhibitory neuronal modulations) or ask whether a connection

For Peer Review

is stronger in one direction (i.e., "undirected" connectivity). Moreover, we also need to point out that functional connectivity between two regions may be mediated by a third region. That is, RSFC and MACM (but not dMRI) may be driven by indirect connections. This could however be alleviated by computing partial correlations, which is closer to direct interaction by summarizing conditional independences (Marrelec et al., 2006). In fact, regression-based estimators, such as dictionary learning (Lee et al., 2011), instead of standard clustering approaches, may be more robust to the issue of third-party influences. Additionally, RSFC and MACM are generally more sensitive in delineating existing connections but more prone to false positives, whereas dMRI is generally less sensitive with frequent false negative results (cf. Jbabdi and Behrens, 2013). Contrarily to RSFC and MACM, the accuracy of dMRI tract tracings decreases with the distance between the considered regions. Only dMRI- and RSFC-CBP can be conducted in individual subjects (Anwander et al., 2007; Kim et al., 2010). dMRI- and RSFC-CBP thus enable detecting interindividual differences in regional functional organization, while MACM-CBP inevitably generalizes across various inter-individual variability sources in the sampled subject population. As an important consequence, dMRI and RSFC can readily parcellate individual brains and infer group aggregates based on between-subject variance. Contrarily, MACM is constrained to aggregating group-level statistics from meta-analytic experiment databases. Hence, MACM-CBP may provide information on the functional parcellation of a region of interest across many experiments, which often shows remarkable congruency with parcellations derived from other modalities, but does not allow any individual, subject-specific parcellation. Taken together, dMRI, RSFC, and MACM grasp different features of connectional brain organization and imply different limitations and promises.

Anatomical and functional connectivity measures are all equally valid for assessing connectivity strengths to perform CBP. dMRI, RSFC, and MACM all lend themselves to whole-brain, node-level, and local CBP. It is important, however, to remember that they are based on fundamentally different concepts of 'brain connectivity'. Roughly, dMRI is most ‘structural/physiological‘ in nature, whereas MACM is exclusively 'functional/psychological.' RSFC, in turn, most likely reflects a mixture of both (with a different set of physiological confounds). These considerations may guide the choice of the employed method when the motivation particularly relates to either functional or anatomical questions. Their different limitations and promises might yield conflicting views on the organization of the ROI, even though first comparative studies show a fairly close convergence (Kelly et al., 2012; Wang et al., 2014). Nevertheless, exploiting distinct connectivity modalities is likely to extend the space of questions that we can ask about functional brain organization.

2.3 Clustering techniques

Clustering uses a similarity measure to group a set of elements into subsets according to their measured similarity. In CBP, the clustering algorithm groups the voxels/vertices/nodes in the ROI into subsets according to similarity of their connectivity profiles, the heart of any CBP approach. As a result of the so-called '*no free lunch' theorem* (Wolpert, 1996), no clustering algorithm performs optimally for all ROIs, types of connectivity information, and study motivations. Rather, methods such as k-means, spectral, and hierarchical clustering have all been frequently employed in CBP studies. While these three clustering algorithms have been used for voxel/vertices-, node-, and whole-brain-level CBP, ICA and boundary detection are popular alternatives for

brain parcellation. We will detail in this section how these algorithms behave in theory and in neuroimaging practice.

==========

Box 1 about here please

==========

K-means. The probably most popular choice in neuroimaging is *K-means clustering* (Lloyd, 1957; Forgy, 1965; Jain, 2010), a *partitional* approach. It divides a ROI into a preselected number of k non-overlapping clusters (Nanetti et al., 2009). In neuroimaging practice, k-means seems to perform best when the subregions in the ROI are expected to be (i) few in number, (ii) of similar size, and (iii) featuring a roughly spherical shape on spatially correlated voxel/vertex/node-wise connectivity (cf. Jain, 2010). Additionally, k-means clustering will converge in the majority of the cases (i.e., seldom early stopping by the tolerance parameter). In a CBP context, the same connectivity data can describe not only one but several stable solutions in ROI parcellation at the *same* preset k (i.e., low reproducibility), such as observed in the human insula (Nanetti et al., 2009). Consequently, the algorithm is conventionally applied many times since k-means fits idiosyncracies in data that may generalize poorly across subjects. As a first practical consequence, the initialization of the cluster centers (cf. Box 1) can be random (Hartigan and Wong, 1979) or based on prior knowledge (e.g., anatomical properties). As a second practical consequence, the 'final' solution can be obtained by an averaging procedure or by selecting the centroids from the best solution (cf. Nanetti et al., 2009). Further, the solutions for different selections of k (i.e., different number of clusters) are independent of each other. Repeating the clustering at different k's does not emulate a hierarchical approach (contrary to hierarchical clustering). That is, the solutions for ROI parcellation at each level (k) are independent of the others, which makes parent-children stratifications possible but by no means necessary. As an attractive k-means variant to address the multi-scale nature of brain organization, investigators can first identify the best fitting k clusters and then test for further separability of each obtained cluster individually (Neubert et al., 2014).

Spectral clustering. One of the first clustering methods in the context of CBP (Johansen-Berg et al., 2004) has been *spectral clustering* (Donath and Hoffman, 1973; von Luxburg, 2007). It can be useful to semi-quantitatively obtain a possible number of clusters by inspection (e.g., Johansen-Berg et al., 2004; Bzdok et al., 2012). Alternatively, when applying an ordinary clustering algorithm, spectral clustering is able to capture clusters that have complicated shape and are discontinuous, yet that are enforced to be roughly equally sized (Craddock et al., 2012; von Luxburg, 2007). Note that the clustering solutions for different cluster numbers are not hierarchically consistent (analogous to k-means, contrary to hierarchical clustering). That is, nestedness of the resulting partitions of the ROI are not methodologically enforced but might still appear as a biological property of the ROI under study. Spatially constrained spectral clustering appears to be stable in capturing connectional similarity features between ROI voxels/vertices/nodes (i.e., high reproducibility). It might however not accurately represent those (i.e., poor model fit, Thirion et al., 2014). In CBP, spectral clustering might be disfavored by some investigators because it strives to enforce simple structure not naturally present in the brain. In particular, as the potentially biggest drawback of spatially constrained spectral clustering, it imposes a strong spatial structure on the data, which thus precludes capturing such structure in the data (Craddock et al., 2012). As

a practical consequence, the more difficult one expects the clusters to separate (e.g., high-level associative brain regions), the more other clustering algorithms should be preferred.

Hierarchical clustering. In contrast to the above-mentioned partitional algorithms, *hierarchical clustering* (Johnson, 1967) represents an *agglomerative* (i.e., *bottom-up*) approach that may reveal connectional similarities at various coarseness levels (Eickhoff et al., 2011). Here, each individual voxel initially represents a separate cluster. These are then progressively merged into a hierarchy by always combining the two most similar clusters. *Divisive* (i.e., *top-down*) hierarchical clustering operates in the opposite direction (start with one cluster, end with one cluster per voxel/vertex), but is seldom used in neuroimaging CBP. The investigator does not need to specify a cluster number because an organizational hierarchy is generated that allows for a functional-connectional multi-level stratification of the ROI. This introduces the advantage of a nested hierarchical solution of ROI parcellations (i.e., enforced parent-children relationships between clustering solutions with different cluster numbers) at the price of large result quantities. There are, however, at least three important drawbacks associated with hierarchical clustering. First, hierarchical clustering is very sensitive to effects in local neighborhoods, which can have a substantial effect on the higher-level solutions in noisy data such as in neuroimaging. Second, the output evidently depends on the investigator-chosen linkage algorithm, i.e., the rules how clusters are combined. This can be remedied by imposing the additional constraint of merging only spatially neighboring clusters, which tends to be better behaved (Thirion et al., 2014). More specifically, as a both biologically plausible and *greedy* (i.e., exploiting computationally convenient simplification) variant, *spatially constrained hierarchical clustering* merges/divides only immediately neighboring clusters. Unfortunately, different linkage algorithms tend to yield different solutions. Finally, there is a tendency for (non-spatially constrained) hierarchical clustering to yield very imbalanced cluster sizes. In the extreme case, one after one voxel/vertex/node is added to a group containing all other ones. This clustering algorithm should be preferred when expecting many clusters (contrarily to k-means). Depending on the merging heuristic, it can however be quite computationally expensive (e.g., complete clustering). Hierarchical clustering captures well the properties of the connectivity differences (i.e., high accuracy) but its solutions may be inconsistent (i.e., low reproducibility, like k-means). As a rule of thumb, accuracy and reproducibility tend to be mutually exclusive across clustering algorithms.

Distance measures & linkage algorithms. We would like to emphasize that neuroimaging data are typically noisy (due to intersubject variability, technical limits, etc.) and smooth (due to Gaussian filtering). Consequently, standard k-means, spectral, and hierarchical clustering often find spatially contiguous clusters, although this is not immanent in the respective algorithms. While these considerations take place in the *brain space*, it is important not to confuse it with the *feature space*. Distance in brain space pertains to the spatial spacing (in, e.g., mm or voxels) between the units to be clustered, whereas distance in feature space pertains to abstract similarity metrics between connectivity measurements. The process of combining voxels/vertices/nodes in the ROI to connectionally homogeneous clusters (i.e., operating in brain space) is strongly influenced by the employed distance measure and linkage algorithm (i.e., operating in feature space) (Hastie et al., 2011). On the one hand, *distance measures* represent the similarity criterion for pairs of connectivity profiles (cf. Jain et al., 2010; Handl et al., 2005). These include the a) Euclidean distance (i.e., squared difference between respective connectivity values; a special case of Minkowski metric at p=2), b) correlation distance (i.e., Pearson's correlation of the connectivity profile vectors), and c) cosine distance (i.e., one minus the cosine of the included angle between

connectivity profile items, acts as normalization). For cosine distance, subtraction of the coefficient from 1 yields a proper distance metric. It can be advantageous in the presence of outliers. If the connectivity data are known to be particularly noisy. It can be advantageous to use cosine/correlation distances or ranked variants of the above distances (i.e., by using Spearman's rather than Pearson's correlation) to improve resistance to outliers. One the other hand, the *linkage algorithm* guides how the measured distances are used to evolve clusters (cf. Stanberry et al., 2003; Timm 2002). The linkage dictates how voxels are combined to clusters based on the computed distance measures. It can be a) 'weighted' (weighting average distances, defined in various ways in the literature), b) 'average' (not weighting average distances; mean between all connectivity values in a first cluster to all connectivity values in a second cluster), and c) 'ward' (replaces distance measures to the minimization of intracluster variance) as well as d) 'single' (i.e., shortest distance, often produces a skewed solutions, i.e., 'chaining phenomenon by always adding the respective next closest element with heterogeneous overall clusters), and e) 'complete' (maximum distance, tends towards compact clusters, less preferred for noisy data). The best linkage method obviously depends on the data properties. Some combinations of distance measure and linkage seem to be better than others. For instance, when using Euclidean distance the ward linkage seems robust to outliers in noisy data. For different distance/linkage choices, the hierarchical clusters can also find spatially contiguous clusters at a similar rate as k-means.

Alternative clustering procedures. While k-means, spectral, and hierarchical clustering algorithms are used in various parcellation scenarios, ICA and boundary detection serve very similar goals in whole-brain parcellation. ICA is an iterative, non-closed-form solution to blind source separation (Hyvarinen, 1999). Applied to fMRI data, it is known to separate out stable, statistically independent, and possibly overlapping spatial activation patterns. Note that the time courses of the nodes of each extracted brain 'network' are identical (Beckmann et al., 2005; Smith et al., 2009). As a first conceptual point, this makes ICA a viable instance of connectivity-based parcellation of functional brain imaging data. As a second conceptual point, ICA is an instance of soft clustering by allowing solutions of spatially overlapping clusters (contrarily to the three clustering algorithms above). ICA is special in computing generative models of the signal, it may be more noise-sensitive than the above hard clustering algorithms (Smith et al., 2013a), and allows extraction of artifactual patterns from the data (assuming additivity), not possible with the above clustering algorithms or border detection. Such continuous and probabilistic, rather than discrete and binary, clusters also result from different alternative clustering methods in neuroimaging, including multi-subject dictionary learning (Varoquaux et al., 2013), fuzzy C-means clustering (e.g., Cauda et al., 2010), deep neural networks (Bengio, 2009; Plis et al., 2014), and Gaussian mixture models (e.g., Shen et al., 2010). They share the advantage of extracting stratifications of overlapping patterns. This has limited gain in parcellating ROIs that cover one or very few cortical areas but will be particularly relevant in whole-brain CBP. Indeed, the neurobiological justification for CBP is the connectional homogeneity of individual cortical areas. Yet, soft clustering approaches can flexibly represent overlapping neurobiological clusters with more expressive parcellation models (cf. Passingham et al., 2002). Boundary mapping, on the other hand, reconceptualizes clustering as the identification of transitions between territories of adjacent brain areas (Cohen et al., 2008; Wig et al., 2013; Wig et al., 2014). High confidence in boundaries (i.e., high 'edge probability') indicates good cluster separability, and vice versa. Detected boundaries are interpreted as localized, abrupt changes in connectivity profiles. Boundary mapping has been instrumental in segregating both circumscribed brain regions (e.g., frontal cortex, Cohen et al., 2008; lateral parietal cortex, Nelson et al., 2010)

and the entire brain (e.g., Wig et al., 2014). Given the possibility of generating probability boundary maps (e.g., by Canny edge detection algorithm), edge modeling qualifies as a mixture between hard and soft clustering. All clustering procedures mentioned up to this point can be applied in single-subject and group analysis.

More globally, each time the investigator choses a clustering algorithm to be applied on the ROI, he or she accepts a number of implicit or explicit assumptions (Hastie et al., 2011). Therefore, *any clustering algorithm unavoidably biases the resulting clustering solution with respect to the number, shape, relative sizes, hierarchy, and contiguity of the clusters*. Consequently, investigators should resist the temptation to promote their CBP study as 'completely model-free', 'purely data-driven', or 'without any assumptions.' Rather, it is important to realize the inevitable assumptions and biases of a clustering algorithm and motivate the choice of a particular one based on the aim of the study, the ROI, and the employed connectivity data (Handl et al., 2005). Moreover, using different connectional modalities and other imaging modalities, the investigator can provide a valuable cross-confirmation of the clusters' biological relevance. Cross-species evidence in favor of a parcellation solution might be especially important (cf. Ramnani et al., 2006; Sallet et al., 2013; Neubert et al., 2014).

2.4 Statistical inference and cluster number selection

Inferential versus exploratory statistics. In short, assessing the significance of brain parcellation results is hard. This is particularly true if *significant* is employed in the strict sense of inferential statistics and not employed in its broader sense of 'interesting' or 'relevant.' The key problem in wanting to assess statistical significance of CBP results is the requirement of a *null hypothesis* to test against. Conceptually, a ROI clustering solution would hence be deemed *statistically significant* if it has a very low probability of being true under the null hypothesis that the investigator seeks to reject. Yet, such a null hypothesis is often difficult to formulate in clustering applications. Instead of *inferential statistics*, which *test for a particular structure* in the clustering results, investigators need to resort to *exploratory statistics*, which *discover and assess structure* in the data (Efron and Tibshirani, 1991; Tukey, 1962; Hastie et al., 2011). While it is true that statistical methods span a continuum between the two poles of inferential and exploratory statistics, comparing the 'importance' or 'pertinence' of clustering results from a CBP analysis is naturally situated more towards the latter. CBP hence represents an unsupervised statistical learning problem that is conventionally addressed by quantitatively comparing *model fit* using *cluster validity criteria*. It may therefore be seen as one instance of a current shift in neuroimaging away from classical inferential towards exploratory approaches, put differently, from voxel/vertex-level mappings to more global assessment of model fit or predictive power (cf. Brodersen, 2009; Cox and Savoy, 2003; Naselaris et al, 2011).

Cluster validity problem. From a broader perspective, the 'true' shape and number of clusters is unknown for most real-world clustering problems, including brain research. Finding an 'optimal' number of clusters represents an unresolved issue (*cluster validity problem*) in computer science, pattern recognition, and machine learning (Handl et al., 2005; Jain et al., 1999). This has prompted the development of diverse *heuristics* (*cluster validity criteria*) to weigh the quality of the obtained clustering solutions. These are necessary because *clustering algorithms will always find subregions in the investigator's ROI, whether these truly exist in nature or not*.

For Peer Review

Cluster separation criteria. Criteria can be based on the separation between clusters such as the silhouette value (which for each element measures how similar that element is to the other ones in its own cluster, when compared to the nearest clusters) or the inter/intra-cluster distance (which compares the distance between the cluster-centers to the distance between the elements within each cluster). Such criteria reflect the goal of CBP, i.e., to form groups such that voxels/vertices/nodes within a group show similar connectivity, while the connectivity is different between groups. Note however that *successively segregating brain connectivity data into clusters tends to result in lower within-cluster and higher between-cluster differences in ever step, regardless of the applied clustering algorithm and the chosen cluster validity criterion*.

Consistence across parcellations. Criteria can also be based on the consistency across parcellations into a given number of clusters. This set of criteria comprises metrics, including variation of information (VI), the Dice coefficient, normalized (NMI) or adjusted (AMI) mutual information as well as *adjusted Rand index* (ARI). These kinds of criteria are often used in multi-subject, possibly also within-subject, settings. That is, when a given ROI is parcellated separately in each subject based on dMRI or RSFC information (not possible with MACM-CBP). In such studies, assessing the quality of a particular clustering solution by testing the consistency across subjects has emerged as an important standard. Nevertheless, the same concept has also been applied to test consistency across parameters such as filter size or to evaluate the stability based on procedures such as bootstrapping (i.e., summarizing statistical results across analyses of resampled data drawn with replacement from the dataset; Efron & Tibshirani, 1994).

Consistence across cluster numbers. Criteria can be based on the hierarchical consistency, such as the VI across clustering solutions (e.g., how different is the information contained in a 4-cluster as compared to a 3-cluster solution) or the hierarchy index (what is the proportion of voxels/vertices/nodes that are originating from the dominant parent cluster). These metrics are only useful in the context of non-hierarchical clustering algorithms, such as spectral and in particular k-means clustering, as hierarchical clustering inevitably yields a perfect hierarchical consistency.

External knowledge. Criteria reflecting convergence of the different cluster solutions with independent, external information, such as CBP performed in other data modalities or a priori anatomical/functional knowledge, finally, represent a very distinct class of criteria (Handl et al., 2005; Jain et al., 2010). On the one hand, they are, in contrast to the aforementioned metrics, independent from the actual data and hence provide external validity. On the other hand, they are based on the problematic assumption (cf. Amunts et al., 2014) that different sources of information should yield the same parcellation of the brain. This has repeatedly been challenged and in fact can result in a strong confirmatory bias.

==========

Table 1 about here please

==========

In sum, the desire to test the 'statistical significance' of a clustering solution is hard to fulfill (cf. Breiman, 2001; Vogelstein et al., 2014). The wish to assess the 'trueness' of clusters within (i.e., cluster comparison) or between (i.e., model comparison) a cluster number choice may be a more legitimate concern. Choosing the clustering solution with the highest model fit represents an unsupervised statistical learning problem that cannot be easily

For Peer Review

framed within the realms of inferential statistics (Jain, 2010). Rather than trying to test whether a clustering solution reaches *statistical significance,* we propose assessing different *cluster validity criteria* to choose among CBP results. More than one single cluster validity criterion should be used because *the choice of one objective cluster validity criterion is still a subjective choice by the investigator*. We would therefore suggest guiding the choice of a final parcellation by majority vote across a number of complementary cluster validity criteria. These evaluate the more or less good model fit of a given clustering solution in the sense of explaining the data. Even more so, an informed and confident decision on the most pertinent ROI segregation should be justified by consistency across different clustering algorithms and cluster validity criteria (Clos et al., 2013). More generally, the neurobiological 'ground truth', unknown to us neuroscientists, is probably hierarchical, modular, and multi-scale (Frackowiak and Markram, 2015).

Table 1: Main characteristics of cluster validity criteria in brain parcellation

Cluster validity criterion	Rationale	Used in previous brain parcellation studies
Across-subject consistency	finds the number of clusters that yields the highest similarity across independent analyses in a number of subjects; this can be done for instance in a cross-validation framework (e.g., leave-one-subject out) or by a split-group approach	Buckner et al., 2011; Liu et al., 2013; Saygin et al., 2011; Solano-Castiella et al., 2010
Across-hemisphere symmetry	computes the percentage of cluster assignments that agree in both hemispheres; can evidently only be performed in paired brain regions outside of the midline	Bzdok et al., 2012; Kahnt et al., 2012
Adjusted Mutual information (MAI)	assesses the similarity between i) the joint distribution of two sets A and B and ii) the marginal distributions of these two sets; it thus weighs how much information is shared between A and B; results in 0 if A and B are independent, in 1 if they are deterministically related; 'adjusted' implies correction for agreement between clusters out of chance (=RAI); it accounts for the fact that the MI is generally higher for two clustering solutions with a larger number of clusters, regardless of whether there is actually more information shared; related to VI	Thirion et al., 2014

For Peer Review

Adjusted Rand index (RAI)	in analogy to mutual information, a measure from probability theory to assess the statistical dependence between two clustering solutions; Rand index is a measure of accuracy between two clusterings; 'adjusted' implies the corrected-for-chance variant of the Rand index (cf. MAI); RAI can be an order of magnitude faster than AMI	Thirion et al., 2014
Akaike's information criterion (AIC)	derived from information theory, it finds the best number of clusters by acknowledging the trade-off between goodness-of-fit and model complexity (i.e., number of clusters); it is based on a probabilistic model and a measure of complexity; this penalty for the cluster numbers is aimed at preventing overfitting, yet is independent of the sample size; as it is a relative measure, it judges the absolute quality of the finally selected model	Zalesky et al., 2010
Bayesian information criterion (BIC)	despite many similarities to AIC, BIC is motivated by a Bayesian approach to model selection; it penalizes the model complexity (i.e., number of clusters, =AIC); in comparison to AIC, BIC imposes higher costs on more complex models (i.e., small cluster numbers are privileged)	Thirion et al., 2014
Cramér's V	assesses the statistical correlation between two groups of discrete values (i.e., voxel/vertex/node-wise cluster assignments) based on chi-square; put differently, it measures the strength of association between two clustering solutions	Liu et al., 2013; Solano-Castiella et al., 2011
Dice coefficient	assesses the similarity between two samples or adjacency matrices; primarily practically justified rather than backed up theoretically; works well in heterogeneous and outlier-prone data; can be used to compare group and single-subject clusterings; can be computed in different ways; it is equivalent to the *Jaccard index* because there is a monotonic transformation between their scores	Blumensath et al., 2013; Craddock et al., 2012; Shen et al., 2013; Wang et al., 2014
Inter- versus intra-cluster distance ratio	assesses cluster separation by the ratio between the average distance of a voxel/vertex/nodes to its cluster centre and the average distance between the cluster centres; a significantly increased ratio compared to the K-1 solution would indicate a better separation of the obtained clusters	Bzdok et al., 2014; Chang et al., 2009
Percentage of misclassification	assesses cluster assignment by the amount of noise and potentially local effects in the clustering; the average percentage of voxels/vertexs/nodes that were assigned to a different cluster compared to the most frequent assignment of these	Bzdok et al., 2014

For Peer Review

	voxels/vertices/nodes; used to compared ways to compute a same or different number of clusters	
Percentage of parent-children congruency	assesses cluster topology by how many voxels/vertices/nodes are not related to the dominant parent cluster compared to the solution with K - 1 clusters; counts voxels/vertices/nodes that do not reflect a hierarchical organization; related to hierarchy index	Clos et al., 2013; Eickhoff et al., in press; Kahnt et al., 2012
Silhouette coefficient	assesses cluster separation by measuring how similar that voxel/vertex/nodes is to voxels/vertices/nodes in its own cluster compared to voxels/vertices/nodes in the nearest cluster; good solutions are those with a higher silhouette value compared to the K-1 solution; this measure of cluster quality is independent of the number of clusters	Bzdok et al., 2014; Craddock et al., 2012; Eickhoff et al., in press; Kannan et al., 2010; Zhang et al., 2011
Variation of information (VI)	assesses how much knowing the cluster assignment for an item in clustering X reduces the uncertainty about the item's cluster in clustering Y; a linear expression involving mutual information and entropy; is not adjusted for chance (contrary to AMI and ARI); used to compare ways to obtain a same or different number of clusters	Bzdok et al., 2014; Clos et al., 2013; Eickhoff et al., in press; Kahnt et al., 2012; Kelly et al., 2010

See also http://en.wikipedia.org/wiki/Cluster_analysis.

For Peer Review

III Integration

3.1 Claims of circularity

Following ROI parcellation, one of the most frequent questions is what features actually drove this distinction. In other words, 'what is the difference in connectivity between the identified clusters?' Thus, the obtained clusters are frequently submitted to supplementary analyses that usually assess the *same* connectivity modality that was initially used to identify the clusters. This kind of follow-up analysis has repeatedly raised the suspicion of circular analysis or *double dipping*.

Double dipping. In neuroimaging, this often refers to the practice of first correlating a behavioral measure with brain activity and then using a hereby identified subset of voxels/vertices for the second 'actual' correlation analysis with the same behavioral measure (Vul et al., 2008). Such 'spurious correlations' do entail overly enthusiastic results as voxels/vertices have been selected twice for the same purpose in a nested, non-independent fashion (Lieberman et al., 2009). More generally, any statistical analysis has been argued to be invalid when the same data is used for selection and then for discriminative analysis if the test statistic depends on the selection criterion (Kriegeskorte et al., 2009).

Is it an instance of double dipping to use the same (connectivity) data to first identify clusters in a ROI and then compute the connectivity profile of the ensuing clusters? We would argue that it is not. While the clustering step takes place in an exploratory statistical framework, aimed at identifying groups within the ROI voxels/vertices that are similar to each other, the characterization step takes place in an inferential statistics framework, aimed at identifying which target voxels/vertices in the brain have above-chance connections to the clusters (null hypothesis: no part of the brain is connected to the current cluster more than by chance). Put differently, the first step discovers structure in the ROI according to voxel/vertex-wise connectivity, whereas the second step explicitly tests this structure by asking which parts of the brain are significantly connected to the ensuing clusters. As a remark, we can however not test whether two derived clusters have different connectivity. This is because it is connectional differences that led to emergence of these clusters. Put differently, the validity of the statistical test depends on the validity of the underlying null hypothesis. A first way to argue against circularity in this scenario would therefore be that the underlying statistical framework/question of both analyses are markedly different.

Descriptive follow-up analyses. Apart from that, there is a completely different and probably much more pragmatic line of argumentation against accusations of circular analysis. The investigator can explicitly frame the cluster connectivity analysis (i.e., second step) as a *non-inferential* and thus *descriptive* follow-up analysis. This would purely serve to 'illustrate which are the strongest connectivity differences that contributed to the cluster formation.' They hence should be considered nothing more than a visualization of what differences caused the cluster formation. Note that this weakens the conclusions on the cluster connectivity results. Nevertheless, purely descriptive cluster characterizations can provide very useful representations of the ROI at hand (e.g., Thirion et al., 2006a; Smith et al., 2013a).

Even if some investigators might be inclined to reproach the post-hoc characterization of connectivity-derived

clusters with circular analysis, it does not appear to hold for connectivity analyses of individual clusters. On the contrary, a comprehensive multi-modal characterization of the obtained clusters in the ROI is strongly recommended. Using the same connectional modality, connectional patterns that drove the parcellation can actually be made explicit in a non-circular fashion.

3.2 Connectivity-derived clusters as priors

CBP for other analyses. CBP can yield reliable cornerstones for a variety of consecutive neuroimaging analyses. Experimental methods requiring a-priori target regions can capitalize on CBP clusters to further characterize their behavioral implications by diverse viewpoints towards cross-modal functional mapping. This might include but is not exclusive to transcranial magnetic stimulation (TMS), voxel-based morphometry (VBM), structural equation models (SEM), Granger causality mapping (GCM), and seed-based experimental fMRI analyses. From a broader perspective, CBP has the potential to enhance any neuroimaging technique reliant on prospective region definitions that critically hinges on proper fit of the topographical priors.

Create atlases as references. The mentioned methods are not only suited for the creation of spatial maps, but for the characterization of known functional segregation. Connectivity-derived clusters can however be seen as novel, potentially untested hypotheses on regional differentiation. These can be tested in hypothesis-driven experimental studies (i.e., anatomical or functional hypotheses operationalized by region definitions). Asking questions on, for instance, the differences in coactivation pattern of dMRI-defined clusters, the relationship between the regional volumes of the identified regions and phenotypical traits in larger samples or the differential affection of a newly described subdivision in clinical samples are all exciting questions. The information gained from these investigations would then provide a detailed and continuously growing characterization of a brain region within the spatial framework of a particular CBP differentiation. A map of functionally distinct subregions might also serve as an independent reference to explain closely neighboring but topographically diverging activation clusters in task-based experimental neuroimaging studies (e.g., Decety & Lamm, 2007; Cieslik et al., 2013), for the characterization of hemodynamic response profiles (Ciuciu et al., 2003), as well as for functional connectome applications (Smith et al., 2013a; Varoquaux & Craddock, 2013). CBP may thus serve as a *post-hoc analysis* to complement interpretation and as a *preceding analysis* to inform the design of experimental neuroimaging investigations.

CBP may therefore fill a vacuum in the current research landscape in providing new spatial maps of brain regions (i.e., *discovering structure in the brain*). These can then be further characterized by a multi-modal investigation of the (differential) structure, connectivity, and function as well as their relation to various phenotypes in health and disease (i.e., *testing structure in the brain*). Moreover, CBP could become a crucial preliminary step to improve the potency of various seed-based neuroimaging methods.

Conclusion

Connectivity-based parcellation is currently one of the most exciting yet also one of the most fluidly evolving approaches in neuroimaging research. In contrast to most existing methods, it may yield maps of the brain that can be seen as spatial hypotheses on functional or structural segregation – a hypothesis that may and should be

For Peer Review

tested by integrative, multi-modal investigations. We hope, however, that this overview has also raised awareness for the various pitfalls that may be encountered when performing or reviewing CBP analyses; from the initial definition of the ROI (which operationalizes the motivation for that particular investigation and constrains all conclusions that can be drawn), to the choice of the clustering algorithm (with each having its specific strengths and biases), cluster number (which should be based on the examination of multiple metrics and with awareness for multi-level biological organization), and finally the difficulty to apply classical inferential statistics in the context of CBP. Brain parcellation not only serves the generation of new hypotheses. Rather, it might allow new insight into the principles of regional organization when conducting CBP based on different aspects of inter-neuronal communication. Such multi-modal comparison, hand-in-hand with results from hypothesis-testing approaches, might provide unprecedented insight into the organization of regional specialization in human brain's structure, connectivity, and function. Ultimately, connectivity-based parcellation methods offer useful simplified views on brain regions that remain complex in nature.

For Peer Review

Acknowledgments

Thanks to Olivier Grisel for very helpful discussion. We also thank Rogier Mars and Gesa Hartwigsen for their thoughtful comments on a previous version of this manuscript. We are indebted to two anonymous reviewers for their insightful comments.

For Peer Review

Figure Captions

Figure 1

Different functional segregation models of the insula

From upper to lower panel: a) depicts a concept of functional segregation in the insula as inferred from connectivity studies by axonal tracing in monkeys (Mesulam & Mufson, 1982b), b) two-cluster solution of the maturing insula from connectivity-based parcellation of newborn and two-year-old humans based on longitudinal resting-state measurements (Alcauter et al., 2013), c) three-cluster solution of the insula of human adults showns subdivision of the anterior insula into ventral and dorsal subregions based on resting-state correlations (Deen et al., 2011), and d) different consensus-cluster solutions of the right and left insula in human adults based on the three connectivity modalities resting-state correlations, task-derived coactivations, and structural covariance (Kelly et al., 2012).

Box 1

Synopsis of clustering algorithms used for CBP

K-means clustering

Clustering depends on free parameters, including i) the cluster initialization, ii) the cluster number k, and iii) the tolerance for iteration stopping. Initially k voxels/vertices in the ROI are randomly chosen to represent the centers of the k desired clusters. Two steps are then iterated multiple times. First, the ROI voxels/vertices/nodes are assigned to the closest cluster center (i.e., 'centroid'), which equates with partitioning the ROI into k clusters. Second, the k cluster centers are recomputed. As soon as the center needs to be shifted by less than the preset distance threshold (early stopping is often needed for the sake of time), the iterative process stops. Note that the final assignments of ROI voxels/vertices/nodes to particular clusters may vary with different cluster initializations and yield non-optimal solutions at local minima. The same connectivity data may thus result in several stable solutions for the ROI parcellation at the *same* preset k (i.e., low reproducibility), as shown, for instance, in the human insula (Nanetti et al., 2009). Consequently, the algorithm is usually repeated many times.

Spectral clustering

It is based on a similarity matrix quantifying the similarity of the connectivity profiles between any pair of voxels/vertices/nodes within the ROI. The first eigenvectors of the (normalized) Laplacian of the similarity matrix are computed. Those then enable transformation into an alternative data representation in a space with reduced dimensionality as the eigenvectors 'summarize' features of the similarity matrix. The output of this *data transformation* can then be used for either i) *spectral reordering* or ii) an ordinary clustering algorithm. Spectral reordering uses the reduced similarity information to reorder the similarity matrix in such manner that voxels/vertices/nodes that are similar to each other are grouped together (Barnard et al., 1995). This can be useful to semi-quantitatively obtain a possible number of clusters by eye inspection (e.g., Johansen-Berg et al., 2004; Bzdok et al., 2012). Note that hierarchical consistency across solutions for different cluster numbers is not methodologically enforced. That is, voxels/vertices/nodes may be assigned to a different cluster when looking for instance at the clustering solutions with 3 or 4 clusters.

Hierarchical clustering

Each individual voxel/vertex/node initially represents a separate cluster. These are then progressively merged into a hierarchy by a) always combining the two most similar clusters (i.e., bottom-up) or b) always dividing the least homogenous cluster (i.e., top-down) in every step. The algorithm implicitly walks through different choices of cluster numbers as these approaches generate a hierarchy that allows for a nested multi-level parcellation of

For Peer Review

the ROI, even if only few of those are interpreted in practice. Such successive clustering trees introduce the advantage of a nested hierarchical solution of ROI parcellations (i.e., enforced parent-children relationships between clustering solutions with different cluster numbers) at the price of large result quantities. Even if this clustering model reflects current views on the hierarchical organization of the brain, a given hierarchical clustering result is not necessarily neurobiologically valid.

For Peer Review

References

Abraham, A., Pedregosa, F., Eickenberg, M., Gervais, P., Mueller, A., Kossaifi, J., Gramfort, A., Thirion, B., Varoquaux, G., 2014. Machine learning for neuroimaging with scikit-learn. Front Neuroinform 8, 14.

Alcauter, S., Lin, W., Keith Smith, J., Gilmore, J.H., Gao, W., 2013. Consistent Anterior-Posterior Segregation of the Insula During the First 2 Years of Life. Cereb Cortex.

Allman, J.M., Watson, K.K., Tetreault, N.A., Hakeem, A.Y., 2005. Intuition and autism: a possible role for Von Economo neurons. Trends Cogn Sci 9, 367-373.

Amunts, K., Kedo, O., Kindler, M., Pieperhoff, P., Mohlberg, H., Shah, N.J., Habel, U., Schneider, F., Zilles, K., 2005. Cytoarchitectonic mapping of the human amygdala, hippocampal region and entorhinal cortex: intersubject variability and probability maps. Anat Embryol (Berl) 210, 343-352.

Amunts, K., Lepage, C., Borgeat, L., Mohlberg, H., Dickscheid, T., Rousseau, M.E., Bludau, S., Bazin, P.L., Lewis, L.B., Oros-Peusquens, A.M., Shah, N.J., Lippert, T., Zilles, K., Evans, A.C., 2013. BigBrain: an ultrahigh-resolution 3D human brain model. Science 340, 1472-1475.

Amunts, K., Schleicher, A., Burgel, U., Mohlberg, H., Uylings, H.B., Zilles, K., 1999. Broca's region revisited: cytoarchitecture and intersubject variability. J Comp Neurol 412, 319-341.

Andrews-Hanna JR, Reidler JS, Sepulcre J, Poulin R, Buckner RL. 2010. Functional-anatomic fractionation of the brain's default network. Neuron 65:550-562.

Anwander, A., Tittgemeyer, M., von Cramon, D.Y., Friederici, A.D., Knosche, T.R., 2007. Connectivity-Based Parcellation of Broca's Area. Cereb Cortex 17, 816-825.

Ashburner, J., Kloppel, S., 2011. Multivariate models of inter-subject anatomical variability. Neuroimage 56, 422-439.

Augustine, J.R., 1996. Circuitry and functional aspects of the insular lobe in primates including humans. Brain Res Rev 22, 229-244.

Bach, D.R., Behrens, T.E., Garrido, L., Weiskopf, N., Dolan, R.J., 2011. Deep and superficial amygdala nuclei projections revealed in vivo by probabilistic tractography. J Neurosci 31, 618-623.

Barnard, S.T., Pothen, A., Simon, H.D., 1995. Numer. Linear Algebra Appl. 2, 317-334.

Beckmann, C.F., DeLuca, M., Devlin, J.T., Smith, S.M., 2005. Investigations into resting-state connectivity using independent component analysis. Philos Trans R Soc Lond B Biol Sci 360, 1001-1013.

Beckmann, M., Johansen-Berg, H., Rushworth, M.F., 2009. Connectivity-based parcellation of human cingulate cortex and its relation to functional specialization. J Neurosci 29, 1175-1190.

Behrens, T.E., Johansen-Berg, H., 2005. Relating connectional architecture to grey matter function using diffusion imaging. Phil. Trans. R. Soc. B 360, 903–911.

Behrens, T.E., Johansen-Berg, H., Woolrich, M.W., Smith, S.M., Wheeler-Kingshott, C.A., Boulby, P.A., Barker, G.J., Sillery, E.L., Sheehan, K., Ciccarelli, O., Thompson, A.J., Brady, J.M., Matthews, P.M., 2003. Non-invasive mapping of connections between human thalamus and cortex using diffusion imaging. Nat Neurosci 6, 750-757.

Bengio, Y., 2009. Learning deep architectures for AI. Foundations and trends in Machine Learning 2, 1-127.

Bellec, P., Rosa-Neto, P., Lyttelton, O.C., Benali, H., Evans, A.C., 2010. Multi-level bootstrap analysis of stable

For Peer Review

clusters in resting-state fMRI. Neuroimage 51, 1126-1139.

Biswal, B., Yetkin, F.Z., Haughton, V.M., Hyde, J.S., 1995. Functional connectivity in the motor cortex of resting human brain using echo-planar MRI. Magn Reson Med 34, 537-541.

Biswal, B.B., Mennes, M., Zuo, X.N., Gohel, S., Kelly, C., Smith, S.M., Beckmann, C.F., Adelstein, J.S., Buckner, R.L., Colcombe, S., Dogonowski, A.M., Ernst, M., Fair, D., Hampson, M., Hoptman, M.J., Hyde, J.S., Kiviniemi, V.J., Kotter, R., Li, S.J., Lin, C.P., Lowe, M.J., Mackay, C., Madden, D.J., Madsen, K.H., Margulies, D.S., Mayberg, H.S., McMahon, K., Monk, C.S., Mostofsky, S.H., Nagel, B.J., Pekar, J.J., Peltier, S.J., Petersen, S.E., Riedl, V., Rombouts, S.A., Rypma, B., Schlaggar, B.L., Schmidt, S., Seidler, R.D., Siegle, G.J., Sorg, C., Teng, G.J., Veijola, J., Villringer, A., Walter, M., Wang, L., Weng, X.C., Whitfield-Gabrieli, S., Williamson, P., Windischberger, C., Zang, Y.F., Zhang, H.Y., Castellanos, F.X., Milham, M.P., 2010. Toward discovery science of human brain function. Proc Natl Acad Sci U S A 107, 4734-4739.

Blumensath, T., Jbabdi, S., Glasser, M.F., Van Essen, D.C., Ugurbil, K., Behrens, T.E., Smith, S.M., 2013. Spatially constrained hierarchical parcellation of the brain with resting-state fMRI. Neuroimage 76, 313-324.

Breiman, L., 2001. Statistical Modeling: The Two Cultures. Statistical Science 16, 199-231.

Brodersen, K.H., 2009. Decoding mental activity from neuroimaging data — the science behind mind-reading. The New Collection, Oxford 4, 50-61.

Brodmann, K., 1909. Vergleichende Lokalisationslehre der Großhirnrinde. Barth, Leipzig.

Buckner, R.L., Krienen, F.M., Yeo, B.T., 2013. Opportunities and limitations of intrinsic functional connectivity MRI. Nat Neurosci 16, 832-837.

Bzdok, D., Laird, A., Zilles, K., Fox, P.T., Eickhoff, S.B., 2012. An investigation of the structural, connectional and functional sub-specialization in the human amygdala. Hum Brain Mapp 34, 3247-3266.

Bzdok, D., Langner, R., Schilbach, L., Engemann, D., Laird, A.R., Fox, P.T., Eickhoff, S.B., 2013a. Segregation of the human medial prefrontal cortex in social cognition. Front Hum Neurosci.

Bzdok, D., Langner, R., Schilbach, L., Jakobs, O., Roski, C., Caspers, S., Laird, A., Fox, P.T., Zilles, K., Eickhoff, S.B., 2013b. Characterization of the temporo-parietal junction by combining data-driven parcellation, complementary connectivity analyses, and functional decoding. Neuroimage 81, 381-392.

Bzdok, D., Eickhoff, S.B., 2015. The resting-state physiology of the human cerebral cortex. In: Toga, A.W. (Ed.), Brain Mapping: An Encyclopedic Reference. Elsevier.

Bzdok, D., Heeger, A., Langner, R., Laird, A., Fox, P., Palomero-Gallagher, N., Vogt, B., Zilles, K., Eickhoff, S., 2014. Subspecialization in the human posterior medial cortex. Neuroimage 106, 55-71.

Cauda, F., Cavanna, A.E., D'Agata, F., Sacco, K., Duca, S., Geminiani, G.C., 2011. Functional connectivity and coactivation of the nucleus accumbens: a combined functional connectivity and structure-based meta-analysis. J Cogn Neurosci 23, 2864-2877.

Cauda, F., Geminiani, G., D'Agata, F., Sacco, K., Duca, S., Bagshaw, A.P., Cavanna, A.E., 2010. Functional connectivity of the posteromedial cortex. PLoS One 5.

Chang, L.J., Yarkoni, T., Khaw, M.W., Sanfey, A.G., 2013. Decoding the role of the insula in human cognition: functional parcellation and large-scale reverse inference. Cereb Cortex 23, 739-749.

Chumbley, J.R., Friston, K.J., 2009. False discovery rate revisited: FDR and topological inference using

For Peer Review

Gaussian random fields. Neuroimage 44, 62-70.

Cieslik, E.C., Zilles, K., Caspers, S., Roski, C., Kellermann, T.S., Jakobs, O., Langner, R., Laird, A.R., Fox, P.T., Eickhoff, S.B., 2013. Is there "one" DLPFC in cognitive action control? Evidence for heterogeneity from co-activation-based parcellation. Cereb Cortex 23, 2677-2689.Clos, M., Amunts, K., Laird, A.R., Fox, P.T., Eickhoff, S.B., 2013. Tackling the multifunctional nature of Broca's region meta-analytically: Co-activation-based parcellation of area 44. Neuroimage 83C, 174-188.

Ciuciu, P., Poline, J.B., Marrelec, G., Idier, J., Pallier, C., Benali, H., 2003. Unsupervised robust non-parametric estimation of the hemodynamic response function for any fMRI experiment. IEEE Trans Med Imaging 22, 1235-1251.

Chumbley, J.R., Friston, K.J., 2009. False discovery rate revisited: FDR and topological inference using Gaussian random fields. Neuroimage 44, 62-70.

Cohen, A.L., Fair, D.A., Dosenbach, N.U., Miezin, F.M., Dierker, D., Van Essen, D.C., Schlaggar, B.L., Petersen, S.E., 2008. Defining functional areas in individual human brains using resting functional connectivity MRI. NeuroImage 41, 45–57.

Corbetta, M., Patel, G., Shulman, G.L., 2008. The reorienting system of the human brain: from environment to theory of mind. Neuron 58, 306-324.

Cox, D.D., Savoy, R.L., 2003. Functional magnetic resonance imaging (fMRI)"brain reading": detecting and classifying distributed patterns of fMRI activity in human visual cortex. Neuroimage 19, 261-270.

Craddock, R.C., James, G.A., Holtzheimer, P.E., 3rd, Hu, X.P., Mayberg, H.S., 2012. A whole brain fMRI atlas generated via spatially constrained spectral clustering. Hum Brain Mapp 33, 1914-1928.

Craig, A.D., 2003. Interoception: the sense of the physiological condition of the body. Curr Opin Neurobiol 13, 500-505.

Craig, A.D., 2005. Forebrain emotional asymmetry: a neuroanatomical basis? Trends Cogn Sci 9, 566-571.

Craig, A.D., 2009. How do you feel - now? The anterior insula and human awareness. Nat Rev Neurosci 10, 59-70.

Damoiseaux, J.S., Rombouts, S.A., Barkhof, F., Scheltens, P., Stam, C.J., Smith, S.M., Beckmann, C.F., 2006. Consistent resting-state networks across healthy subjects. Proc Natl Acad Sci U S A 103, 13848-13853.

Decety, J., Lamm, C., 2007. The role of the right temporoparietal junction in social interaction: how low-level computational processes contribute to meta-cognition. Neuroscientist 13, 580-593.

Dehaene, S., Piazza, M., Pinel, P., Cohen, L., 2003. Three parietal circuits for number processing. Cogn Neuropsychol 20, 487-506.

Devlin JT, Poldrack RA. 2007. In praise of tedious anatomy, NeuroImage. 37, 1033-1041.

Donath, W.E., Hoffman, A.J., 1973. Lower bounds for the partitioning of graphs. IBM J. Res. Develop. 17, 420-425.

Efron, B., Tibshirani, R., 1991. Statistical data analysis in the computer age. Science 253, 390-395.

Efron, B., & Tibshirani, R. J. (1994). An introduction to the bootstrap. CRC press.

Eickhoff, S.B., Bzdok, D., Laird, A.R., Roski, C., Caspers, S., Zilles, K., Fox, P.T., 2011. Co-activation patterns distinguish cortical modules, their connectivity and functional differentiation. Neuroimage 57, 938-949.

For Peer Review

Eickhoff, S.B., Jbabdi, S., Caspers, S., Laird, A.R., Fox, P.T., Zilles, K., Behrens, T.E., 2010. Anatomical and functional connectivity of cytoarchitectonic areas within the human parietal operculum. J Neurosci 30, 6409-6421.

Eickhoff, S.B., Laird, A.R., Grefkes, C., Wang, L.E., Zilles, K., Fox, P.T., 2009. Coordinate-Based Activation Likelihood Estimation Meta-Analysis of Neuroimaging Data: A Random-Effects Approach Based on Empirical Estimates of Spatial Uncertainty. Hum Brain Mapp. 30, 2907-2926.

Eickhoff, S.B., Laird, A.R., Fox, P.T., Bzdok, D., Hensel, L., in press. Functional segregation of the human dorsomedial prefrontal cortex. Cereb Cortex.

Evans, A.C., Collins, D.L., Milner, B., 1992. An MRI-based stereotactic atlas from 250 young normal subjects. Soc Neurosci Abstr 18.

Evans, A.C., 2013. Networks of anatomical covariance. Neuroimage 80, 489-504.

Fischl, B., van der Kouwe, A., Destrieux, C., Halgren, E., Segonne, F., Salat, D.H., Busa, E., Seidman, L.J., Goldstein, J., Kennedy, D., Caviness, V., Makris, N., Rosen, B., Dale, A.M., 2004. Automatically parcellating the human cerebral cortex. Cereb Cortex 14, 11-22.

Felleman, D.J., Van Essen, D.C., 1991. Distributed hierarchical processing in the primate cerebral cortex. Cereb Cortex 1, 1-47.

Forgy, E.W., 1965. Cluster analysis of multivariate data: efficiency versus interpretability of classifications. Biometrics 21, 768-769.

Formisano, E., Kriegeskorte, N., 2012. Seeing patterns through the hemodynamic veil--the future of pattern-information fMRI. Neuroimage 62, 1249-1256.

Fox, P.T., Friston, K.J., 2012. Distributed processing; distributed functions? Neuroimage 61, 407-426.

Fox, P.T., Lancaster, J.L., 2002. Opinion: Mapping context and content: the BrainMap model. Nat Rev Neurosci 3, 319-321.

Fox, C.J., Iaria, G., Barton, J.J., 2009. Defining the face processing network: optimization of the functional localizer in fMRI. Hum Brain Mapp 30, 1637-1651.

Frackowiak R. and Markram H., 2015. The future of human cerebral cartography: a novel approach. Phil. Trans. R. Soc. B 370.

Friston, K.J., Frith, C.D., Fletcher, P., Liddle, P.F., Frackowiak, R.S., 1996. Functional topography: multidimensional scaling and functional connectivity in the brain. Cereb Cortex 6, 156-164.

Friston, K.J., Rotshtein, P., Geng, J.J., Sterzer, P., Henson, R.N., 2006. A critique of functional localisers. Neuroimage 30, 1077-1087.

Gordon, E.M., Laumann, T.O., Adeyemo, B., Huckins, J.F., Kelley, W.M., Petersen, S.E., 2014. Generation and Evaluation of a Cortical Area Parcellation from Resting-State Correlations. Cereb Cortex.

Gorgolewski, K.J., Varoquaux, G., Rivera, G., Schwarz, Y., Ghosh, S.S., Maumet, C., Nichols, T.E., Poldrack, R.A., Poline, J.P., Yarkoni, Y., Margulies, D.S., 2014. NeuroVault.org: A web-based repository for collecting and sharing unthresholded statistical maps of the human brain, in press.

Grosbras, M.H., Paus, T., 2002. Transcranial magnetic stimulation of the human frontal eye field: effects on visual perception and attention. J. Cogn. Neurosci. 14, 1109-1120.

For Peer Review

Handl, J., Knowles, J., Kell, D.B., 2005. Computational cluster validation in post-genomic data analysis. Bioinformatics 21, 3201-3212.

Hartigan, J.A., Wong, M.A., 1979. A k-means clustering algorithm. Appl. Stat. 28, 100–108.

Hastie, T., Tibshirani, R., Friedman, J., 2011. The Elements of Statistical Learning. Springer Series in Statistics, Heidelberg, Germany.

Hensel, L., Bzdok, D., Müller, V.I., Zilles, K., Eickhoff, S.B., in press. Neural correlates of explicit social judgments on vocal stimuli. Cereb Cortex.fH

Honey, C.J., Sporns, O., Cammoun, L., Gigandet, X., Thiran, J.P., Meuli, R., Hagmann, P., 2009. Predicting human resting-state functional connectivity from structural connectivity. Proc Natl Acad Sci U S A 106, 2035-2040.

Igelström, K. M., Webb, T. W., & Graziano, M. S., 2015. Neural Processes in the Human Temporoparietal Cortex Separated by Localized Independent Component Analysis.
The Journal of Neuroscience 35, 9432-9445.

Jbabdi, S., Behrens, T.E., 2013. Long-range connectomics. Ann N Y Acad Sci 1305, 83-93.

Jain, A.K., 2010. Data clustering: 50 years beyond K-means. Pattern Recognition Letters 31, 651–666.

Jain, A.K., Murty, M.N., Flynn, P.J., 1999. Data clustering: a review. ACN Computing Surveys 31, 264-323.

Johansen-Berg, H., Behrens, T.E., Robson, M.D., Drobnjak, I., Rushworth, M.F., Brady, J.M., Smith, S.M., Higham, D.J., Matthews, P.M., 2004. Changes in connectivity profiles define functionally distinct regions in human medial frontal cortex. Proc Natl Acad Sci U S A 101, 13335-13340.

Aapo Hyvarinen. Fast and robust fixed-point algorithms for independent component analysis. Neural Networks, IEEE Transactions on, 10(3):626–634, 1999.

Johansen-Berg, H., Behrens, T.E., Sillery, E., Ciccarelli, O., Thompson, A.J., Smith, S.M., Matthews, P.M., 2005. Functional-anatomical validation and individual variation of diffusion tractography-based segmentation of the human thalamus. Cereb Cortex 15, 31-39.

Johansen-Berg, H., Rushworth, M.F., 2009. Using diffusion imaging to study human connectional anatomy. Annu Rev Neurosci 32, 75-94.

Johnson, S.C., 1967. Hierarchical clustering schemes. Psychometrika 32, 241–254.

Jones, D.K., 2008. Studying connections in the living human brain with diffusion MRI. Cortex 44, 936-952.

Kahnt, T., Chang, L.J., Park, S.Q., Heinzle, J., Haynes, J.D., 2012. Connectivity-based parcellation of the human orbitofrontal cortex. J Neurosci 32, 6240-6250.

Kanwisher N, McDermott J, Chun MM. 1997. The fusiform area: A module in human extrastriate cortex specialized for face perception. The Journal of Neuroscience 17, 4301-4311.

Kelly, C., Toro, R., Di Martino, A., Cox, C.L., Bellec, P., Castellanos, F.X., Milham, M.P., 2012. A Convergent Functional Architecture of the Insula Emerges Across Imaging Modalities. Neuroimage 61, 1129–1142.

Kim, J.H., Lee, J.M., Jo, H.J., Kim, S.H., Lee, J.H., Kim, S.T., Seo, S.W., Cox, R.W., Na, D.L., Kim, S.I., Saad, Z.S., 2010. Defining functional SMA and pre-SMA subregions in human MFC using resting state fMRI: functional connectivity-based parcellation method. Neuroimage 49, 2375-2386.

Kleinberg, J., 2002. An impossibility theorem for clustering. NIPS 15, 463-470.

Kochunov, P., Glahn, D.C., Fox, P.T., Lancaster, J.L., Saleem, K., Shelledy, W., Zilles, K., Thompson, P.M.,

For Peer Review

Coulon, O., Mangin, J.F., Blangero, J., Rogers, J., 2010. Genetics of primary cerebral gyrification: Heritability of length, depth and area of primary sulci in an extended pedigree of Papio baboons. Neuroimage 53, 1126-1134.

Koster-Hale J, Saxe R, Dungan J, Young LL. 2013. Decoding moral judgments from neural representations of intentions. Proceedings of the National Academy of Sciences of the United States of America 110:5648-5653.

Krall SC, Rottschy C, Oberwelland E, Bzdok D, Fox PT, Eickhoff SB, Fink GR, Konrad K. 2015. The role of the right temporoparietal junction in attention and social interaction as revealed by ALE meta-analysis. Brain Struct Funct 220:587-604.

Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S., Baker, C.I., 2009. Circular analysis in systems neuroscience: the dangers of double dipping. Nat Neurosci 12, 535-540.

Kurth, F., Zilles, K., Fox, P.T., Laird, A.R., Eickhoff, S.B., 2010. A link between the systems: functional differentiation and integration within the human insula revealed by meta-analysis. Brain Struct Funct 214, 519-534.

Laird, A.R., Lancaster, J.L., Fox, P.T., 2009. Lost in localization? The focus is meta-analysis. Neuroimage 48, 18-20.

Lee, K., Tak, S., Ye, J.C., 2011. A Data-Driven Sparse GLM for fMRI Analysis Using Sparse Dictionary Learning With MDL Criterion. Medical Imaging, IEEE Transactions, 1076-1089.

Lieberman, M.D., Berkman, E.T., Wager, T.D., 2009. Correlations in Social Neuroscience Aren't Voodoo: Commentary on Vul et al. Perspectives on Psychological Science 4.

Lloyd, S., 1957, published 1982 in IEEE Transactions on Information Theory. Least squares quantization in PCM. Technical report 28, 128-137.

Logothetis, N.K., Pauls, J., Augath, M., Trinath, T., Oeltermann, A., 2001. Neurophysiological investigation of the basis of the fMRI signal. Nature 412, 150-157.

Logothetis, N.K., Wandell, B.A., 2004. Interpreting the BOLD signal. Annu Rev Physiol 66, 735-769.

Marrelec, G., Krainik, A., Duffau, H., Pelegrini-Issac, M., Lehericy, S., Doyon, J., Benali, H., 2006. Partial correlation for functional brain interactivity investigation in functional MRI. Neuroimage 32, 228-237.

Mars, R.B., Jbabdi, S., Sallet, J., O'Reilly, J.X., Croxson, P.L., Olivier, E., Noonan, M.P., Bergmann, C., Mitchell, A.S., Baxter, M.G., Behrens, T.E., Johansen-Berg, H., Tomassini, V., Miller, K.L., Rushworth, M.F., 2011. Diffusion-weighted imaging tractography-based parcellation of the human parietal cortex and comparison with human and macaque resting-state functional connectivity. J Neurosci 31, 4087-4100.

Mars, R.B., Sallet, J., Schuffelgen, U., Jbabdi, S., Toni, I., Rushworth, M.F., 2012. Connectivity-Based Subdivisions of the Human Right "Temporoparietal Junction Area": Evidence for Different Areas Participating in Different Cortical Networks. Cereb Cortex 22, 1894-1903.

Mennes, M., Kelly, C., Colcombe, S., Castellanos, F.X., Milham, M.P., 2013. The Extrinsic and Intrinsic Functional Architectures of the Human Brain are not equivalent. Cereb Cortex 23, 223-229.

Mesulam, M.M., 1978. Tetramethyl benzidine for horseradish peroxidase neurohistochemistry: a non-carcinogenic blue reaction product with superior sensitivity for visualizing neural afferents and efferents. J Histochem Cytochem 26, 106-117.

Mesulam, M.M., Mufson, E.J., 1982a. Insula of the old world monkey. I. Architectonics in the insulo-orbito-

For Peer Review

temporal component of the paralimbic brain. J Comp Neurol 212, 1-22.

Mesulam, M.M., Mufson, E.J., 1982b. Insula of the old world monkey. III: Efferent cortical output and comments on function. J Comp Neurol 212, 38-52.

Mesulam, M.M., 1998. From sensation to cognition. Brain 121 (Pt 6), 1013-1052.

Mitchell, J.P., 2008. Activity in right temporo-parietal junction is not selective for theory-of-mind. Cereb Cortex 18, 262-271.

Mufson, E.J., Mesulam, M.M., 1982. Insula of the old world monkey. II: Afferent cortical input and comments on the claustrum. J Comp Neurol 212, 23-37.

Muhle-Karbe, P.S., Derrfuss, J., Lynn, M.T., Neubert, F.X., Fox, P.T., Brass, M., Eickhoff, S.B., 2015. Co-Activation-Based Parcellation of the Lateral Prefrontal Cortex Delineates the Inferior Frontal Junction Area. Cereb Cortex.

Mutihac, R., Braun, A., Balkin, T.J., 2005. Exploratory versus Confirmatory Analysis in Imaging Neuroscience. Recent Advances in Intelligent Control, Modelling and Simulation.

Nanetti, L., Cerliani, L., Gazzola, V., Renken, R., Keysers, C., 2009. Group analyses of connectivity-based cortical parcellation using repeated k-means clustering. Neuroimage 47, 1666-1677.

Naselaris, T., Kay, K.N., Nishimoto, S., Gallant, J.L., 2011. Encoding and decoding in fMRI. Neuroimage 56, 400-410.

Neubert, F.X., Mars, R.B., Thomas, A.G., Sallet, J., Rushworth, M.F., 2014. Comparison of human ventral frontal cortex areas for cognitive control and language with areas in monkey frontal cortex. Neuron 81, 700-713.

Ng, B., Varoquaux, G., Poline, J., Thirion, B., 2013. Implications of Inconsistencies between fMRI and dMRI on Multimodal Connectivity Estimation MICCAI.

O'Toole, A.J., Jiang, F., Abdi, H., Penard, N., Dunlop, J.P., Parent, M.A., 2007. Theoretical, statistical, and practical perspectives on pattern-based classification approaches to the analysis of functional neuroimaging data. J Cogn Neurosci 19, 1735-1752.

Passingham, R.E., Stephan, K.E., Kotter, R., 2002. The anatomical basis of functional localization in the cortex. Nat Rev Neurosci 3, 606-616.

Petrides, M., Tomaiuolo, F., Yeterian, E.H., Pandya, D.N., 2012. The prefrontal cortex: comparative architectonic organization in the human and the macaque monkey brains. Cortex 48, 46-57.

Poldrack, R.A., Gorgolewski, K.J., 2014. Making big data open: data sharing in neuroimaging. Nat Neurosci 17, 1510-1517.

Plis, S.M., Hjelm, D.R., Salakhutdinov, R., Allen, E.A., Bockholt, H.J., Long, J.D., Johnson, H.J., Paulsen, J.S., Turner, J.A., Calhoun, V.D., 2014. Deep learning for neuroimaging: a validation study. Front Neurosci 8, 229.

Pradal, C., Varoquaux, G., 2013. Publishing scientific software matters. J. Computational Science.

Radua, J., Mataix-Cols, D., 2009. Voxel-wise meta-analysis of grey matter changes in obsessive-compulsive disorder. Br J Psychiatry 195, 393-402.

Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., Matelli, M., 1988. Functional organization of inferior area 6 in the macaque monkey. II. Area F5 and the control of distal movements. Exp Brain Res 71, 491-507.

For Peer Review

Robinson, J.L., Laird, A.R., Glahn, D.C., Lovallo, W.R., Fox, P.T., 2010. Metaanalytic connectivity modeling: delineating the functional connectivity of the human amygdala. Hum Brain Mapp 31, 173-184.

Sallet, J., Mars, R.B., Noonan, M.P., Neubert, F.X., Jbabdi, S., O'Reilly, J.X., Filippini, N., Thomas, A.G., Rushworth, M.F., 2013. The organization of dorsal frontal cortex in humans and macaques. J Neurosci 33, 12255-12274.

Saxe, R., Brett, M., Kanwisher, N., 2006. Divide and conquer: a defense of functional localizers. Neuroimage 30, 1088-1096; discussion 1097-1089.

Saygin, Z.M., Osher, D.E., Koldewyn, K., Reynolds, G., Gabrieli, J.D., Saxe, R.R., 2012. Anatomical connectivity patterns predict face selectivity in the fusiform gyrus. Nat Neurosci 15, 321-327.

Scannell, J.W., Blakemore, C., Young, M.P., 1995. Analysis of connectivity in the cat cerebral cortex. J Neurosci 15, 1463-1483.

Schilbach, L., Eickhoff, S.B., Rotarska-Jagiela, A., Fink, G.R., Vogeley, K., 2008. Minds at rest? Social cognition as the default mode of cognizing and its putative relationship to the "default system" of the brain. Conscious Cogn 17, 457-467.

Schurz M, Radua J, Aichhorn M, Richlan F, Perner J. 2014. Fractionating theory of mind: a meta-analysis of functional brain imaging studies. Neuroscience and biobehavioral reviews 42:9-34.

Shehzad, Z., Kelly, A.M., Reiss, P.T., Gee, D.G., Gotimer, K., Uddin, L.Q., Lee, S.H., Margulies, D.S., Roy, A.K., Biswal, B.B., Petkova, E., Castellanos, F.X., Milham, M.P., 2009. The resting brain: unconstrained yet reliable. Cereb Cortex 19, 2209-2229.

Shen, X., Papademetris, X., Constable, R.T., 2010. Graph-theory based parcellation of functional subunits in the brain from resting-state fMRI data. Neuroimage 50, 1027-1035.

Shen, X., Tokoglu, F., Papademetris, X., Constable, R.T., 2013. Groupwise whole-brain parcellation from resting-state fMRI data for network node identification. Neuroimage 82, 403-415.

Silani, G., Lamm, C., Ruff, C.C., Singer, T., 2013. Right supramarginal gyrus is crucial to overcome emotional egocentricity bias in social judgments. The Journal of Neuroscience 33:15466-15476.

Smith, S.M., 2012. The future of FMRI connectivity. Neuroimage.

Smith, S.M., Fox, P.T., Miller, K.L., Glahn, D.C., Fox, P.M., Mackay, C.E., Filippini, N., Watkins, K.E., Toro, R., Laird, A.R., Beckmann, C.F., 2009. Correspondence of the brain's functional architecture during activation and rest. Proc Natl Acad Sci U S A 106, 13040-13045.

Smith, S.M., Miller, K.L., Salimi-Khorshidi, G., Webster, M., Beckmann, C.F., Nichols, T.E., Ramsey, J.D., Woolrich, M.W., 2011. Network modelling methods for FMRI. Neuroimage 54, 875-891.

Smith, S.M., Beckmann, C.F., Andersson, J., Auerbach, E.J., Bijsterbosch, J., Douaud, G., Duff, E., Feinberg, D.A., Griffanti, L., Harms, M.P., Kelly, M., Laumann, T., Miller, K.L., Moeller, S., Petersen, S., Power, J., Salimi-Khorshidi, G., Snyder, A.Z., Vu, A.T., Woolrich, M.W., Xu, J., Yacoub, E., Ugurbil, K., Van Essen, D.C., Glasser, M.F., Consortium, W.U.-M.H., 2013a. Resting-state fMRI in the Human Connectome Project. Neuroimage 80, 144-168.

Smith, S. M., Vidaurre, D., Beckmann, C. F., Glasser, M. F., Jenkinson, M., Miller, K. L., and Van Essen, D. C., 2013b. Functional connectomics from resting-state fMRI. Trends in cognitive sciences, 17, 666-682.

Solano-Castiella, E., Anwander, A., Lohmann, G., Weiss, M., Docherty, C., Geyer, S., Reimer, E., Friederici,

For Peer Review

A.D., Turner, R., 2010. Diffusion tensor imaging segments the human amygdala in vivo. Neuroimage 49, 2958-2965.

Stanberry L, Nandy R, Cordes D., 2003. Cluster analysis of fMRI data using dendrogram sharpening. Human brain mapping 20:201-219.

Talairach, J., Tournoux, P., 1988. Co-planar stereotaxic atlas of the human brain. Thieme, New York.

Thirion, B., Flandin, G., Pinel, P., Roche, A., Ciuciu, P., Poline, J.B., 2006b. Dealing with the shortcomings of spatial normalization: multi-subject parcellation of fMRI datasets. Hum Brain Mapp 27, 678-693.

Thirion, B., Varoquaux, G., Dohmatob, E., Poline, J.B., 2014. Which fMRI clustering gives good brain parcellations? Front Neurosci 8, 167.

Tibshirani, R., Walther, G., Hastie, T., 2001. Estimating the number of clusters in a data set via the gap statistic. J. Roy. Statist. Soc. B, 411–423.

Timm, N.H. 2002. Applied Multivariate Analysis. Springer, New York

Toro, R., Fox, P.T., Paus, T., 2008. Functional coactivation map of the human brain. Cereb Cortex 18, 2553-2559.

Tzourio-Mazoyer, N., Landeau, B., Papathanassiou, D., Crivello, F., Etard, O., Delcroix, N., Mazoyer, B., Joliot, M., 2002. Automated anatomical labeling of activations in SPM using a macroscopic anatomical parcellation of the MNI MRI single-subject brain. Neuroimage 15, 273-289.

Tukey, J.W., 1962. The future of data analysis. Annals of Statistics 33, 1-67.

Tukey, J.W., 1977. Exploratory Data Analysis. Addison-Wesley.

Van Essen, D.C., Anderson, C.H., Felleman, D.J., 1992. Information processing in the primate visual system: an integrated systems perspective. Science 255, 419-423.

Van Essen, D.C., 1985. Functional organization of primate visual cortex. In: Peters, A., Jones, E.G. (Eds.), Cereb Cortex. Plenum Press, New York, pp. 259-329.

Varoquaux, G., & Craddock, R. C. (2013) Learning and comparing functional connectomes across subjects. *Neuroimage* 80: 405-415.

Varoquaux G, Gramfort A, Pedregosa F, Michel V, Thirion B. 2011. Multi-subject dictionary learning to segment an atlas of brain spontaneous activity. Information processing in medical imaging : proceedings of the conference 22:562-573.

Vogt, C., Vogt, O., 1919. Allgemeinere Ergebnisse unserer Hirnforschung. Journal für Psychologie und Neurologie 25, 279-461.

Vogelstein, J.T., Park, Y., Ohyama, T., Kerr, R.A., Truman, J.W., Priebe, C.E., Zlatic, M., 2014. Discovery of brainwide neural-behavioral maps via multiscale unsupervised structure learning. Science 344, 386-392.

von Economo, C., 1926. Eine neue Art Spezialzellen des Lobus cinguli und Lobus insulae. Z Ges Neurol Psychiatr 100, 706-712.

von Luxburg, U., 2007. A tutorial on spectral clustering. Statistics and Computing. 17, 395–416.

Vul, E., Harris, C., Winkielman, P., Pashler, H., 2008. Voodoo Correlations in Social Neuroscience. Psychol Sci.

Wager, T.D., Lindquist, M., Kaplan, L., 2007. Meta-analysis of functional neuroimaging data: current and future directions. Soc Cogn Affect Neurosci 2, 150-158.

For Peer Review

Wandell, B. A., Dumoulin, S. O., & Brewer, A. A. (2007). Visual field maps in human cortex. Neuron 56, 366-383.

Wang, J., Yang, Y., Fan, L., Xu, J., Li, C., Liu, Y., Fox, P.T., Eickhoff, S.B., Yu, C., Jiang, T., 2014. Convergent functional architecture of the superior parietal lobule unraveled with multimodal neuroimaging approaches. Hum Brain Mapp.

Weiner, K.S., Golarai, G., Caspers, J., Chuapoco, M.R., Mohlberg, H., Zilles, K., Amunts, K., Grill-Spector, K., 2014. The mid-fusiform sulcus: a landmark identifying both cytoarchitectonic and functional divisions of human ventral temporal cortex. Neuroimage 84, 453-465.

Wiegell, M.R., Tuch, D.S., Larsson, H.B., Wedeen, V.J., 2003. Automatic segmentation of thalamic nuclei from diffusion tensor magnetic resonance imaging. Neuroimage 19, 391-401.

Wig, G.S., Laumann, T.O., Cohen, A.L., Power, J.D., Nelson, S.M., Glasser, M.F., Miezin, F.M., Snyder, A.Z., Schlaggar, B.L., Petersen, S.E., 2013. Parcellating an individual subject's cor- tical and subcortical brain structures using snowball sampling of resting-state correla- tions. Cereb. Cortex.

Wig, G.S., Laumann, T.O., Petersen, S.E., 2014. An approach for parcellating human cortical areas using resting-state correlations. Neuroimage 93, 276-291.

Wolpert, D., 1996. The lack of a priori distinctions between learning algorithms. Neural Computation 8, 1341–1390.

Yarkoni, T., Poldrack, R.A., Nichols, T.E., Van Essen, D.C., Wager, T.D., 2011. Large-scale automated synthesis of human functional neuroimaging data. Nat Methods 8, 665-670.

Yeo, B.T., Krienen, F.M., Sepulcre, J., Sabuncu, M.R., Lashkari, D., Hollinshead, M., Roffman, J.L., Smoller, J.W., Zollei, L., Polimeni, J.R., Fischl, B., Liu, H., Buckner, R.L., 2011. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. J Neurophysiol 106, 1125-1165.

Young, M.P., 1993. The organization of neural systems in the primate cerebral cortex. Proc Biol Sci 252, 13-18.

Young, M.P., Hilgetag, C.C., Scannell, J.W., 2000. On imputing function to structure from the behavioural effects of brain lesions. Philos Trans R Soc Lond B Biol Sci 355, 147-161.

Zalesky, A., Fornito, A., Harding, I.H., Cocchi, L., Yucel, M., Pantelis, C., Bullmore, E.T., 2010. Whole-brain anatomical networks: does the choice of nodes matter? Neuroimage 50, 970-983.

Zhang, D., Raichle, M.E., 2010. Disease and the brain's dark energy. Nat Rev Neurol 6, 15-28.

Zhang, S., Li, C., 2012. Functional connectivity mapping of the human precuneus by resting state fMRI. Neuroimage 59, 3548-3562.

Zhang, Y., Fan, L., Zhang, Y., Wang, J., Zhu, M., Zhang, Y., Yu, C., Jiang, T., 2014. Connectivity-based parcellation of the human posteromedial cortex. Cereb Cortex 24, 719-727.

Zilles, K., Amunts, K., 2010. Centenary of Brodmann's map--conception and fate. Nat Rev Neurosci 11, 139-145.

Zilles, K., Schleicher, A., Langemann, C., Amunts, K., Morosan, P., Palomero-Gallagher, N., Schormann, T., Mohlberg, H., Burgel, U., Steinmetz, H., Schlaug, G., Roland, P.E., 1997. Quantitative analysis of sulci in the human cerebral cortex: development, regional heterogeneity, gender difference, asymmetry, intersubject variability and cortical architecture. Hum Brain Mapp 5, 218-221.

For Peer Review