

# The brain of the beholder: honouring individual representational idiosyncrasies

Ian Charest<sup>a,\*</sup>, Nikolaus Kriegeskorte<sup>a</sup>

<sup>a</sup>Medical Research Council Cognition and Brain Sciences Unit, Cambridge, UK

## Abstract

In the early days of neuroimaging, brain function was investigated by averaging across voxels within a region, stimuli within a category and individuals within a group. These three forms of averaging discard important neuroscientific information. Recent studies have explored analyses that combine the evidence in better-motivated ways. Multivariate pattern analyses enable researchers to reveal representations in distributed population codes, honouring the unique information contributed by different voxels (or neurons). Condition-rich designs more richly sample the stimulus space and can treat each stimulus as a unique entity. Finally, each individual's brain is unique and recent studies have found ways to model and analyse the interindividual representational variability. Here we review our field's journey towards more sophisticated analyses that honour these important idiosyncrasies of brain representations. We describe an emerging framework for investigating individually unique pattern representations of particular stimuli in the brain. The framework models stimuli, responses and individuals multivariately and relates representations by means of representational dissimilarity matrices. Important components are computational models and multivariate descriptions of brain and behavioural responses. These recent developments promise a new paradigm for studying the individually unique brain at unprecedented levels of representational detail.

**Keywords:** object representations, individual differences, fMRI, representational similarity analyses

We would like to understand “the brain”. However, every brain is different. Our unique brains, products of our genes and individual experience, give rise to our unique personalities. Even the same person's brain is constantly in flux, with its plasticity adapting the individual to a changing environment [1, 2, 3]. Given the idiosyncratic and plastic nature of any individual brain, it is amazing that brain science has been quite successful with a research paradigm that assumes that all brains are identical. At the gross scale of the global layout of functional regions, brain imaging studies have documented a significant degree of consistency across individual brains in both the functional decomposition and the approximate localisation of functional components. However, we know that functional correspondence between individual brains must break down at some level [4]. As far as we know, no neuron in a higher level cortical region in one person's brain has an exact functional equivalent in another person's brain.

In this paper, we argue that neuroscience needs to honour the uniqueness of the individual brain, the unique contribution of each patch of an individual cortex (and ultimately each neuron) to brain representations and processing and the particular properties and meaning of each particular stimulus (Figure 1). Most studies in neuroimaging have averaged across individuals, across cortical columns (within functional regions) and across particular stimuli. Reducing the complexity by averaging has provided a natural starting point for our investigations. In the domain of vision, early functional magnetic resonance imaging

(fMRI) reported several cortical regions that selectively respond to particular categories of images. For example [5], contrasted brain activations elicited by faces and other objects and analysed the fMRI signal averaged across the voxels within each region, across stimuli within each category and across subjects (in a random-effects analysis). They observed that the activity within a region in the fusiform gyrus was stronger for face images compared to images depicting objects from other categories. Regions selective for other categories, including places [6] and bodies [7], have also been discovered with this approach.

The approach of analysing three-way average activation has served the purposes of revealing the big picture, focusing analyses on what is consistent across individuals, simplifying the inference and increasing power by combining the evidence. However, averaging across individuals, cortical patches and stimuli also discards a lot of information in the brain-activity data that will ultimately be essential to understanding brain function.

An emerging literature is beginning to honour the idiosyncrasies that are at the heart of how the individual brain gives rise to the individual mind, endowing a complex world with a unique meaning and producing successful behaviour. We sketch a framework for multivariate analyses that link stimulus, brain representations and behaviour without averaging across stimuli, brain locations or individuals. In the context of this framework, we review previous studies in object vision that have made forays along these dimensions of progress. Finally, we describe what elements are still missing for cognitive neuroscience to engage the challenges of individuality.

\*Corresponding author

Email address: [ian.charest@mrc-cbu.cam.ac.uk](mailto:ian.charest@mrc-cbu.cam.ac.uk) (Ian Charest)

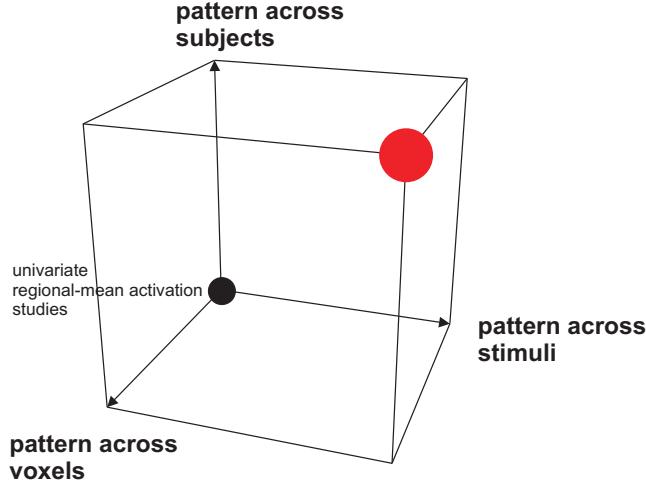


Figure 1: Honouring the unique response of each individual brain, in each cortical patch or column to each particular stimulus. Early neuroimaging studies have focused on analyses of regional-mean activation, averaging across voxels within the region of interest (e.g. a face region), across stimuli within a given category (e.g. individual faces) and across individual subjects (black circle). The field has begun to honour these distinctions, by analysing patterns of activity within each region, responses to single particular images and the variation across individuals. However, only recently have all these components of progress been combined (red circle) in a single study.

### Representations: linking neurons to the computational goals of the brain

The concept of representation is central to the brain and cognitive sciences. Researchers often refer to a characteristic pattern of neuronal activity that reliably occurs when a particular stimulus is presented as a 'representation' of the stimulus. When we refer to an activity pattern as a representation, we go beyond a statement of the statistical dependency between stimulus and response pattern established by data analysis. The term 'representation' implies a functional interpretation, namely that the brain-activity pattern in question serves the function of representing the stimulus in the context of the organism's overall brain information processing [8, 9]. This functional interpretation, although questionable in each particular case, has been extraordinarily helpful in building theories of brain function.

David Marr famously proposed the pursuit of brain science at three levels of description [10, 11]. The highest level is that of the computational goals of the system. The intermediate level is that of representations and algorithm. And the lowest level is that of neuronal implementation. This framework continues to guide theoretical and empirical brain science [12]. The representational interpretation, thus, provides the link between neuronal activity and the computational goals of the brain.

Cognitive psychology has tested cognitive theories with behavioural data, linking theory to experiment at a high level of description. At the opposite end of the spectrum, computational neuroscience has tested single neuron computational models with activity recorded from single neurons, linking theory and experiment at Marr's lowest level. Both approaches are limited: behavioural data do not provide sufficient empirical constraints to explain brain function, and single-neuron computational models will never explain complex cognitive processes. Neuronal representations reside at an intermediate level that

promises to link neurons to cognition. Neuronal representations are commonly associated with the activity of populations of neurons within a functional area. They exist at a spatial scale that lies between the level of single neurons, whose activity is classically recorded with electrodes, and the regional average activation of functional regions, which has been characterised by classical brain imaging methods. A major current challenge is to link theory and experiment at this crucial intermediate level of description, the level of neuronal population representations (Figure 2).

In the domain of visual perception, the recognition of an object takes place along a hierarchy of visual areas [13], whose representational content ranges from local image features to representations of the parts of natural objects and their relationships, and on to semantic properties, as information moves forward along the ventral stream [14, 15, 16]. The cascade of representations arises through feedforward, lateral recurrent and feedback signalling between interconnected regions. Higher cognitive processes make use of the representations to produce successful behaviour and construct or update the organism's knowledge [17]. Understanding the representation within each area would provide a major stepping stone towards understanding the brain as a whole.

### From univariate selectivity to pattern information

Brain science has experienced a paradigm shift from univariate analyses of selectivity to multivariate analyses of pattern information. This paradigm shift started with the theoretical concept of a neuronal "population code" and has more recently led to widespread multichannel measurement and multivariate pattern-information analyses (Figure 3) in both cell recording and functional imaging [18].

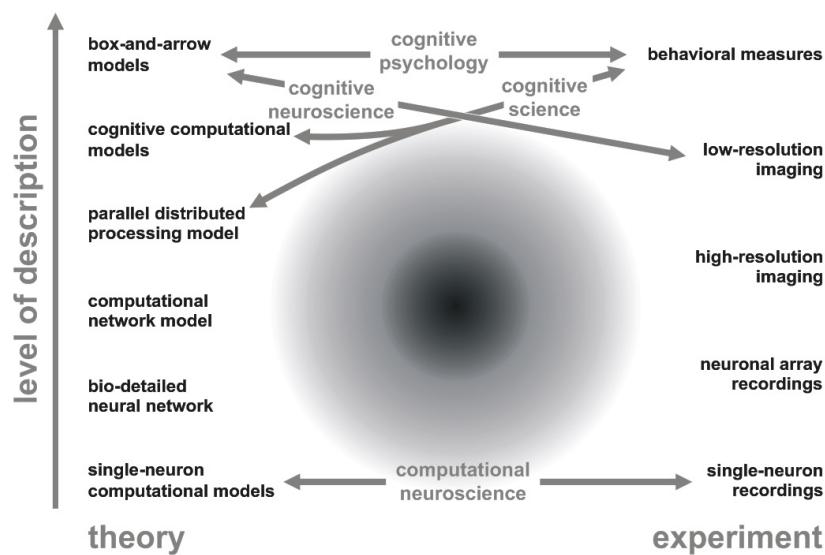


Figure 2: Levels of description in theory and experiment. The explanatory gap between single neurons (bottom) and cognition is bridged on the side of theory (left) by models at multiple levels of description. On the experimental side (right), the gap is bridged by neuronal array recordings and high-resolution imaging. However, the cognitive and brain sciences have yet to learn how to connect theory and experiment at the crucial intermediate level of description. The study of representational geometry offers one avenue of addressing this challenge.

In neuroimaging, the paradigm shift towards pattern information analyses was pioneered by Haxby et al. [19], whose pattern-decoding approach revealed widely distributed category information in the ventral temporal cortex. This study was among the first to honour the distinct contribution to the representation of each little patch of cortex measured by an fMRI voxel. However, like previous work on category selectivity, response patterns were averaged across many distinct stimuli (which were presented in a category-block design), and although activity patterns were not averaged across individuals, individual idiosyncrasies were not investigated.

The approach described in Haxby et al. [19] was an example of the important concept of late combination of the evidence [20]. Combining the evidence across multiple measurements is essential when dealing with noisy data, as it improves the signal-to-noise ratio. The combination of the evidence can be achieved early on in the analysis procedure, by averaging activation levels across voxels, stimuli or individuals. However, averaging is not the only way to combine the evidence. Haxby et al. (2001) did not average across voxels within a region. Instead, they correlated category-related response patterns across voxels as part of their pattern-decoding approach. Pattern-information analyses, such as multivariate decoding, combine the evidence across voxels without averaging the activity levels. This exploits a major strength of fMRI, the large number of response channels (voxels), and can greatly enhance sensitivity [21, 4]. The case for late combination of the evidence has been articulated for evidence distributed across voxels [19, 4] across stimuli [20, 22] and across individuals [23, 19].

### From category averages to rich sets of particular stimuli

Following Haxby et al. [19], multivoxel decoding became popular in neuroimaging [24, 21, 25]. Most studies took a simple-decoding approach, asking, for example, whether regional activity patterns contain information about a particular stimulus dichotomy. While engaging the complexity of distributed representations, the literature largely ignored individual differences and seldom analysed the representation of individual stimuli. Kriegeskorte and Bandettini [4] investigated the pattern representations of particular stimuli. However, in order to obtain stable estimates of the response patterns, the study was limited to four particular object images, two faces and two houses.

A number of studies have explored more complex stimulus spaces, while averaging across voxels and focusing on commonalities across subjects [26, 27, 28]. Mur et al. [26] investigated whether the category preferences of regions in the visual ventral stream held for every exemplar of a set of 96 object images. They found that face and place regions exhibit almost perfectly categorical ranking of the single-image activations, but also graded responses within the preferred and non-preferred categories. This study also took a step towards honouring individually unique representations by investigating both subject-average and subject-unique activation profiles.

Three fMRI studies published in 2008 explored the pattern representations of richer sets of particular stimuli. Mitchell et al. [29] investigated the representation of noun concepts. They showed that a semantic model fitted on the basis of response patterns elicited by 58 word picture pairs could predict response patterns elicited by novel noun concepts (not used in training the model). Kay et al. [30] took a similar approach in

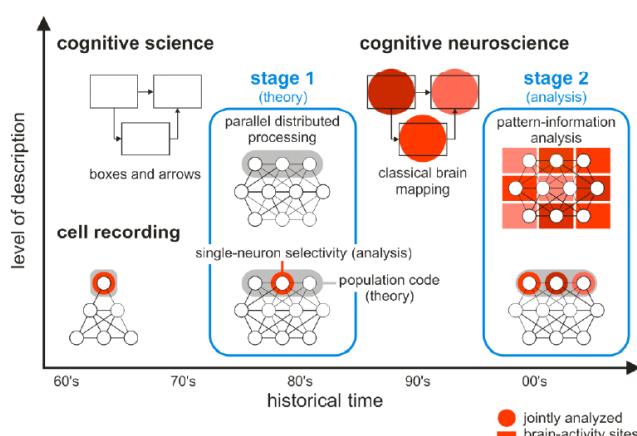


Figure 3: Population-coding theory and pattern-information analysis: a two-stage paradigm shift for systems neuroscience. The paradigm shift from univariate selectivity to pattern information in neuronal population codes had two historical stages. In the 1960s, cell recording studies mostly used single electrodes to measure single-neuron (or single-site) response amplitudes. The responses were analysed univariately to assess neuronal selectivity. In theory, similarly, the dominant idea was that information important to the animal should be explicitly represented in single-cell responses. Cognitive science proceeded quite separately from neurophysiology in this period, constructing box-and-arrow models based on behavioural data. Stage 1 of the paradigm shift concerned theory and occurred when population coding became the prevalent concept in neurophysiology. Cognitive science concurrently embraced the concept of parallel-distributed processing, which brought it closer to brain science. Although distributed population coding was the dominant theoretical concept, measurement and analysis of neuronal responses (red) remained mostly univariate. Then the advent of neuroimaging led to the inception of cognitive neuroscience. However, spatial resolution was initially low (in the cm range with positron emission tomography) and so the initial approach was to link brain regions to the boxes of the box-and-arrow models of cognitive psychology. Researchers assessed each region's overall involvement in different tasks with univariate analyses. More recently, cell-array recordings and high-resolution functional imaging have enabled us to measure distributed representations in unprecedented detail. Stage 2 of the paradigm shift concerned analysis and occurred when the field began to measure large numbers of responses simultaneously within a region and to analyse them jointly with multivariate pattern-analysis techniques.

the domain of vision, investigating the representation of particular images in early visual cortex. They showed that a Gabor-filter model fitted on the basis of response patterns elicited by 1750 training images could predict response patterns elicited by novel images.

Kriegeskorte and Bandettini [4] investigated the representation of 92 object images in the inferior temporal cortex of humans (hIT) and monkeys (IT). They found that the patterns associated with individual images formed clusters corresponding to natural categories and that these clusters (along with the within-category representational dissimilarities) matched closely between human and monkey. These three studies, reviewed in [31], all honoured the distinct contributions of individual voxels and the representations of particular stimuli. By exploiting the late combination of evidence, they managed to forgo averaging across stimuli, while exploring richer sets of stimuli than previous studies.

## From pattern information to representational geometry and tests of computational models

The three papers just mentioned [30, 22, 29] also took the analysis of representational patterns in a novel direction. While previous pattern analyses used generic statistical models to demonstrate the presence of information about the stimuli in a brain region, these three studies tested computational models of brain information processing, which predicted not merely the presence of information about particular stimulus dimensions, but the format in which the information was represented. In population receptive-field modelling [32, 30, 29], computational models are used to predict the responses of individual voxels. In representational similarity analysis [20, 22] computational models predict the dissimilarity relationships of the response patterns (Figure 4 and Figure 5). The fMRI activity pattern elicited by each particular stimulus is compared to the pattern elicited by each other stimulus. All pairwise comparisons are assembled in a representational dissimilarity matrix (RDM). RDMs are useful because they capture not only the information present, but also the format in which it is represented. Moreover, RDMs from brain regions can be directly compared to RDMs predicted by computational models. Population receptive-field modelling, by contrast, requires a separate data-set for fitting a linear model that predicts the response of each voxel from the computational model representation. RDMs also enable straightforward comparisons between brain regions, between individuals, between species and between modalities of brain-activity measurement [9].

## From stable to task-flexible representations

The representational geometry of a brain region has recently been shown to be somewhat influenced by top-down mechanisms such as attention and behavioural goals [33, 34, 35, 36]. Çukur et al. [33] used fMRI to study the impact that searching for an object category during a natural movie has on the semantic representation measured in the brain. The volunteers were asked to either “search for humans” or “search for vehicles” while their brain activities were recorded. Similarly to the procedure described in Huth et al. [37], the authors used WordNet to label object and action categories in natural movies. Using voxel-wise modelling and regularised regression, the authors showed how attending one category distorted the semantic structure of the neural representations of both attended and unattended categories, with category-attended expansion of the representational geometry at the cost of a compression of the distant category [33]. In another study, Harel et al. [34] used fMRI to investigate how the neural representations in regions of the ventral temporal cortex vary as a function of task. They compared the brain-activity patterns for a single set of objects under six different tasks (fixation, colour, tilt, content, movement and size). Their results demonstrated the presence of flexible task-dependent neural representations in the lateral occipital and in the lateral prefrontal cortex, indicating that object processing is highly influenced by the aim of the observer [34]. These two studies clearly demonstrated the task-adaptive

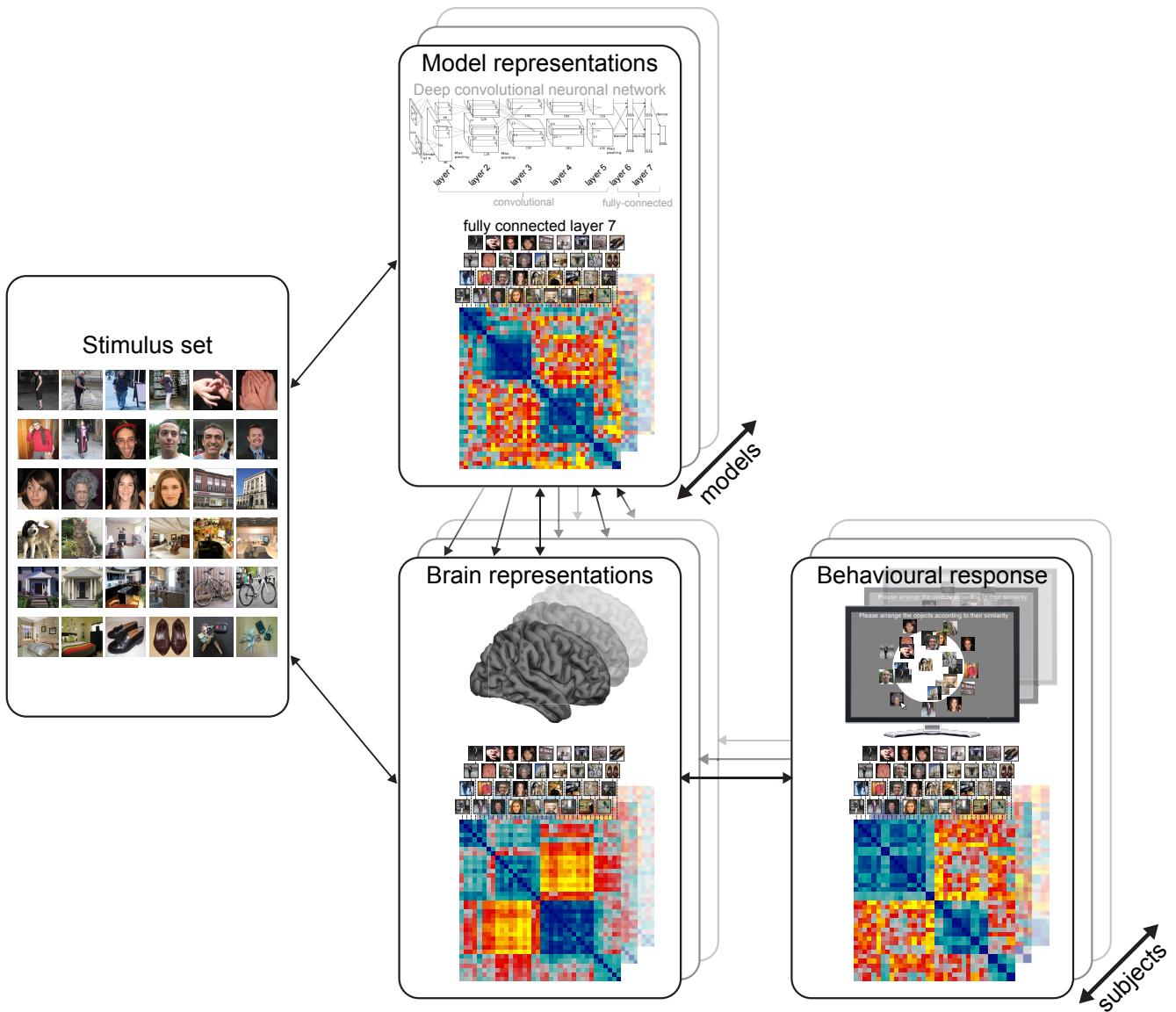


Figure 4: Combining multivariate descriptions of stimuli, brain representations, and behavioural responses. With recent advances in pattern information analyses, progress has been made in understanding object-vision processes in the brain, and to relate them to behaviour. Advances in computer vision enable us to model the similarity structure of stimulus properties with ecologically valid, neuropsychology inspired and biologically plausible computational models. One example of such computer vision model is the deep convolutional neuronal network (DCNN) model of vision, which made significant advances in object recognition and offer similar representational performance to human inferior temporal (hIT) cortex. For example, there is a great deal of correspondence in representational geometry between the fully connected layer 7 of the DCNN and the representational geometry in individual subjects hIT cortex. Future research will seek to establish whether understanding the computations achieved throughout the model's architecture can help us understand the computational mechanisms that enable an individual's brain to recognise objects. These brain representations can also be compared across subjects, and related to multidimensional accounts of subjective behaviour.

flexibility of the representations of the ventral temporal cortex and of the prefrontal cortex. Both these studies were designed to characterise commonalities across individuals in the brain's ability to flexibly adapt to task demands.

### From group characterisations to individually unique representations

Honouring individual voxels and stimuli, the cited studies have made important advances. However, most of this literature tacitly assumed that representations are consistent across individuals [38, 39, 24, 37, 30, 40, 22, 29, 41, 42, 43, 44, 45, 46]. Several studies have explicitly demonstrated representational commonalities between individual [19, 47] and even between species [22].

Even if all individuals had functionally identical brains and the spatial layout of functional units were similar across individuals, the precise anatomical location of functional units in a common brain space might still be variable. This would reflect limitations of our alignment methods (defining the common brain space) rather than true functional differences between individuals. Brains are commonly aligned using volume-based (Talairach or MNI) or surface-based (FreeSurfer, BrainVoyager, AFNI-SUMA) methods or individual functional localiser experiments.

Considering the variability of functional localisation across individuals (as it appears in a given common brain space) can improve the outcome of activation analyses. For example, Fedorenko and colleagues have stressed the importance of defining functional regions in a subject-specific manner, before averaging activation across subjects [48, 49]. Neuroimaging has a long history of individually defined functional regions of interest (e.g. Kanwisher et al., 1997). In pattern-information studies, similarly, activity patterns are usually not averaged across individuals [19, 21]. A recent study demonstrated that structure-function relationships can be consistent between individuals, even when the precise localisation in a given common brain space is variable. Saygin et al. [50] demonstrated that the precise location of face-selective responses in the fusiform gyrus of an individual subject can be predicted on the basis of anatomical connectivity measured with diffusion-weighted imaging. Face-selective responses in the fusiform gyrus were associated with a particular fingerprint of anatomical connectivity across the rest of the brain. The prediction model was cross-validated across subjects, demonstrating the consistency across individuals of the relationship between structure (anatomical connectivity) and function [50].

The most widely used paradigm of group analysis in neuroimaging assumes some common brain space and treats across-subject variation in activation levels as a random effect. In contrast to fixed-effects analysis, where the variability across subjects is not modelled at all, random-effects analysis models interindividual variation as noise [51, 52, 53]. It is important to note that neither of these approaches treats the interindividual variation as an effect of interest.

Despite the widespread focus on commonalities, there is also an expanding neuroimaging literature on individual differences

(for a review, see [54]), with most studies addressing individual differences in regional-average activation. Individual differences in regional-average activation are commonly investigated in memory research, higher order cognitive functions, intelligence research and social neuroscience [55, 56, 57, 58].

As an example from the object-vision literature, Furl et al. [55] showed that regional-average face responses predict a person's ability to recognise faces. Other studies have investigated how interindividual variability in connectivity [59] and brain anatomical measures relates to behavioural differences between individuals [60, 61, 62].

Most studies that investigated individual differences in brain function focused on regional-average activation and its relationship to behaviour. While pattern-information studies usually allow for the precise location of representational units to vary across individuals, few studies to date have addressed hypotheses about the individuality of brain representations. Can brain representational idiosyncrasies explain a person's unique perception and behaviour?

Raizada et al. [63] investigated interindividual differences in brain representations and the degree to which they predict interindividual differences in perception. Using pattern-information analyses, they compared brain representations of the phonemes /ra/ and /la/ in English and Japanese speakers, showing that the behavioural ability to discriminate these vocalisations could be predicted from the discriminability of their representational patterns in auditory cortex [63, 64]. This relationship held across groups (English vs. Japanese speakers) and even across individuals within the two groups.

Another study investigated the role of the hippocampus in episodic memory using pattern-information analyses [65]. The authors showed that the activity patterns observed in the hippocampus carried information about the temporal positions of objects in learned sequences. Furthermore, individuals who performed better in object sequence retrieval had more robust hippocampal object-position binding as indicated by larger hippocampal voxel pattern similarity. The studies by Raizada et al. [63] and Hsieh et al. [65] used pattern representations measured in each subject to predict a unidimensional subject covariate (behavioural performance). A recent study by Charest et al. [23] investigated the relationship between high-dimensional perceptual judgements of the similarity among a set of objects and the multivariate brain representations of the objects. Earlier work has shown that object similarity judgements exhibit similar categorical divisions and similar within-category similarity structure as ventral temporal representational patterns [66]. Charest et al. [23] showed that even individual idiosyncrasies of the perceptual judgements could be predicted from idiosyncrasies of the brain representational geometries. The methodological approach and key result of this paper are shown in Figure 5.

### From univariate to multivariate descriptions of stimuli and behavioural responses

We have argued that brain representations are inherently multivariate and should thus be described and analysed multivari-

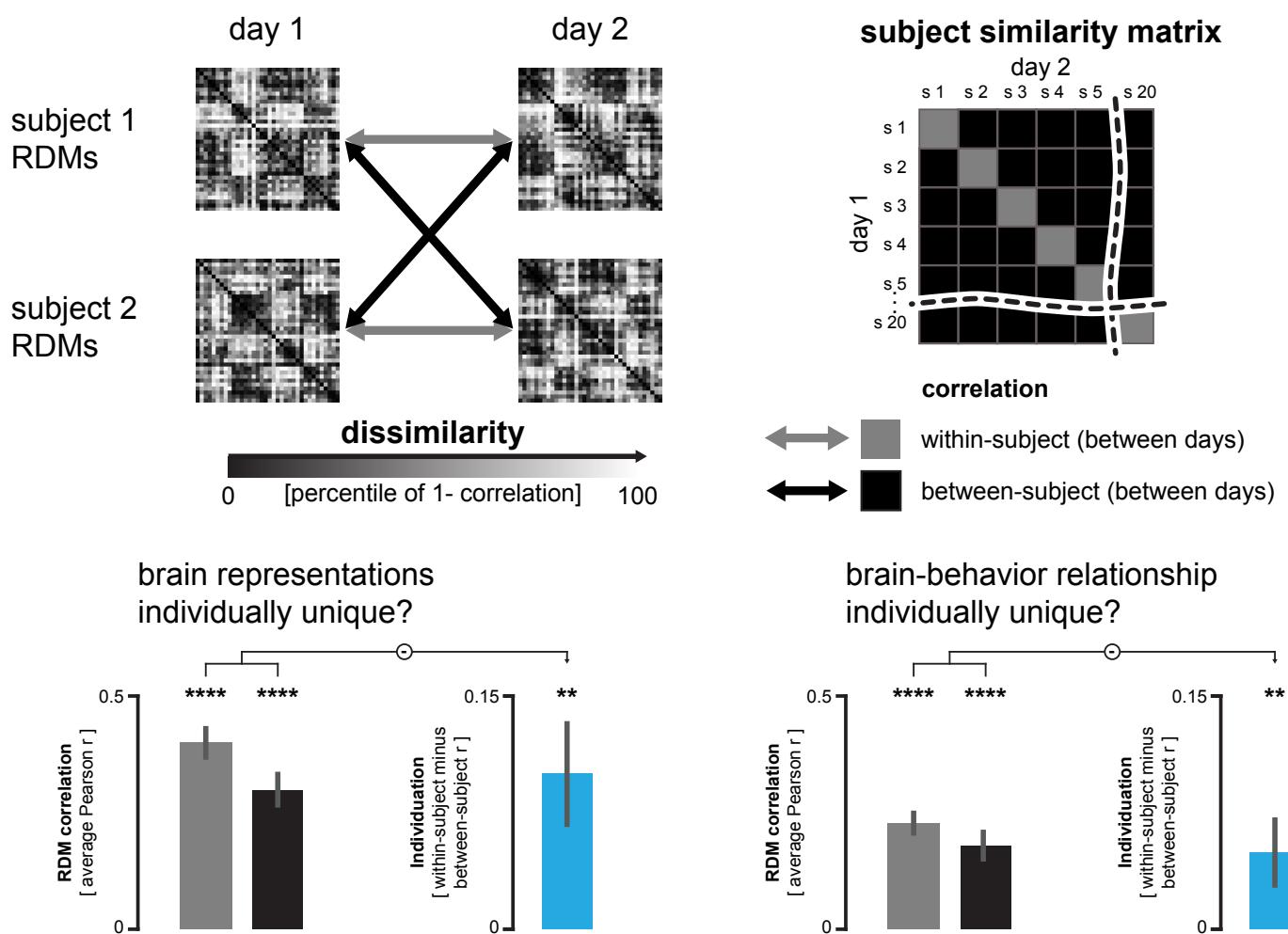


Figure 5: Characterising individually unique object representations. (a) Individual subjects are scanned on two separate days. This allows characterising the stability of a representational geometry within and between subjects (red and black bars). This also enables us to compute an individuation index, consisting of the average within-subject RDM correlation across days, minus the average between-subject RDM correlation across days. If the within-subject RDM correlation is significantly larger than the between-subject RDM correlation, this indicates some degree of individual uniqueness in the representation. After defining whether components of the brain representations are individually unique, one can also investigate whether the brain representations are predictive of the idiosyncrasies in behaviour using the above mentioned framework. (b) Actual results from a recent study on individually unique brain representations. The left panel shows the stability of the representational geometry of the hIT activity patterns to visual objects (within-and between-subjects; red and black bar). The within-subject RDM correlation was significantly larger than the between-subject RDM correlation across days, reflecting the idiosyncrasies in the hIT representation (blue bar). The right panel shows a significant correlation between the brain representations (averaged across the two scanning days) and the similarity judgments obtained from the multiple arrangements task (within-and between subjects; red and black bar). The brain-behaviour correlation was significantly larger within than between subjects, indicating that an individual's similarity judgments are better predicted by that individual's mental representations than by another's.

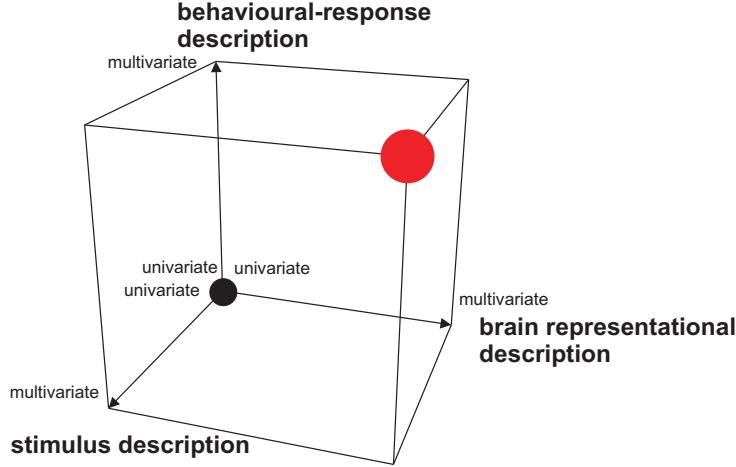


Figure 6: Multivariate description of stimulus, brain representation, and behaviour. Early studies have focussed on particular dimensions of stimulus, brain response, and behaviour. However, the sensory information, its representation in the brain, and behaviour are fundamentally multivariate phenomena. Recent studies have begun to engage this inherent complexity, linking stimulus to brain response, brain response to behaviour, or all three elements, while treating several of them in a multivariate framework.

ately. The same is true for the objects in the external world and our behavioural responses to them (Figure 6). The study by Charest et al. [23], cited above, demonstrated how we can investigate the individually unique relationships between multivariate brain representations and multivariate behavioural measures. The multivariate behavioural measure used in this study was similarity judgements. Judgements for all pairwise object comparisons were acquired using a novel multi-arrangement task [67, 68]. In this method, subjects communicate perceived object similarity by arranging object images on a computer screen by mouse drag-and-drop operations. In an adaptive measurement procedure, the subject is asked to arrange subsets of the objects in 2D according to their similarity. The higher dimensional perceptual similarity space is then inferred from the multiple arrangements by inverse MDS [68].

Recent studies have increasingly employed multivariate descriptions of the stimuli. One example cited above is the study by Mitchell et al. [29], who used a 25-dimensional semantic space to describe their stimuli. A more recent example is the study by Huth et al. [37] who defined detailed semantic predictors describing a movie stimulus. The predictors described the presence of objects from different categories in the scene and utilised the is-a hierarchy of the WordNet model [69]. Using a regularised linear regression on the voxel activity patterns, they showed how the continuous semantic space provided by the WordNet labels predicted the activity of each voxel in the brain [37].

The semantic models of Mitchell et al. [29] and Huth et al. [37] require the stimuli to come with labels. The computational model linking stimuli to brain representations, thus, is not fully explicit in these studies. The cited study by Kay et al. [30] provides an example of a test of a simple image-computable model (Gabor model). Characterising the multivariate nature of the

stimuli in the external world is now also possible for higher level representations, using more complex computational models. One influential model of object vision is the HMAX model [70, 71]. HMAX is a computational model inspired by the neuroscience literature, which aims at characterising the representations along the visual hierarchy. Efforts have been made to compare the representational geometry of the HMAX model to the representational geometry of activity patterns along the visual ventral stream [72, 20, 22, 73]. The classical implementation of the HMAX model so far fails to satisfactorily explain the categorical divisions of higher level object representations.

Recent advances with deep convolutional neural networks (DCNNs) have improved the performance of computers at visual object recognition [74, 75, 76]. These models are loosely biologically inspired in their hierarchical layered architecture and acquire their representations through supervised learning with large category-labelled image sets. Recent studies suggest that the performance achieved by the DCNNs in object classification approaches that of IT cortex in human and non-human primates [77]. The DCNN developed by A. et al. [74] outperformed other computer vision models at predicting the representational geometry of human and monkey IT [78] for isolated object images. These authors showed that the models strong categorical boundaries (acquired through strong supervision) contributed to its better performance at predicting the IT representational geometry. For the contextualised images from Charest et al. [23] as well, we observed that the A. et al. [74] model captures major categorical distinctions similarly to human IT (Figure 4, previously unpublished data). These recent studies have given us a better understanding of the computational mechanisms underlying object vision.

	<b>Do the analyses honour the patterns of response across...</b>		
	<b>voxels?</b>	<b>stimuli?</b>	<b>subjects?</b>
<b>Talairach/MNI block-design activation mapping</b> (e.g. Haxby et al., 1994)	No: Single-subject maps are smoothed for inferential mapping and/or voxels within an ROI averaged.	No: Responses to distinct stimuli within a block are averaged.	No: Single-subject maps (or ROI activations) are averaged across subjects for fixed- or random-effects analysis.
<b>Category-block activation group studies</b> (e.g. Kanwisher et al., 1997)	No: Voxels within a functional region are averaged across.	No: Responses to stimuli within each predefined category are averaged.	ROIs defined in each subject, but no detailed analysis of individual differences.
<b>Pattern-information category-block group studies</b> (e.g. Haxby et al., 2001)	ROI based analyses of voxel pattern information, but no detailed analyses of voxel-to-voxel variation.	No: Responses to stimuli within each predefined category are averaged.	Each subject allowed different representational patterns, but no detailed individual differences analysis.
<b>Single-stimulus representational geometry studies</b> (e.g. Kriegeskorte, Mur, Ruff, et al., 2008)	ROI based analyses of pattern representational geometry, no detailed mapping.	Each stimulus is a separate condition and is analysed inferentially with multiple-testing correction.	Each subject allowed different representational patterns. Interspecies, but no individual differences are analysed.
<b>Single-voxel encoding models</b> (e.g. Huth et al., 2012)	Detailed descriptive mapping of single-voxel responses (no inferential mapping with control of false-positives rate).	Movie responses modelled categorically (but extremely richly). Inferences pool effects over stimuli.	Each subject allowed different representational patterns, but no detailed analysis of individual differences.
<b>Common representational geometry studies</b> (e.g. Haxby et al., 2011)	Each voxel allowed a unique response profile and response pattern spaces linearly mapped between subjects.	Rich stimulus-space sample (movie) without categorical averaging, but inferences pool effects over stimuli.	Subject-unique patterns, but common geometries. Subject relationships modelled, but not analysed in detail.
<b>Individual representational geometry studies</b> (e.g. Charest et al., in press)	ROI based analyses of pattern representational geometry, no detailed mapping.	Each stimulus is a separate condition, but inferences pool effects over stimuli.	Subject-unique representational geometries demonstrated to be distinct and predictive of judgment idiosyncrasies.

no  
(averaged)
yes  
(but not analysed)
yes  
(target of analysis)

Figure 7: Honouring idiosyncrasies of voxels, stimuli, and subjects. Early neuroimaging studies (top row) tended to average across voxels, stimuli, and subjects. These studies gave us a coarse-scale view of brain function, revealing how the regional-average activation of large chunks of brain differed between tasks (each involving complex processing of particular stimuli, which were represented by block averages). The resulting brain maps were typically averaged across subjects. Recent studies have attempted to honour the idiosyncrasies of voxels, stimuli, and subjects. Here we compare particular exemplary studies (cited in the left column) with regard to the way they treated variation across these dimensions. We distinguish two levels of honouring differences between voxels, stimuli, and subjects: (1) The differences are modelled and not treated as noise, but they are not the target of the analyses (light gray). The motivation for this approach is usually the greater power such analyses can confer for testing hypotheses of interest. For example, pattern-information studies that do not average activation across voxels or subjects tend to have greater power for detecting information encoded in subject-unique patterns of activity within a region. (2) The differences are not only modelled, but form the target of the analyses. An additional more technical dimension that we do not consider here is the distinction between fixed- and random-effects analyses (across subjects, stimuli, or voxels). An analysis that treats subject as a random-effects factor is useful for generalising from the sample of subjects to the population. However, the variation across subjects is part of the noise model (thus not honoured) and the inference targets the population mean.

## Honouring the idiosyncrasies

Figure 7 explores the extent to which a number of exemplary studies conducted over the past two decades honoured the idiosyncrasies of voxels, stimuli and subjects. We see a clear development from averaging to more sophisticated data modelling that engages these idiosyncrasies. Future studies will push further along these dimensions, linking multivariate descriptions of the world via explicit computational models of brain information processing to multivariate measurements of brain activity in individual organisms and to multivariate measurements of behaviour. One framework for pursuing these directions is the analysis of representational dissimilarity matrices (Figure 4 and Figure 5; for a review, see [9]), which enable us to relate multivariate descriptions of stimuli, brain regions, model representations and behavioural responses without complex fitting procedures.

Beyond basic science, the characterisation of individually unique brain function is likely to contribute to our understanding of disorders and of the continuous variation across patients. If functional brain imaging is to become useful in the diagnosis and monitoring of patients, we will need to develop a rich repertoire of methods for characterising subtle functional differences between individual brains and minds. Disorders are increasingly recognised to fall on a complex manifold where every patient has a unique place. For example, individuals with an autism spectrum condition (ASC) tend to exhibit atypical perceptual processing [79]. However, every case is different. Grandin and Panek [80] describe three challenges that neuroimaging researchers of autism face: the absence of apparent brain structural abnormalities [81], the heterogeneity of causes of ASC and the heterogeneity of behavioural symptoms. In the future, it might become possible to characterise the biological underpinnings of the unique behavioural dysfunction of a given individual patient. Multivariate characterisations of an individual's brain representations might one day help tailor therapeutic interventions in the context of personalised medicine.

## Acknowledgement

We would like to thank Philip Pell for helpful comments on the manuscript.

## Disclosure statement

No potential conflict of interest was reported by the authors.

## Funding

This work was supported by the UK Medical Research Council [grant number MC-A060-5PR60] and by a European Research Council [grant number ERC-2010-StG 261352] to NK.

## References

- [1] D. H. Hubel and T. N. Wiesel. The period of susceptibility to the physiological effects of unilateral eye closure in kittens. *Journal of Physiology*, 206(5):419–436, 1970. [a](#)
- [2] A. Pascual-Leone, A. Amedi, F. Fregni, and L. B. Merabet. The plastic human brain cortex. *Annual Review of Neuroscience*, 28:377401, 2005. doi: 10.1146/annurev.neuro.27.070203.144216. [a](#)
- [3] A. Pascual-Leone, C. Freitas, L. Oberman, J. C. Horvath, M. Halko, M. Eldaief, S. Bashir, M. Vernet, M. Shafii, B. Westover, A. M. Vahabzadeh-Hagh, and A. Rotenberg. Characterizing brain cortical plasticity and network dynamics across the age-span in health and disease with tms-eeg and tms-fmri. *Brain Topography*, 24:302–315, 2011. doi: 10.1007/s10548-011-0196-8. [a](#)
- [4] N. Kriegeskorte and P. A. Bandettini. Analyzing for information, not activation, to exploit high-resolution fmri. *Neuroimage*, 38:649–662, 2007. doi: 10.1016/j.neuroimage.2007.02.022. [a](#), [a](#), [a](#)
- [5] N. Kanwisher, J. McDermott, and M. M. Chun. The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience*, 17:43024311, 1997. [a](#)
- [6] R. A. Epstein and N. Kanwisher. A cortical representation of the local visual environment. *Nature*, 392:598–601, 1998. doi: 10.1038/33402. [a](#)
- [7] P. E. Downing. A cortical area selective for visual processing of the human body. *Science*, 293(5539):24702473, 2001. doi: 10.1126/science.1063414. [a](#)
- [8] D. C. Dennett. *The intentional stance*. The MIT press, Cambridge, MA, 1989. [a](#)
- [9] N. Kriegeskorte and R. A. Kievit. Representational geometry: Integrating cognition, computation, and the brain. *Trends in Cognitive Sciences*, 17(8):401–412, 2013. doi: 10.1016/j.tics.2013.06.007. [a](#), [a](#), [a](#)
- [10] D. Marr. *Vision: A computational approach*. Freeman, New-York, NJ, 1982. [a](#)
- [11] D. Marr. *Vision: A computational investigation into the human representation and processing of visual information*. The MIT Press, Cambnridge, MA, 2010. [a](#)
- [12] K. Grill-Spector and K. S. Weiner. The functional architecture of the ventral temporal cortex and its role in categorization. *Nature Reviews Neuroscience*, 15:536–548, 2014. doi: 10.1038/nrn3747. [a](#)
- [13] L. G. Ungerleider and J. V. Haxby. "What" and "Where" in the human brain. *Current Opinion in Neurobiology*, 4:157–165, 1994. doi: 10.1016/0959-4388(94)90066-3. [a](#)
- [14] R. M. Cichy, D. Pantazis, and A. Oliva. Resolving human object recognition in space and time. *Nature Neuroscience*, 17:455462, 2014. doi: 10.1038/nn.3635. [a](#)
- [15] A. Clarke and L. K. Tyler. Object-specific semantic coding in human perirhinal cortex. *Journal of Neuroscience*, 34(14):4766–4775, 2014. doi: 10.1523/JNEUROSCI.2828-13.2014. [a](#)
- [16] C. R. Gallistel. *The organization of learning*, volume viii. The MIT Press, Cambridge, MA, 1990. [a](#)
- [17] J. H. Holland, K. J. Holyoak, R. E. Nisbett, P. R. Thagard, and S. W. Smoliar. *Induction: Processes of inference, learning, and discovery*. The MIT Press, Cambridge, MA, 1987. [a](#)
- [18] N. Kriegeskorte and G. Kreiman. *Visual population codes: Toward a common multivariate framework for cell recording and functional imaging*. The MIT Press, Cambridge, MA, 2012. [a](#)
- [19] J. V. Haxby, M. I. Gobbini, M. L. Furey, A. Ishai, J. L. Schouten, and P. Pietrini. Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293:24252430, 2001. doi: 10.1126/science.1063736. [a](#), [a](#), [a](#)
- [20] N. Kriegeskorte, M. Mur, and P. A. Bandettini. Representational similarity analysis connecting the branches of systems neuroscience. *Frontiers in Systems Neuroscience*, 2(4), 2008. doi: 10.3389/neuro.06.004.2008. [a](#), [a](#), [a](#)
- [21] Y. Kamitani and F Tong. Decoding the visual and subjective contents of the human brain. *Nature Neuroscience*, 8(5):679–685, 2005. doi: 10.1038/nn1444. [a](#), [a](#), [a](#)
- [22] N. Kriegeskorte, M. Mur, D. A. Ruff, R. Kiani, J. Bodurka, H. Esteky, K. Tanaka, and P. A. Bandettini. Matching categorical object representations in inferior temporal cortex of man and monkey. *Neuron*, 60(6):11261141, 2008. doi: 10.1016/j.neuron.2008.10.043. [a](#), [a](#), [a](#)
- [23] I. Charest, R. A. Kievit, T. W. Schmitz, D. Deca, and N. Kriegeskorte. Unique semantic space in the brain of each beholder predicts perceived

- similarity. *Proceedings of the National Academy of Sciences*, 111(40):1456514570, 2014. doi: 10.1073/pnas.1402594111. [a](#), [a](#), [a](#)
- [24] J. D. Haynes and G. Rees. Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience*, 7:523–534, 2006. doi: 10.1038/nrn1931. [a](#)
- [25] N. Kriegeskorte, R. Goebel, and P. A. Bandettini. Information-based functional brain mapping. *Proceedings of the National Academy of Sciences*, 103(10):38633868, 2006. doi: 10.1073/pnas.0600244103. [a](#)
- [26] M. Mur, D. A. Ruff, J. Bodurka, P. De Weerd, P. A. Bandettini, and N. Kriegeskorte. Categorical, yet graded - single-image activation profiles of human category-selective cortical regions. *Journal of Neuroscience*, 32:8649–8662, 2012. doi: 10.1523/JNEUROSCI.2334-11.2012. [a](#)
- [27] I. D. Popivanov, J. Jastorff, W. Vanduffel, and R. Vogels. Heterogeneous single-unit selectivity in an fmri-defined body-selective patch. *Journal of Neuroscience*, 34(1):95–111, 2013. doi: 10.1523/JNEUROSCI.2748-13. 2014. [a](#)
- [28] V. Troiani, A. Stigliani, M. E. Smith, and R. A. Epstein. Multiple object properties drive scene-selective regions. *Cerebral Cortex*, 24(4):883–897, 2012. doi: 10.1093/cercor/bhs364. [a](#)
- [29] T. M. Mitchell, S. V. Shinkareva, A. Carlson, K. M. Chang, V. L. Malave, R. A. Mason, and M. A. Just. Predicting human brain activity associated with the meanings of nouns. *Science*, 320:1191–1195, 2012. doi: 10.1126/science.1152876. [a](#), [a](#), [a](#)
- [30] K. N. Kay, T. Naselaris, R. J. Prenger, and J. L. Gallant. Identifying natural images from human brain activity. *Nature*, 452:352–355, 2008. doi: 10.1038/nature06713. [a](#), [a](#), [a](#)
- [31] N. Kriegeskorte. Pattern-information analysis: From stimulus decoding to computational-model testing. *Neuroimage*, 56:411421, 2011. doi: 10.1016/j.neuroimage.2011.01.061. [a](#)
- [32] S. O. Dumoulin and B. A. Wandell. Population receptive field estimates in human visual cortex. *Neuroimage*, 39(2):647–660, 2008. doi: 10.1016/j.neuroimage.2007.09.034. [a](#)
- [33] T. Çukur, S. Nishimoto, A. G. Huth, and J. L. Gallant. Attention during natural vision warps semantic representation across the human brain. *Nature Neuroscience*, 16:763–770, 2013. doi: 10.1038/nn.3381. [a](#)
- [34] A. Harel, D. J. Kravitz, and C. I. Baker. Task context impacts visual object processing differentially across the cortex. *Proceedings of the National Academy of Sciences*, 111(10):E962–E971, 2014. doi: 10.1073/pnas.1312567111. [a](#)
- [35] T. T. Rogers, J. Hocking, U. Noppeney, A. Mechelli, M. L. Gorno-Tempini, K. Patterson, and C. J. Price. Anterior temporal cortex and semantic memory: Reconciling findings from neuropsychology and functional imaging. *Cognitive, Affective & Behavioral Neuroscience*, 6:201–203, 2006. doi: 10.3758/CABN.6.3.201. [a](#)
- [36] L. K. Tyler, E. A. Stamatakis, P. Bright, K. Acres, S. Abdallah, J. M. Rodd, and H. E. Moss. Processing objects at different levels of specificity. *Journal of Cognitive Neuroscience*, 16:351–362, 2012. [a](#)
- [37] A. G. Huth, S. Nishimoto, A. T. Vu, and J. L. Gallant. A continuous semantic space describes the representation of thousands of object and action categories across the human brain. *Neuron*, 76:1210–1224, 2012. doi: 10.1016/j.neuron.2012.10.014. [a](#), [a](#), [a](#)
- [38] G. K. Aguirre. Continuous carry-over designs for fmri. *Neuroimage*, 35(4):1480–1494, 2007. doi: 10.1016/j.neuroimage.2007.02.005. [a](#)
- [39] A. Harel, D. J. Kravitz, and C. I. Baker. Deconstructing visual scenes in cortex: Gradients of object and spatial layout information. *Cerebral Cortex*, 23:947–957, 2013. doi: 10.1093/cercor/bhs091. [a](#)
- [40] N. Kriegeskorte, E. Formisano, B. Sorger, and R. Goebel. Individual faces elicit distinct response patterns in human anterior temporal cortex. *Proceedings of the National Academy of Sciences*, 104(51):20600–20605, 2007. doi: 10.1073/pnas.0705654104. [a](#)
- [41] K. A. Norman, S. M. Polyn, G. J. Detre, and J. V. Haxby. Beyond mind-reading: Multi-voxel pattern analysis of fmri data. *Trends in Cognitive Sciences*, 10:424–430, 2006. doi: 10.1016/j.tics.2006.07.005. [a](#)
- [42] A. J. O’Toole, F. Jiang, H. Abdi, and J. V. Haxby. Partially distributed representations of objects and faces in ventral temporal cortex. *Journal of Cognitive Neuroscience*, 17:580590, 2005. doi: 10.1162/0898929053467550. [a](#)
- [43] R. D. S. Raizada and A. C. Connolly. What makes different peoples representations alike: Neural similarity space solves the problem of across-subject fmri decoding. *Journal of Cognitive Neuroscience*, 24:868877, 2012. doi: 10.1162/jocn\_a.00189. [a](#)
- [44] M. Spiridon and N. Kanwisher. How distributed is visual category information in human occipito-temporal cortex? an fmri study. *Neuron*, 35:1157–1165, 2002. doi: 10.1016/S0896-6273(02)00877-2. [a](#)
- [45] K. S. Weiner and K. Grill-Spector. Sparsely-distributed organization of face and limb activations in human ventral temporal cortex. *Neuroimage*, 52:1559–1573, 2010. doi: 10.1016/j.neuroimage.2010.04.262. [a](#)
- [46] K. S. Weiner and K. Grill-Spector. Neural representations of faces and limbs neighbor in human high-level visual cortex: Evidence for a new organization principle. *Psychological Research*, 77(1):74–97, 2011. doi: 10.1007/s00426-011-0392-x. [a](#)
- [47] S. V. Shinkareva, V. L. Malave, M. A. Just, and T. M. Mitchell. Exploring commonalities across participants in the neural representation of objects. *Human Brain Mapping*, 33:13751383, 2012. doi: 10.1002/hbm.21296. [a](#)
- [48] E. Fedorenko, J. Duncan, and N. Kanwisher. Broad domain generality in focal regions of frontal and parietal cortex. *Proceedings of the National Academy of Sciences*, 110(41):1661616621, 2013. doi: 10.1073/pnas.1315235110. [a](#)
- [49] E. Fedorenko, P. J. Hsieh, A. Nieto-Castañon, S. Whitfield-Gabrieli, and N. Kanwisher. New method for fmri investigations of language: Defining rois functionally in individual subjects. *Journal of Neurophysiology*, 104(41):11771194, 2010. doi: 10.1152/jn.00032.2010. [a](#)
- [50] Z. M. Saygin, D. E. Osher, K. Koldewyn, G. Reynolds, J. D. E. Gabrieli, and R. R. Saxe. Anatomical connectivity patterns predict face selectivity in the fusiform gyrus. *Nature Neuroscience*, 15:321–327, 2011. doi: 10.1038/nn.3001. [a](#)
- [51] K. J. Friston, C. D. Frith, P. F. Liddle, R. J. Dolan, A. A. Lammertsma, and R. S. J. Frackowiak. The relationship between global and local changes in pet scans. *Journal of Cerebral Blood Flow & Metabolism*, 10:458466, 1990. doi: 10.1038/jcbfm.1990.88. [a](#)
- [52] K. J. Friston, A. P. Holmes, K. J. Worsley, J. B. Poline, C. D. Frith, and R. S. J. Frackowiak. Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, 2(4):189–210, 1994. doi: 10.1002/hbm.460020402. [a](#)
- [53] K. J. Worsley, A. C. Evans, S. Marrett, and P. Neelin. A three-dimensional statistical analysis for cbf activation studies in human brain. *Journal of Cerebral Blood Flow & Metabolism*, 12:900–918, 1992. doi: 10.1038/jcbfm.1992.127. [a](#)
- [54] R. Parasuraman and Y. Jiang. Individual differences in cognition, affect, and performance: Behavioral, neuroimaging, and molecular genetic approaches. *Neuroimage*, 59(1):70–82, 2012. doi: 10.1016/j.neuroimage.2011.04.040. [a](#)
- [55] N. Furl, L. Garrido, R. J. Dolan, J. Driver, and B. Duchaine. Fusiform gyrus face selectivity relates to individual differences in facial recognition ability. *Journal of Cognitive Neuroscience*, 23(7):1732–1740, 2011. doi: 10.1162/jocn.2010.21545. [a](#)
- [56] S. Martens, M. Kandula, and J. Duncan. Restricted attentional capacity within but not between sensory modalities: An individual differences approach. *PLoS One*, 5(12):e15280, 2010. doi: 10.1371/journal.pone.0015280. [a](#)
- [57] M. B. Miller, C. L. Donovan, C. M. Bennett, E. M. Aminoff, and R. E. Mayer. Individual differences in cognitive style and strategy predict similarities in the patterns of brain activity between individuals. *Neuroimage*, 59(1):83–93, 2012. doi: 10.1016/j.neuroimage.2011.05.060. [a](#)
- [58] K. N. Ochsner, S. A. Bunge, J. J. Gross, and J. D. E. Gabrieli. Rethinking feelings: An fmri study of the cognitive regulation of emotion. *Journal of Cognitive Neuroscience*, 14:12151229, 2002. doi: 10.1162/089892902760807212. [a](#)
- [59] M. W. Cole, T. Yarkoni, G. Repovs, A. Anticevic, and T. S. Braver. Global connectivity of prefrontal cortex predicts cognitive control and intelligence. *Journal of Neuroscience*, 32(26):89888999, 2012. doi: 10.1523/JNEUROSCI.0536-12.2012. [a](#)
- [60] M. J. Chadwick, H. M. Bonnici, and E. A. Maguire. Ca3 size predicts the precision of memory recall. *Proceedings of the National Academy of Sciences*, 111(29):10720–10725, 2014. doi: 10.1073/pnas.1319641111. [a](#)
- [61] D. S. Schwarzkopf and G. Rees. Subjective size perception depends on central visual cortical magnification in human v1. *PLoS One*, 8:e60550, 2013. doi: 10.1371/journal.pone.0060550. [a](#)
- [62] D. S. Schwarzkopf, C. Song, and G. Rees. The surface area of human v1 predicts the subjective experience of object size. *Nature Neuroscience*, 14(1):28–30, 2010. doi: 10.1038/nn.2706. [a](#)

- [63] R. D. S. Raizada, F. M. Tsao, H. M. Liu, and P. K. Kuhl. Quantifying the adequacy of neural representations for a cross-language phonetic discrimination task: Prediction of individual differences. *Cerebral Cortex*, 20(1):1–12, 2010. doi: 10.1093/cercor/bhp076. [a](#)
- [64] R. D. S. Raizada, F. M. Tsao, H. M. Liu, I. D. Holloway, D. Ansari, and P. K. Kuhl. Linking brain-wide multivoxel activation patterns to behaviour: Examples from language and math. *Neuroimage*, 51:462–471, 2010. doi: 10.1016/j.neuroimage.2010.01.080. [a](#)
- [65] L. T. Hsieh, M. J. Gruber, L. J. Jenkins, and C. Ranganath. Hippocampal activity patterns carry information about objects in temporal context. *Neuron*, 81(5):1165–1178, 2014. doi: 10.1016/j.neuron.2014.01.015. [a](#)
- [66] M. Mur, M. Meys, J. Bodurka, R. Goebel, P. A. Bandettini, and N. Kriegeskorte. Human object-similarity judgments reflect and transcend the primate-it object representation. *Frontiers in Psychology*, 4(128), 2013. doi: 10.3389/fpsyg.2013.00128. [a](#)
- [67] R. Goldstone. An efficient method for obtaining similarity data. *Behavior Research Methods, Instruments, & Computers*, 26:381–386, 1994. doi: 10.3758/BF03204653. [a](#)
- [68] N. Kriegeskorte and M. Mur. Inverse mds: Inferring dissimilarity structure from multiple item arrangements. *Frontiers in Psychology*, 3(245), 2012. doi: 10.3389/fpsyg.2012.00245. [a](#)
- [69] G. A. Miller. Wordnet: A lexical database for english. *Communications of the ACM*, 38(11):39–41, 1995. doi: 10.1145/219717.219748. [a](#)
- [70] M. Riesenhuber and T. Poggio. Hierarchical models of object recognition in cortex. *Nature Neuroscience*, 2:1019–1025, 2010. doi: 10.1038/14819. [a](#)
- [71] T. Serre, A. Oliva, and T. Poggio. A feedforward architecture accounts for rapid categorization. *Proceedings of the National Academy of Sciences*, 104(15):64246429, 2007. doi: 10.1073/pnas.0700622104. [a](#)
- [72] N. Kriegeskorte. Relating population-code representations between man, monkey, and computational models. *Frontiers in Neuroscience*, 3:363–373, 2009. doi: 10.3389/neuro.01.035.2009. [a](#)
- [73] D. D. Leeds, D. A. Seibert, J. A. Pyles, and M. J. Tarr. Comparing visual representations across human fmri and computational vision. *Journal of Vision*, 13(25):1–27, 2013. doi: 10.1167/13.13.25. [a](#)
- [74] Krizhevsky A., I. Sutskever, and G. E. Hinton. Imagenet classification with deep convolutional neural networks. In P Bartlett, editor, *Advances in neural information processing systems*, page 10971105. Neural Information Processing Systems ( NIPS ), December 2012. [a](#)
- [75] Q. V. Le, M. A. Ranzato, R. Monga, M. Devin, G. Corrado, K. Chen, J. Dean, and A. Ng. Building high-level features using large scale unsupervised learning. In John Langford and Joelle Pineau, editors, *Proceedings of the 29th International Conference on Machine Learning (ICML-12)*, pages 81–88. Omnipress, July 2012. [a](#)
- [76] M. D. Zeiler and R. Fergus. Visualizing and understanding convolutional networks. *CoRR*, abs/1311.2901, 2013. [a](#)
- [77] C. F. Cadieu, H. Hong, D. L. K. Yamins, N. Pinto, D. Ardila, E. A. Solomon, N.J. Majaj, and J. J. DiCarlo. Deep neural networks rival the representation of primate it cortex for core visual object recognition. *ArXiv e-prints*, 1406.3284, 2014. [a](#)
- [78] S. M. Khaligh-Razavi and N. Kriegeskorte. Deep supervised, but not unsupervised, models may explain it cortical representation. *PLoS Computational Biology*, 10(11):e1003915, 2014. doi: 10.1371/journal.pcbi.1003915. [a](#)
- [79] D. R. Simmons, A. E. Robertson, L. S. McKay, E. Toal, P. McAleer, and F. E. Pollick. Vision in autism spectrum disorders. *Vision Research*, 49: 27052739, 2009. doi: 10.1016/j.visres.2009.08.005. [a](#)
- [80] T. Grandin and R. Panek. *The autistic brain: Helping different kinds of minds succeed*. Houghton Mifflin Harcourt, Boston, MA, 2014. [a](#)
- [81] S. Haar, S. Berman, M. Behrmann, and I. Dinstein. Anatomical abnormalities in autism? *Cerebral Cortex*, page epub ahead of press, 2014. [a](#)