Title
Quantifying spatial pattern similarity in multivariate analysis using functional anisotropy

Key words: (max of 6)
fMRI, Group MVPA, multivariate statistics, functional anisotropy, searchlight

Authors:
Roee Gilron¹, Jonathan Rosenblatt², Oluwasanmi Koyejo³, Russell A. Poldrack³, Roy Mukamel¹

Institutions:
¹Sagol School of Neuroscience; School of Psychological Sciences, Tel Aviv University, Tel Aviv, Israel, ²Department of Industrial Engineering and Management; Zlotowsky Center for Neuroscience Ben Gurion University, Beer Sheva, Israel, ³Department of Psychology, Stanford University, Stanford, CA

Abstract:
Multivoxel pattern analysis (MVPA) has gained enormous popularity in the neuroimaging community over the past few years. At the group level, most MVPA studies adopt an “information based” approach in which the sign of the effect of individual subjects is discarded and a non-directional summary statistic is carried over to the second level. This is in contrast to a directional “activation based” approach which is typical in univariate group level analysis, in which both signal magnitude and sign are taken into account. The transition from examining effects in one voxel at a time vs. several voxels (univariate vs. multivariate) has thus tacitly entailed a transition from directional to non-directional signal definition at the group level. While a directional MVPA approach implies that individuals share multivariate spatial patterns of activity, in a non-directional approach each individual may have a distinct spatial pattern of activity. Here we show using an experimental dataset that indeed directional and non-directional MVPA approaches uncover distinct brain regions with some overlap. Moreover, we developed a descriptive measure to quantify the degree to which subjects share spatial patterns of activity. Our approach is based on adapting the Fractional Anisotropy (FA) measure, originally developed for examining diffusion MRI signals, in a novel way to quantify the degree to which subjects share a spatial pattern of activity. We term this measure “Functional Anisotropy” (FuA). Applying FuA to an auditory task, we found higher values in primary auditory regions compared to secondary and control regions. This highlights the potential of the FuA measure in second-level MVPA analysis to detect differences in the consistency of spatial patterns across subjects and their relationship to functional domains in the brain.
Intro:

In the last decade, the use of multivoxel pattern analysis (MVPA) to analyse fMRI data has grown substantially and is now commonplace (Haxby, 2012; Haynes and Rees, 2006; Kriegeskorte et al., 2006a; Poldrack and Farah, 2015; Tong and Pratte, 2012). In addition, many authors have shifted to the out-of-sample classification approach to MVPA, leading to additional concerns such as lower statistical power when compared to the analogous within-sample statistical test (Allefeld and Haynes, 2014; Ramdas et al., 2016). These shortcomings are most evident in multivariate group inference in which multivariate results from single subjects (first level) are carried over to group inference (second level). Here we expose some underlying tacit assumptions and suggest a novel measure – Functional Anisotropy - to characterize the level of consistency of multivariate patterns at the group level.

We focus on examples in which the signal of two conditions is compared. In a typical mass-univariate analysis, the BOLD signal in each individual voxel is examined separately by comparing values in each condition at the individual subject level (first level). This is typically done by performing a t-test examining the null hypothesis that the expected response is not different across conditions. In multivariate approaches, a distributed spatial pattern of activity across several voxels is compared (Haxby et al., 2001). Commonly in such cases, supervised machine learning approaches such as linear discriminant analysis or support vector machines (Kragel et al., 2012; Misaki et al., 2010; Mur et al., 2009; Tong and Pratte, 2012) are used, and their results are compared against an empirical null distribution - putatively centred around chance classification levels. At the group level (second level) analysis, univariate studies use a random effects (RFX) analysis to examine whether the difference between two conditions is consistent across
subjects. If the mean difference between conditions is significantly different from zero across subjects (as examined using a t-test for example) the voxel is declared responsive at the group level. Therefore to reject the null, one must show a directional group-wise effect (Fig. 1A). A directional effect is one in which most subjects display a consistent (either positive or negative) effect size in a given voxel. This takes into account both magnitude and sign (direction) of the effect. This directional effect was termed an “activation based” approach (Kriegeskorte et al., 2006b). If, for example, we had a cohort of subjects in which half of the sample showed an increase in their response to one condition relative to the other while the other half showed a decrease of equal magnitude in their response – a second level directional analysis would not define such an effect as signal. Put differently, this additional variability is part of the RFX null hypothesis and not part of the alternative. Although there is a strong effect size at the individual subject level, at the group level there is not a significant effect under such a directional definition of signal. Indeed, this directional approach is the commonly adopted signal definition in second-level mass-univariate RFX analysis.
Figure 1- Univariate & Multivariate Signal Definition

A  Univariate Signal - Group Level

**Nondirectional**

- Average difference between conditions in one subject
- Group Average

**Directional**

- Grey coloured circles represent the average difference (contrast) between conditions of interest (A and B) of individual subjects. The mean (group) average is represented by a filled black circle. In a directional univariate analysis (top right) the average group signal must be different from zero. In contrast, in a non-directional univariate analysis (left) the mean of the contrast across subjects may be zero.

B  Multivariate Signal - Group Level

**Nondirectional**

- Empty circles represent single trials, filled circles represent average difference, and black filled circle is group average.

**Directional**

Figure Caption:

Univariate and multivariate signal definition. This schematic diagram represents the different signal definitions in univariate and multivariate approaches employing either a directional or non-directional analysis. (A) Univariate group level analysis. Grey coloured circles represent the average difference (contrast) between conditions of interest (A and B) of individual subjects. The mean (group) average is represented by a filled black circle. In a directional univariate analysis (top right) the average group signal must be different from zero. In contrast, in a non-directional univariate analysis (left) the mean of the contrast across subjects may be zero. (B) Multivariate group level analysis. Empty circles represent single trials, filled circles represent average difference, and black filled circle is group average. In a non-
directional multivariate analysis (left) the spatial pattern of each subject does not have to be overlapping, as long as it is not centred around zero. Note that the group average can be centred around zero. In contrast, in a directional multivariate analysis (right) subjects share a spatial pattern of activity such that the population mean is not centred around zero.

In contrast, the large majority of MVPA studies to date have adopted a non-directional (information based) definition of signal at the group level (Fig. 1B). In a non-directional analysis, a certain measure (usually classification accuracy) is calculated at the individual subject level, and this measure is then carried over to the second level (Haxby et al., 2014; Kriegeskorte et al., 2006b; Tong and Pratte, 2012). Note that as opposed to the t-test metric, the accuracy metric is directionless, thus the sign of the effect at the first level is lost and only its magnitude (or a corresponding statistical metric) is passed on to the second level analysis. In the example (Fig 1A – left) half of the subjects show an increase in their response to one condition vs. the other while the other half of subjects show a decrease of equal magnitude. Thus effect size at the individual subject level is large and would be reflected in a corresponding high statistical metric value (e.g. classification accuracy) that is carried to the second level. Since all subjects have a large effect size, such a case would be detected by a non-directional 2nd level analysis, irrespective of the fact that different subjects show completely opposite patterns of responses. The equivalent univariate null hypothesis of a non-directional signal definition is that across subjects, the expected absolute value of the effect in a given voxel (e.g. average difference between conditions) is zero. Thus a signal would be detected even if in some subjects there is a positive effect and in other subjects there is a negative effect in a particular voxel. However, this non-directional univariate approach is seldom taken when conducting group-level univariate analysis since the biological significance of such an effect would be deemed suspect. That is, it would be challenging to interpret a study in which half of the subjects show an increase in BOLD response whereas the other half show a decrease in BOLD response to a given contrast. The
transition from examining effects in one voxel at a time vs. several voxels (univariate vs. multivariate) has thus tacitly entailed a transition from directional to non-directional signal definition at the 2nd level. Studies that opted for an MVPA directional signal definition are rare, primarily occurring in cases where classifiers are used to predict activity in new individuals (e.g. - Helfinstein et al., 2014).

Importantly, 2nd level multivariate non-directional analysis implies a fundamentally different definition of signal compared to a traditional univariate directional analysis. This represents an implicit paradigm shift the field has undergone. There is no a-priori reason to believe that moving from univariate (single voxel) to multivariate analysis (2 voxels or more) requires a redefinition of the null and alternative hypothesis. Moreover the original motivation of multivariate approaches was to uncover weak distributed signals as well as information at finer spatial scales than fMRI affords (Haxby, 2012; Haxby et al., 2001; Haynes and Rees, 2006; Kamitani and Tong, 2005; Kriegeskorte et al., 2006b). These papers did not make any explicit hypothesis about the differing nature of the signal across subjects. In addition to differences in signal definition between directional and non-directional analysis, both types of analysis carry over a statistical measure to the second level. The practice of carrying over a statistical measure to second level analysis is subject to a variety of confounds that may provide alternative explanations for the results (Todd et al., 2013).

Using an empirical data set we show that divergent definitions of the null hypothesis governing 2nd-level directional and non-directional analysis yield different results. Furthermore, we developed a descriptive measure to quantify the degree to which subjects share the same spatial pattern of activity in a given region of interest (ROI). Our measure employs an adaptation of the metric of Fractional Anisotropy (FA) often used in analysis of
diffusion weighted imaging (DWI). In DWI, low FA values are interpreted as water molecule displacement that is equiprobable along the three cardinal axes and high FA values point to a restriction of water displacement in a certain direction. These differences in FA values are interpreted as evidence that cellular microstructure (such as axons) is oriented in a particular direction. Since FA captures multivariate asymmetry in space, we exploit this measure in our analysis to quantify the degree of asymmetry (directionality) of the multivariate signal across subjects. Since we employ this metric on BOLD functional data rather than DWI we suggest the term functional (instead of fractional) anisotropy and denote it FuA.

**Material and Methods:**

Data Set:

We used an fMRI dataset which involved a localizer used to identify areas sensitive to human voices (vs. non-man-made voice) in the auditory cortex (Pernet et al., 2015). The experimental procedure is described in detail in their original paper, therefore we describe it here in brief. The voice localizer is a 10:20 min block design fMRI experiment. The experiment consisted of forty 8 second long blocks of human vocalizations (20 blocks) or non-vocal (20 blocks) stimuli. A few periods of silence (10 seconds) were interspersed between the experimental blocks to allow the hemodynamic response to relax. Vocal blocks were primarily sounds of human vocal origin obtained from 47 speakers while non-vocal sounds are mostly from natural or man-made sources (like cars). Scans were acquired using a 3T Siemens (Erlangen, Germany) Tim Trio using a repetition time (TR) of 2 s with an echo time (TE) of 30ms. Additional scan parameters can be found in the original paper. Data and analysis were graciously shared by the authors and can be found at
https://openfmri.org/dataset/ds000158. Data were acquired at a 3x3x3.3mm resolution, and the rostral part of the frontal cortex was not scanned in some of the subjects.

fMRI data analysis:

We used the pre-processing analysis code used by Pernet and colleagues which can be found on the OpenfMRI link above. The original data set contained 218 subjects. Whole brain functional coverage was acquired for 150 subjects, from which we randomly chose 20 subjects. We chose 20 subjects in order to obtain a sample size concordant with many fMRI studies and to avoid trivial power gains. Data were analysed using SPM12b (r6225 – Welcome Department of Cognitive Neurology, University College London). Pre-processing consisted of slice time correction, motion correction (6 parameters), co-registration of the structural image to the mean functional image and normalization of the structural image to the Montreal Neurological Institute (MNI) space (diffeomorphic normalization with the forward deformation field computed during segmentation, data was resampled at 2 mm isotropic with 4th degree B-spline interpolation). These spatial transformations were then applied to the functional images to achieve normalization to the Montreal Neurological Institute (MNI) space. Data were high pass filtered (1/128s) to remove slow drift.

Pre-processing:

We used a design matrix that contained separate regressors for each trial. Each regressor was modelled by convolving a boxcar function describing the timing of stimulus events with the canonical hemodynamic response function (HRF) used in SPM. Since strong correlations between trial-wise beta estimates still existed in the data after the use of the default AR(1)
serial correlation model, we used the AR(3) serial correlation model (implemented in SPM12) to remove correlations in the beta estimates and model each beta value separately (Penny et al., 2005). Thus a separate beta value was estimated for each block resulting in a total of 40 beta values per subject (20 vocal, 20 non-vocal). Since scan coverage was not identical across subjects, we created a Boolean ‘AND’ map of all subjects’ functional data masks in order to allow us to easily compare only signals in voxels that are common across subjects. This resulted in a matrix of 40 beta values X 32,482 voxels per subject. In both the directional and non-directional analysis detailed below, we used a searchlight approach analysed in a similar fashion to that employed in a previous paper (Krasovsky et al., 2014). For each center voxel, beta values from its 26 closest voxels were used in the data analysis. Thus each searchlight beam in a single subject was represented by a 40 x 27 matrix corresponding to 40 beta values per voxel (20 vocal, 20 non-vocal trials) by 27 voxels (center voxel + 26 closest neighbours using Euclidian distance).

Detecting signal in a searchlight beam – the statistical test

When facing a multivariate comparison between two conditions, most neuroimaging studies have employed a supervised machine learning approach in which performance is assessed through testing of out-of-sample generalization (e.g., via cross-validation). While this approach is useful in assessing the generalizability of the results, it is also substantially more conservative, and a number of studies have suggested the use of in-sample hypothesis testing over out-of-sample classification for multivariate comparison (Allefeld and Haynes, 2014; Kriegeskorte et al., 2006a; Ramdas et al., 2016). There is a wide body of statistical literature concerned with detecting multivariate differences between populations (Harris, 2014). Allefeld and Haynes (2014) proposed a variation on Hotelling’s Trace as their test statistic.
(Hotelling, 1944; Lawley, 1938). Here we use a related multivariate statistical test which is better equipped to deal with cases in which the number of features (voxels in the searchlight) is larger than the number of observations (trials or subjects in our case). This test, recently developed by Strivastava and Du (2008) is both quick to compute, and more powerful than Hotteling type tests for the dimensions of a searchlight used in a typical MVPA fMRI setup. Both the directional and non-directional tests have the same general structure, which consists of testing for the expectation of the group effect, based on some first (subject) level summary statistic (Mumford and Nichols, 2009). We thus denote by $T_i$ subject $i$’s summary statistic of a beam centred at voxel $v$ (voxel index omitted). We denote by $T$ the group level summary of the same beam. Under the summary statistic approach, $T = (T_1, ..., T_n)$, where $n$ is the number of subjects. We will also denote by $p$ the number of voxels in a beam, and by $m$ the number of repeats of each stimulus (trials), which is the same for the two stimuli in our balanced design.

In the most general case, each beam is fitted with a multivariate general linear model, and then signal detection can be performed with any test for the coefficients of a MANOVA such as Wilk’s Lambda, Pillai-Bartlett Trace, Lawley-Hotelling Trace, or Roy’s Greatest Root test (Anderson, 2003). This was indeed the framework in Allefeld and Haynes (2014). In our two-stimuli case, all these tests collapse to the classic Hotelling $T^2$ test, which is perhaps the best known multivariate test. It is however notoriously low powered when the number of parameters ($p$) is in the same order as the number of samples ($n$ or $m$) (Dempster, 1963). In our analysis, we found the Srivastava-Du (2008) statistic to be the most powerful metric for search-light MVPA when compared to Hotelling’s $T^2$, and Dempster’s statistic (1963).

We now present our multivariate directional and non-directional tests. In order to assess statistical significance we use the permutation scheme of Stelzer et al. (2013) as discussed in the “significance testing” section.
Whole brain searchlight MVPA directional analysis

Our statistic for detecting directional signal consists of applying the one-sample multivariate test described in (Srivastava and Du, 2008) to the directional summary from the first level. Formally, let $c_i$ be subject $i$’s vector valued estimated contrast of interest. In our example, each of the $p$ coordinates of $c_i$ encodes the difference in the mean response between vocal and non-vocal response. More generally, it may be the output of any contrast in a multivariate linear model.

The directional test we propose consists of the following two levels:

$$T_i^D := c_i,$$

$$T^D := T^{2008}(T_1^D, \ldots , T_n^D) = T^{2008}(c_1, \ldots , c_n).$$

$T^{2008}(\ldots )$ is the one-sample Srivastava-Du test, which is defined as

$$T^{2008}(c_1 \ldots c_n) := \frac{n \bar{\epsilon}^t D^{-1} \bar{\epsilon} - \frac{n^* p}{n^* - 2}}{\sqrt{2d(tr(R^2) - \frac{p^2}{n^2})}}$$

where $n^* = n - 1; \bar{\epsilon} := 1/n \sum_i c_i; S := 1/n^* \sum_i (c_i - \bar{\epsilon})(c_i - \bar{\epsilon})^t; D := diag(S), R := D^{-1/2} SD^{-1/2} ;$ and $d := 1 + \frac{tr(R^2)}{p^{3/2}} .$

In this test each subject is essentially summarized by its raw contrast estimate, which serves to test for a null group multivariate mean. The first level statistic, $c_i$, is trivially directional. The test statistic $T^{2008}$ can be seen as Hotelling’s $T^2$ computed under a spatial (between voxel) independence assumption, and then corrected to relax this assumption. For more on the design and motivation of this test statistic, see (Srivastava and Du, 2008).
Whole brain searchlight MVPA non-directional analysis

For the non-directional version of the group test, each subject is summarized by a non-directional measure of signal. Hotelling’s two group test is a natural candidate, but again, we will want to replace it with a high-dimensional version in which the number of features (voxels) can be larger than the number of samples (trials in our case).

Seeing the two conditions in our example as a balanced block design so that \( x_{i,j} \) denotes subject \( i \)'s \( j \)th response to a vocal stimulus, and \( y_{i,j} \) the same for a non vocal stimulus, the non-directional test we propose has the following form:

\[
T_{i}^{\text{ND}} := T_{2013}^{(x_{i,1}, \ldots x_{i,m}, y_{i,1}, \ldots y_{i,m})},
\]

\[
T^{\text{ND}} := \frac{1}{n} \sum_{i=1}^{n} T_{i}^{\text{ND}}.
\]

where \( T_{2013}^{(x_{i,1}, \ldots x_{i,m}, y_{i,1}, \ldots y_{i,m})} \) is the two-sample Srivastava-Du test defined as

\[
T_{2013}^{(x_{i,1}, \ldots x_{i,m}, y_{i,1}, \ldots y_{i,m})} := \frac{m \frac{1}{2} \delta D^{-1} \delta - p}{\sqrt{2d(\text{tr}(R^{2}) - \frac{p^{2}}{m^{2}})}},
\]
where $m^* = m - 1; \bar{\delta} := \bar{x} - \bar{y}; S_x := 1/m^* \sum_i (x_i - \bar{x})(x_i - \bar{x})^t; S_y := 1/m^* \sum_i (y_i - \bar{y})(y_i - \bar{y})^t; S := (S_x + S_y)/2; D := diag(S); R := D^{-1/2}SD^{-1/2};$ and $d := 1 + \frac{\text{tr}(R^2)}{p^{3/2}}$.

Like the single sample case, this test can be seen as a two sample Hotelling $T^2$ test, corrected for the relaxation of the assumption of (spatial) independence.

In this test each subject is summarized by a beam-wise difference between conditions, which is later averaged over subjects. To verify that the first level statistic is non-directional, one may observe that $T^2_{2013}$ is a scaled and shifted quadratic form in the difference between group means ($\bar{\delta}$). As such, and just like the squared univariate $t$-statistic, it grows when $\bar{x} > \bar{y}$, and also when $\bar{x} < \bar{y}$. We also note that while we assumed a balanced design, this assumption is relaxed in (Srivastava et al., 2013) Section 4.4.
Figure 2- Analysis Scheme

2\textsuperscript{nd} level Nondirectional MVPA

Single Subjects

Group Level

2\textsuperscript{nd} level Directional MVPA

Single Subjects

Group Level
Analysis scheme. The top panel describes the analysis scheme for non-directional MVPA. At the first level, for each center voxel, in each subject, a matrix (trials x voxels) from each condition is used in order to calculate $T^{2013}$. The circles represent trial labels and the squares activity in a particular voxel (feature) and a particular trial. The $T^{2013}$ value is calculated for each center voxel for each subject. Note that $T^{2013}$ is an unsigned statistic. On the second level, the single subject $T^{2013}$ values are averaged to create a group $T^{ND}$ map composed of average $T^{2013}$ across subjects for each voxel.

The bottom panel describes the analysis scheme for directional MVPA. Here the first level summary statistic of each subject ($T^D_i$) is simply the difference between the average activity in each condition. Then, on the second level these average difference matrices from each subject are aggregated and the group $T^D$ value is calculated in one step for each center voxel.

Significance Testing

To threshold our directional and non-directional group level analysis we employ the same non-parametric permutation scheme described by Stelzer et al. (2014). We shuffle the condition labels across trials within each subject, and compute $T^D$ and $T^{ND}$ value using the same pipeline described above to generate a shuffled map. For each shuffled permutation map we use the same shuffling scheme across all searchlight beams so that spatial correlations in the noise are conserved. It is important to note that this dataset still had strong correlations between trial-wise beta estimates before the use of the AR(3) model to whiten the noise process. If one were to use the default AR(1) model in SPM, a naïve permutation scheme would underestimate the number of significant voxels due to dependence between trial-wise estimates. Once we used an AR(3) model to whiten the noise, trial-wise estimates were no longer correlated and in accordance with Stelzer (2013), we computed 1,000 shuffled label whole brain searchlight maps for each subject. We created group level shuffled-label maps by averaging randomly selected maps from each subject’s shuffled maps (with replacement). Within each voxel we used the distribution of shuffled values to compute a
corresponding voxel-wise p-value for both the $T^{ND}$ and $T^D$ maps. In this way we associate a p-value with each voxel in the brain and can submit these p-values to false discovery rate FDR control with the BH (Benjamini and Hochberg, 1995) procedure to create a binary map of the center voxels which pass significance. An example of this implementation can be found in the accompanying code.

From Fractional to Functional Anisotropy - Computing a Multivariate Similarity Measure

A distinction between directional and non-directional signal is a distinction between symmetrically distributed effects across subjects and non-symmetric ones. Multivariate symmetry is more delicate than univariate symmetry, as discussed in (Serfling, 2004). Several multivariate measures of symmetry can be found in the non-parametric statistics literature (e.g. Oja and Randles, 2004). This is simply because any rank-based test for location is sensitive to the violation of the symmetry assumption, and can thus serve as an asymmetry measure.

An alternative to the rank-based symmetry measures is motivated by the diffusion-imaging literature. Viewing each subject’s estimated contrast in a beam, as a vector of dimension $p$, the data of $n$ subjects is thus a cloud of $n$ points in $\mathbb{R}^p$. If each subject shows a unique spatial pattern, this cloud will fill the space in an approximately symmetric fashion. If, on the other hand, subjects show consistent spatial patterns, this cloud will break symmetry, and “point” in some direction in $\mathbb{R}^p$. The singular value of that direction in a singular value decomposition (SVD) will stand out from singular values in other directions. This suggests that the variance of the singular values of the contrasts in the beam can serve as a measure of symmetry. Instead, we prefer the variance of the squared singular values, since it is equivalent to the FA in diffusion imaging. We also note that since we seek symmetry about
the origin, and not about the centre of the data, then the data should not be centred when computing the SVD.

Formally, let $c_i$ be subject $i$’s vector-valued contrast of interest, $n$ the number of subjects, and $C$ the $n \times p$ data matrix in a beam. By the SVD decomposition we have that $C = USV^t$ where $U$ is a $n \times n$ unitary matrix, $S$ is a $n \times p$ rectangular diagonal matrix of singular values, and $V$ is a $p \times p$ unitary matrix. Denoting by $s_j$ the $j$’th singular value of $C$, we define

$$FuA := b \sqrt{\frac{\sum_{j=1}^{p} (s_j^2 - \bar{s}^2)^2}{\sum_{j=1}^{p} s_j^4}}$$  \hspace{1cm} (7)$$

where $\bar{s}^2 = \frac{1}{p} \sum s_j^2$, and $b$ is some known constant, set at $b = \sqrt{3/2}$ for compatibility with FA in DTI. The definition in Eq. (7), is illuminated by the well known relation between singular values and eigenvalues. We recall that for centred data, $s_j^2 = \lambda_j(n - 1)$, where $\lambda_j$ is the $j$’th eigenvalue of the empirical covariance $C'C/(n-1)$. Eq.(7) thus amounts to

$$FuA := b \sqrt{\frac{\sum_{j=1}^{p} (\lambda_j - \bar{\lambda})^2}{\sum_{j=1}^{p} \lambda_j^2}}$$  \hspace{1cm} (8)$$

which is the definition of the FA of the diffusion tensor when viewing the empirical covariance as a diffusion tensor.
**Results:**

Our whole brain directional MVPA searchlight analysis found a number of regions which survived voxelwise control (FDR $\leq 0.05$) including primary auditory cortex, inferior frontal gyrus and orbitofrontal gyrus (see Table 1). A total of 1,438 voxels passed significance (Fig 3. red and blue). In addition, our whole brain MVPA non-directional searchlight analysis revealed a number of partially overlapping brain regions which survived voxelwise control (FDR $\leq 0.05$) including primary auditory cortex, inferior frontal gyrus and bilateral insula (Table 1). A total of 973 voxels passed significance (Fig. 3, green and blue). Overlapping voxels that are common to both the directional and non-directional analysis are shown in blue (a total of 621 voxels).

| Type       | x    | y    | z    | Cluster Size | Anatomical Label                  |
|------------|------|------|------|--------------|-----------------------------------|
| Directional| 65   | -29  | 12   | 291          | RH Superior Temporal Gyrus        |
|            | 53   | 32   | 4    | 27           | RH Inferior Frontal Gyrus         |
|            | 45   | 8    | 44   | 24           | RH Middle Frontal Gyrus           |
|            | 15   | -5   | 9    | 23           | RH Thalamus                       |
|            | -24  | -41  | -3   | 50           | LH Sub-Gyral                      |
|            | -30  | -80  | 32   | 31           | LH Cuneus                         |
|            | -36  | -29  | 12   | 267          | LH Transverse Temporal Gyrus      |
|            | -39  | -7   | 39   | 22           | LH Precentral Gyrus               |
| Nondirectional | 56  | 21   | 4    | 86           | RH Inferior Frontal Gyrus         |
|            | 53   | -4   | 42   | 56           | RH Precentral Gyrus               |
|            | -30  | 15   | 19   | 39           | LH Insula                         |
|            | -30  | 26   | -1   | 44           | LH Insula                         |
|            | -50  | -40  | 16   | 28           | LH Superior Temporal Gyrus        |
| Common     | 65   | -20  | 4    | 365          | RH Superior Temporal Gyrus        |
|            | -48  | -37  | 10   | 232          | LH Superior Temporal Gyrus        |

Table caption:

Details regarding all clusters found across directional and non-directional analysis. LH is left hemisphere, RH is right hemisphere. Cluster size refers to the number of voxels found in the cluster. Only clusters sizes above 20 contiguous voxels are shown.
Figure 3- Unique Contributions of Directional and Non-Directional Signal

Figure Caption:

Significant voxels defined by directional and non-directional analysis. These maps represent an overlay of voxels which were declared significant using either a directional analysis only (red), a non-directional analysis only (green), or both (blue).

Our descriptive FuA measure captures how similar the spatial pattern of activation is within each searchlight across subjects. We used the regions discovered in the directional and non-
directional MVPA searchlight analysis as a mask and overlaid their corresponding FuA values (Fig 4A). We also present the values in these brain regions in histogram form (Fig. 4B). Note the higher FuA values in voxels defined by the directional analysis and common regions – consistent with the notion of similarity in activity patterns across subjects obtained using a directional analysis. We performed this calculation both on the 20 subjects we randomly drew from the original data set, and also on all 150 subjects.

Figure 4 - FuA values

![Figure 4](image)

Figure Caption:
FuA values. The top panel overlays the FuA values on the significant regions from Figure 3. Numbers on sagittal slices represent Y plan coordinates in MNI space. The bottom panel
describes the distribution of FuA values in voxels obtained with directional (red) and non-directional (green) analysis. Since the number of voxels obtained in the directional and non-directional analysis is not identical, the histogram was normalized such that the Y axis represents the proportion of voxels with a particular FuA value (for details see Table 1). We computed the FuA values both in 20 subjects (bottom, left) and 150 subjects (bottom, right).

Finally, we computed the FuA value in a number of ROIs defined by the Harvard – Oxford atlas including, primary and secondary auditory, motor and sensorimotor cortices. We hypothesized the regions lower in the auditory hierarchy would show common multivariate spatial patterns, reflected by higher FuA values, whereas regions higher up in auditory processing hierarchy would be more variable across subjects and thus show lower FuA values. Indeed, we found higher FuA values in primary auditory regions relative to the control regions (Fig. 5).
Figure 5- Anatomical Regions for ROI - FuA analysis

Figure Caption:
Top panel - anatomical ROI’s from the probabilistic Harvard-Oxford atlas used in FuA analysis. Bottom panel - FuA values are larger in primary auditory regions compared to secondary auditory and control regions.
Discussion:

We show that in the move from univariate to multivariate group analysis, the field underwent a paradigm shift in the definition of the null hypothesis. Most MVPA studies to date implicitly employ a non-directional analysis. This means that subjects do not necessarily share the same pattern of neural activity and may even have opposite activity patterns. However, the motivation of the first papers that popularized MVPA was that it would allow researchers to discover patterns at lower than single voxel resolution - such as orientation columns in visual cortex (Haynes and Rees, 2006; Kamitani and Tong, 2005), or discover weak, subthreshold effects (Haxby, 2012). The bias towards non-directional effects in MVPA analysis is in stark contrast to the hypothesis underlying univariate group analysis - namely that subjects share the same pattern of information. Indeed we found that employing a directional MVPA analysis uncovers regions that partially overlap with non-directional analysis, but importantly also new regions. The early papers that popularized MVPA did not make any explicit hypothesis with regard to shared spatial patterns across subjects. Perhaps if these early studies tested not only non-directional but also directional tests they would uncover overlapping representations across subjects. Indeed, we find strong directional patterns in primary auditory cortex which we do not find using a non-directional test alone. The fact that directional and non-directional analysis reveal partially overlapping brain areas may also help shed light on the poor correspondence that has sometimes been observed between multivariate and univariate analysis (Jimura and Poldrack, 2012).

By proposing an informed choice between directional and non-directional tests, we believe the neuroimaging community will gain a better understanding of the type of signal one is discovering. A sharper definition of the nature of spatial patterns of activity across subjects can have important implications for the study of patient populations and design of brain computer interfaces. For example, in some decoding applications, learning a model which
works for all subjects may be desirable. Such a scheme will only be fruitful in regions showing a *directional* signal – or a shared spatial pattern of information across subjects (Helfinstein et al., 2014). In contrast, the spatial activity patterns in regions with non-directional signal (i.e. each subject has a unique spatial pattern of information) are expected to generalize to a lesser extent from one subject to another.

Importantly, performing a directional or non-directional analysis does not require the acquisition of new data and both analysis approaches can be employed using the same data set. However, the two have different definition of signal and consequently yield non-identical results.

To quantify the differences between directional and non-directional analysis, we developed a novel tool to measure multivariate pattern overlap across subjects. We borrow the Fractional Anisotropy (FA) measure, used in analysis of diffusion-weighted MRI and adapt it to functional data (EPI) used in MVPA. We call this measure *functional* anisotropy (FuA). This allows measuring the degree of directional agreement over subjects instead of a-priori committing to one particular type of test.

Fractional anisotropy (FA) is commonly used in diffusion scans and describes the degree to which a multivariate pattern is clustered along one of 3 dimensions in order to infer the spatial orientation of white matter tracts. However, it is also possible to generalize FA to arbitrary number of dimensions, as a measure of multivariate asymmetry. The actual numerical values of FuA depend on the number of dimensions (voxels) in a searchlight beam / ROI and also on the number of subjects. Therefore it is not comparable with typical FA values that are commonly reported in which the input signal is very different (diffusion weighted data from one subject in 3D vs EPI data with p-dimensional searchlight beams from multiple subjects). For this reason comparing FuA values from the same data using different
searchlight sizes is also not trivial. The optimal number of subjects needed in order to observe a difference in FuA values is also a subject to further research.

Unique spatial patterns across subjects imply symmetry, and thus low FuA values. Indeed we find low FuA in control regions, and in regions higher up in auditory processing – where each subject has its own unique spatial patterns (Fig. 5). In contrast we see higher FuA values corresponding to directional signal in primary auditory regions. Computing FuA can prove informative in many cases. For example, it is possible that primary sensory regions show similar spatial patterns of activity across subjects (e.g. due to similar tonotopic representations), whereas higher level association areas are more idiosyncratic and display specific activity patterns unique to each subject. Moreover, once a set of brain regions is discovered to be associated with a certain task, one can probe sub regions using FuA to characterize degrees of multivariate pattern “personalization” across subjects as they relate to hierarchical models of neural processing. For example it can be used to highlight regions that process tasks in a similar fashion across subjects versus regions which have unique processing across subjects. Last, using FuA could aid in targeting regions for brain computer interface (BCI) development. Since one of the challenges of modern BCI implementation is the need to learn a specific classifier for each subject, using FuA can help identify regions which have stable multivariate patterns across subjects so that the same model (classification model – such as SVM) could be used across subjects.

A natural extension of this work would be to assess the replicability of multivariate signals across studies in both directional and non-directional analysis frames. For instance, the different definitions of signals in directional and non-directional hypothesis may also have differential replicability prospects on both the single subject and study level.

**Acknowledgments:**

The study was supported by the I-CORE Program of the Planning and Budgeting Committee and The Israel Science Foundation (grant No. 51/11), Human Frontiers Science Project
Organization (HFSPO) (CDA00078/2011-C) and Israel Science Foundation (grants No. 1771/13 and 2043/13) to R.M. and the National Science Foundation (OCI-1131441) to R.A.P.

* We offer code at https://github.com/roeegilron/Multi-TandFuA and maps at http://neurovault.org/collections/978/, Original data can be found https://openfmri.org/dataset/ds000158.
References

Allefeld, C., Haynes, J.-D., 2014. Searchlight-based multi-voxel pattern analysis of fMRI by cross-validated MANOVA. Neuroimage 89, 345-357.

Anderson, T.W., 2003. An Introduction to Multivariate Statistical Analysis, 3 edition ed. Wiley- Interscience, Hoboken, NJ.

Benjamini, Y., Hochberg, Y., 1995. Controlling the false discovery rate - a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society Series B- Methodological 57, 289-300.

Dempster, A.P., 1963. Multivariate Theory for General Stepwise Methods. The Annals of Mathematical Statistics 34, 873-883.

Harris, R.J., 2014. A Primer of Multivariate Statistics, 3rd ed. ed.

Haxby, J.V., 2012. Multivariate pattern analysis of fMRI: The early beginnings. Neuroimage 62, 852-855.

Haxby, J.V., Connolly, A.C., Guntupalli, J.S., 2014. Decoding Neural Representational Spaces Using Multivariate Pattern Analysis. Annual Review of Neuroscience, Vol 37 37, 435-456.

Haxby, J.V., Gobbini, M.I., Furey, M.L., Ishai, A., Schouten, J.L., Pietrini, P., 2001. Distributed and overlapping representations of faces and objects in ventral temporal cortex. Science 293, 2425-2430.

Haynes, J.-D., Rees, G., 2006. Decoding mental states from brain activity in humans. Nature Reviews Neuroscience 7, 523-534.

Helfinstein, S.M., Schonberg, T., Congdon, E., Karlsgodt, K.H., Mumford, J.A., Sabb, F.W., Cannon, T.D., London, E.D., Bilder, R.M., Poldrack, R.A., 2014. Predicting risky choices from brain activity patterns. Proc Natl Acad Sci U S A 111, 2470-2475.

Hotelling, H., 1944. Some Improvements in Weighing and Other Experimental Techniques. 297-306.

Jimura, K., Poldrack, R.A., 2012. Analyses of regional-average activation and multivoxel pattern information tell complementary stories. Neuropsychologia 50, 544-552.

Kamitani, Y., Tong, F., 2005. Decoding the visual and subjective contents of the human brain. Nat Neurosci 8, 679-685.

Kragel, P.A., Carter, R.M., Huettel, S.A., 2012. What makes a pattern? Matching decoding methods to data in multivariate pattern analysis. Frontiers in Neuroscience 6.

Krasovsky, A., Gilron, R., Yeshurun, Y., Mukamel, R., 2014. Differentiating Intended Sensory Outcome from Underlying Motor Actions in the Human Brain. Journal of Neuroscience 34, 15446-15454.

Kriegeskorte, N., Goebel, R., Bandettini, P., 2006a. Information-based functional brain mapping. Proc Natl Acad Sci U S A 103, 3863-3868.

Kriegeskorte, N., Goebel, R., Bandettini, P., 2006b. Information-based functional brain mapping. Proc Natl Acad Sci U S A 103, 3863-3868.

Lawley, D.N., 1938. A Generalization of Fisher's z Test. Biometrika 30, 180-187.

Misaki, M., Kim, Y., Bandettini, P.A., Kriegeskorte, N., 2010. Comparison of multivariate classifiers and response normalizations for pattern-information fMRI. Neuroimage 53, 103-118.

Mumford, J.A., Nichols, T., 2009. Simple group fMRI modeling and inference. Neuroimage 47, 1469-1475.

Mur, M., Bandettini, P.A., Kriegeskorte, N., 2009. Revealing representational content with pattern-information fMRI—an introductory guide. Soc Cogn Affect Neurosci 4, 101-109.

Oja, H., Randles, R.H., 2004. Multivariate Nonparametric Tests. Statistical Science 19, 598-605.
Penny, W.D., Trujillo-Barreto, N.J., Friston, K.J., 2005. Bayesian fMRI time series analysis with spatial priors. Neuroimage 24, 350-362.

Pernet, C.R., McAleer, P., Latinus, M., Gorgolewski, K.J., Charest, I., Bestelmeyer, P.E.G., Watson, R.H., Fleming, D., Crabbe, F., Valdes-Sosa, M., Belin, P., 2015. The human voice areas: Spatial organization and inter-individual variability in temporal and extra-temporal cortices. Neuroimage 119, 164-174.

Poldrack, R.A., Farah, M.J., 2015. Progress and challenges in probing the human brain. Nature 526, 371-379.

Ramdas, A., Singh, A., Wasserman, L., 2016. Classification Accuracy as a Proxy for Two Sample Testing. ArXiv e-prints.

Serfling, R.J., 2004. Multivariate Symmetry and Asymmetry. Encyclopedia of Statistical Sciences. John Wiley & Sons, Inc.

Srivastava, M.S., Du, M., 2008. A test for the mean vector with fewer observations than the dimension. Journal of Multivariate Analysis 99, 386-402.

Srivastava, M.S., Katayama, S., Kano, Y., 2013. A two sample test in high dimensional data. Journal of Multivariate Analysis 114, 349-358.

Stelzer, J., Buschmann, T., Lohmann, G., Margulies, D.S., Trampel, R., Turner, R., 2014. Prioritizing spatial accuracy in high-resolution fMRI data using multivariate feature weight mapping. Frontiers in Neuroscience 8, 1-8.

Stelzer, J., Chen, Y., Turner, R., 2013. Statistical inference and multiple testing correction in classification-based multi-voxel pattern analysis (MVPA): Random permutations and cluster size control. Neuroimage 65, 69-82.

Todd, M.T., Nystrom, L.E., Cohen, J.D., 2013. Confounds in multivariate pattern analysis: Theory and rule representation case study. Neuroimage 77, 157-165.

Tong, F., Pratte, M.S., 2012. Decoding Patterns of Human Brain Activity. In: Fiske, S.T., Schacter, D.L., Taylor, S.E. (Eds.), Annual Review of Psychology, Vol 63, pp. 483-509.