# Dependencies between stimuli and spatially independent fMRI sources: Towards brain correlates of natural stimuli

Jarkko Ylipaavalniemi [a,*,1], Eerika Savia [a,b,1], Sanna Malinen [c,d], Riitta Hari [c,d], Ricardo Vigário [a], Samuel Kaski [a,b]

[a] Adaptive Informatics Research Centre, Department of Information and Computer Science, Helsinki University of Technology, P.O. Box 5400, FI-02015 TKK, Finland
[b] Helsinki Institute for Information Technology HIIT, Department of Information and Computer Science, Helsinki University of Technology, P.O. Box 5400, FI-02015 TKK, Finland
[c] Brain Research Unit, Low Temperature Laboratory, Helsinki University of Technology, P.O. Box 5100, FI-02015 TKK, Finland
[d] Advanced Magnetic Imaging Centre, Helsinki University of Technology, P.O. Box 3000, FI-02015 TKK, Finland

## ARTICLE INFO

## ABSTRACT

Natural stimuli are increasingly used in functional magnetic resonance imaging (fMRI) studies to imitate real-life situations. Consequently, challenges are created for novel analysis methods, including new machine-learning tools. With natural stimuli it is no longer feasible to assume single features of the experimental design alone to account for the brain activity. Instead, relevant combinations of rich enough stimulus features could explain the more complex activation patterns.

We propose a novel two-step approach, where independent component analysis is first used to identify spatially independent brain processes, which we refer to as *functional patterns*. As the second step, temporal dependencies between stimuli and functional patterns are detected using canonical correlation analysis. Our proposed method looks for combinations of stimulus features and the corresponding combinations of functional patterns.

This two-step approach was used to analyze measurements from an fMRI study during multi-modal stimulation. The detected complex activation patterns were explained as resulting from interactions of multiple brain processes. Our approach seems promising for analysis of data from studies with natural stimuli.

© 2009 Elsevier Inc. All rights reserved.

## Introduction

The focus of functional magnetic resonance imaging (fMRI) studies has been in rather simple block designs aimed to optimize stimulus control and signal-to-noise ratio. When genuinely natural stimuli are used, block designs are no longer appropriate. The aim of this paper is to take the first step forward by inferring brain correlates of natural stimuli with possibly overlapping stimuli. In such uncontrolled setups, it is extremely difficult to differentiate the stimulus-related processes from ongoing brain activity and, analogously, brain activity-related stimulus properties from all other aspects of the natural environment. The statistical hypotheses are no longer self-evidently derived from the experimental setup, which has conventionally constrained the set of possible hypotheses. Instead, identifying the correct statistical hypotheses is a goal of the analysis in itself. Therefore, we propose a data-driven analysis to define testable research questions.

Purely hypothesis-driven methods have been used extensively in neuroimaging studies. Earlier, the most typical setup of an fMRI experiment has consisted of alternatingly repeating blocks of rest and controlled stimulation, often using only one type of stimuli (see, *e.g.*, Worsley and Friston, 1995). The hypothesis-driven methods are well suited for these setups. However, the situation gets quickly more complicated when one infers correlates of more complex brain processes. Currently, the focus is starting to shift from simple unimodal stimuli towards integration of multiple sensory stimuli to study cognitive processes and, generally, brain activation related to natural stimuli. The recent interest in more real-life-like experimental setups has triggered the first experiments with natural stimuli, such as movies (see, *e.g.*, Hasson et al., 2004; Bartels and Zeki, 2005a; Damoiseaux et al., 2006).

Purely data-driven independent component analysis (ICA, Hyvärinen et al., 2001) is able to separate underlying sources of brain activity, or functional patterns (see, *e.g.*, McKeown et al., 1998; Bartels and Zeki, 2005b). ICA looks for spatially independent patterns of activity without any prior knowledge on the location or temporal dynamics of the activity. It has quickly become a common tool to analyze fMRI data. The majority of the found independent patterns are typically left without an explanation in terms of the stimulus features (see, *e.g.*, Bartels and Zeki, 2005a; Damoiseaux et al., 2006; Malinen et al., 2007).

* Corresponding author.
    E-mail addresses: jarkko.ylipaavalniemi@tkk.fi (J. Ylipaavalniemi), eerika.savia@tkk.fi (E. Savia), sanna@neuro.hut.fi (S. Malinen), riitta.hari@tkk.fi (R. Hari), ricardo.vigario@tkk.fi (R. Vigário), samuel.kaski@tkk.fi (S. Kaski).
  [1] These authors contributed equally to this work.

Some of the unexplained components can actually be unrelated to the stimulation. It is possible, however, to extend the analysis to identify some of the components as stimulus-related, by considering temporal relations between them (see, *e.g.*, Calhoun et al., 2002).

In addition to ICA, other statistical machine-learning methods, such as support vector machines (Vapnik, 1998) and Gaussian-process classifiers (Rasmussen and Williams, 2005), have recently been used to find brain correlates of individual stimulus features (see, *e.g.*, Kamitani and Tong, 2005; Haynes and Rees, 2005). It has even been possible to predict the stimulus features from the brain activity with a trained predictive model (see, *e.g.*, Schneider et al., 2006; Sona et al., 2007). After training, the methods can be very accurate, but they require training data with correct target classes. Thus, they are best suited for repeatable studies where the experimental design readily defines the target classes of interest, such as, visual object categories in a multiple regression analysis.

Regression analysis is based on pre-selected regressors, typically the stimulus time courses, that are used to fit a regression function to explain the response variables, *e.g.*, the brain measurements. In contrast to regression approaches, it is also possible to search for dependencies between stimuli and brain measurements without pre-specifying target classes, by data-driven analysis that is able to find the most relevant target classes automatically. In some suggested approaches the aim has been to search for statistical dependencies between brain activations, inferred from fMRI data and features of the applied stimuli (*cf.*, Friman et al., 2001; Hardoon et al., 2007; Ylipaavalniemi et al., 2007). Even dependencies between fMRI and other measurements, such as EEG, have been studied (Mantini et al., 2007). A classical linear method capable of detecting such dependencies is canonical correlation analysis (CCA, see, *e.g.*, Timm, 2002). These approaches have, however, been hampered by the huge dimensionality of the fMRI data in contrast to the dimensionality of the available stimulus data. Without regularization, the methods would find strong correlations with all sorts of noise structures. In many cases the problem has been solved by constraining the methods to manually-selected brain regions of interest. In our approach, ICA can be seen as a kind of regularization that imposes a well-justified constraint to the findings.

We propose a two-step approach for finding stimulus-related brain activity. The method assumes information only about the measured brain activity and about the stimuli. In controlled setups, the experiment has been designed so that the stimulus sequence defines how the value of each experimental variable changes. We call these variables *labels*. Additional physical features can be extracted from the stimuli to augment the expressiveness of the designed labels. These additional features are also available in uncontrolled or natural setups. Both brain activity and stimuli contain aspects or parts that are not mutually related, as well as parts that are related. Canonical correlation analysis identifies the correlated parts that we call *functional combinations*. It is particularly attractive that the result is invariant to any aspects of the stimuli that are not reflected in brain activity, and to any brain activity unrelated to the stimuli.

As the first step we search for spatially independent brain regions, which we will call *functional patterns*. We assume them to represent independent subunits of more complex activity. As the second step we look for combinations of the functional patterns that are temporally correlated with combinations of the stimulus features. We assume these functional combinations to represent complex stimulus-related processes involving several functional patterns. As a result, each functional combination exhibits temporal correlation between its spatially independent patterns and, on the other hand, the different combinations are uncorrelated with each other.

We target answering the following three questions: First of all, are we able to identify reasonable functional patterns describing brain activity with ICA? Second, do the features extracted from the stimuli explain the found functional combinations well, or even better than the original labels of the experimental design? Finally, and most importantly, do the functional patterns bind to reasonable combinations as a result of our analysis?

## Materials and methods

The analysis consisted of two steps. First, the data were decomposed into spatially independent sources with a variant of ICA, see Independent component analysis section. In the second step, the temporal interactions of the sources were modeled with the help of stimulus labels and features using CCA. The approach is illustrated schematically in Fig. 1; for details see Canonical correlation analysis section.

### Imaging and preprocessing

The analyzed fMRI data are the same as used in a previous publication by Malinen et al., (2007). In that study, the data were successfully analyzed by means of ICA that identified spatial brain activation patterns related to various aspects of the stimuli. Six healthy young adults participated in two identical sessions, in which they received a continuous 8-min sequence comprising of auditory, visual and tactile stimuli in blocks of 6–33 s. The stimuli of different senses never overlapped.

Whole-head volumes were acquired with a Signa VH/i 3.0 T MRI scanner (General Electric, Milwaukee, WI) using a gradient EPI sequence (TR = 3 s, TE = 32 ms, FOV = 20 cm, flip = 90°, $64 \times 64 \times 44$ voxels with resolution $3 \times 3 \times 3$ mm$^3$). In each session, 165 volumes were recorded with the 4 first time points excluded from further analysis. For further details, see Malinen et al., (2007). Preprocessing of the fMRI data included realignment, normalization with skull stripping, and smoothing. For additional details on the measurements and applied preprocessing, see Malinen et al., (2007).

In this study, the measurements of all subjects and all trials were further normalized and temporally concatenated, resulting in single set with $6 \times 2 \times 161 = 1932$ volumes.

### Labels and features of the stimuli

The stimulation sequence consisted of seven different types of stimuli, summarized in Table 1.

In contrast to traditional artificial setups, for natural stimuli the designed labels of the experimental setup may not be expressive enough. We therefore augmented the labels by extracting six features, as an example, from the spectrogram of the actual auditory stimuli. Hence, we had time courses for altogether 13 stimulus labels and features. The spectrogram was calculated using time windows that matched the fMRI acquisition times. The acoustic features we extracted from the spectrogram are commonly used, *e.g.*, in speech recognition (see, *e.g.*, ISO, 2002; Zolnay et al., 2007). Table 2 describes the six extracted acoustic features in terms of qualitative sound properties. Due to the nature of the stimuli, some of the extracted features share a strong resemblance with the original labels; compare for example tones and kurtosis. However, since kurtosis is a physical feature based on the sound signal, it is not identical with any of the designed labels. Including the extracted features brings in more information about the stimuli and, potentially, makes more refined matching with brain processes possible.

Time courses of the stimulus labels were taken as binary-valued functions, whereas the extracted acoustic features had continuous-valued time courses. All label and feature time courses were normalized to the range 0–1 and convolved with a standard hemodynamic response function (HRF) using default latency values of SPM2 toolbox
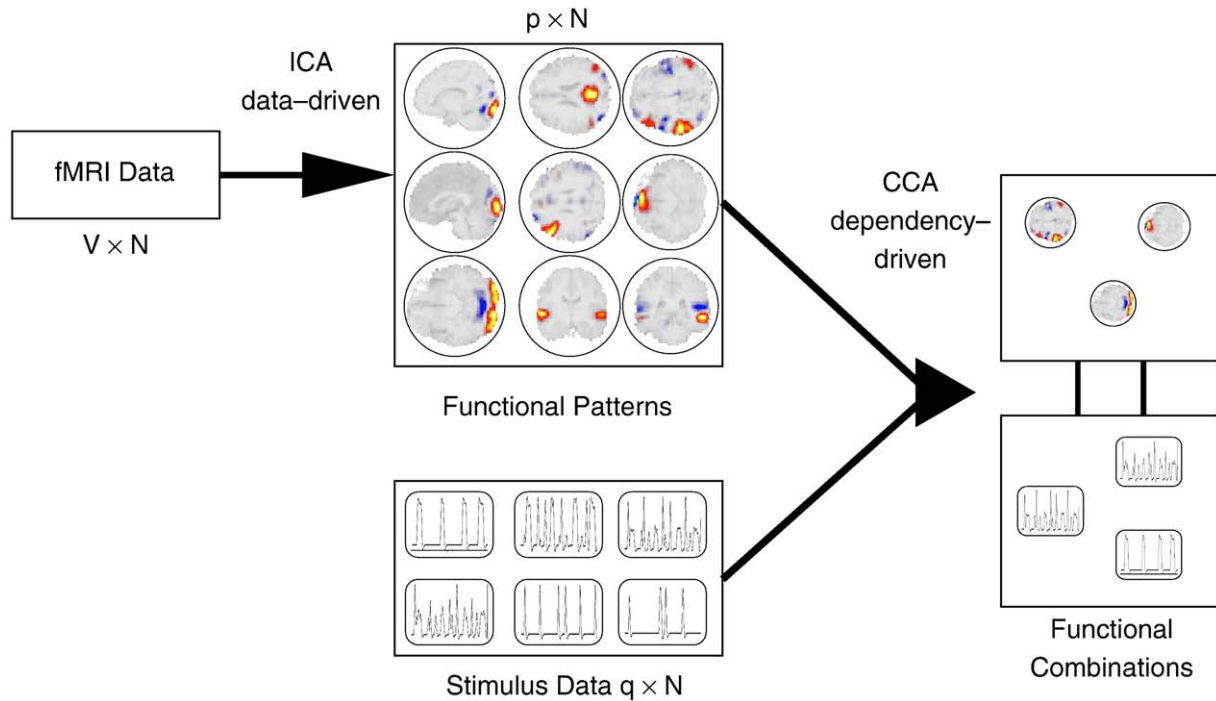
**Fig. 1.** Illustration of the analysis framework. As the first step, ICA is applied to the fMRI measurements ($V$ = number of voxels, $N$ = number of measurement time points) to find spatially independent patterns of brain activity, which we call functional patterns ($p$ = number of reliable ICA components). The second step uses CCA to identify functional combinations based on the temporal dynamics of both the stimuli and the ICA components ($q$ = number of stimulus time courses).

(SPM2, 2002). Fig. 2 shows the resulting time courses of the original labels of the experiment and extracted features of the auditory parts.

*Independent component analysis*

Independent component analysis (ICA, Hyvärinen et al., 2001) is one of the most popular methods for solving the blind source separation (BSS) problem in a purely data-driven manner. BSS consists of finding solutions to the mixture $\mathbf{X} = \mathbf{A}S$, when only the observed data $\mathbf{X}$ are known. ICA assumes only statistical independence of the sources $\mathbf{S}$ and full rank of the mixing matrix $\mathbf{A}$. In fMRI analysis, independence is typically considered in the spatial domain (McKeown et al., 1998), and mixing reveals the temporal dynamics of the identified independent patterns.

We analyzed the fMRI measurements with a reliable ICA method (Ylipaavalniemi and Vigário, 2008), available in the Matlab toolbox

Arabica (Ylipaavalniemi and Soppela, 2009). The inherent stochasticity of the ICA algorithm leads to variability between ICA runs. Arabica takes into account the algorithmic variability and the uncertainty of the sources in the data by performing multiple runs of ICA. Components that are consistent across several runs are considered reliable. Their mean time courses, and the corresponding source volumes, are retained for the second step. We refer to these reliable components as functional patterns. The functional patterns were also ranked according to their consistency over the trials. Fig. 3 shows an example of such an independent component (IC) and its time course. In the rest of the figures we will omit the time courses of the ICs and show only the spatial activation patterns. See the online Supplementary material for all the time courses.

Parameter values for Arabica were derived from testing several alternative values to avoid under- and overfitting the model parameters. The method requires more degrees of freedom than regular ICA and good performance is typically seen with a whitened space having twice the number of dimensions than ICs. In each run, 38 ICs were sought in a 76-dimensional whitened space. The dimensionality of 76 was chosen to

**Table 1**
Labels derived from the original stimulus sequence.

| Sense Label | Description |
| --- | --- |
| Tactile | |
| Touch | 4-Hz tactile pulses delivered bilaterally to the 2nd–4th fingers of both hands in a random order |
| Auditory | |
| Tones | Binaural 100-ms tones with 5 different pitches presented in a random order |
| History | A male voice reading a chapter on the history of the local university |
| Instruction | The same voice reading about guitar fingering for beginners |
| Visual | |
| Faces | Video clips containing close-ups of people's faces |
| Hands | Video clips containing close-ups of people's hands |
| Buildings | Video clips with slow movement through outdoor scenes |

For the corresponding time courses, see panel a in Fig. 2.

**Table 2**
Features extracted from the spectrum of the auditory stimuli.

| Sense Feature | Description |
| --- | --- |
| Auditory | |
| Amplitude | Total sound energy |
| Mean | Centroid pitch of the sound |
| Standard deviation (std) | Spread of sound energy across all frequencies |
| Skewness | Concentration of sound energy to lower frequencies |
| Kurtosis | Concentration of sound energy to only a few frequencies |
| Spectrum derivative feature (sdf) | Amount of increase or decrease of total sound energy |

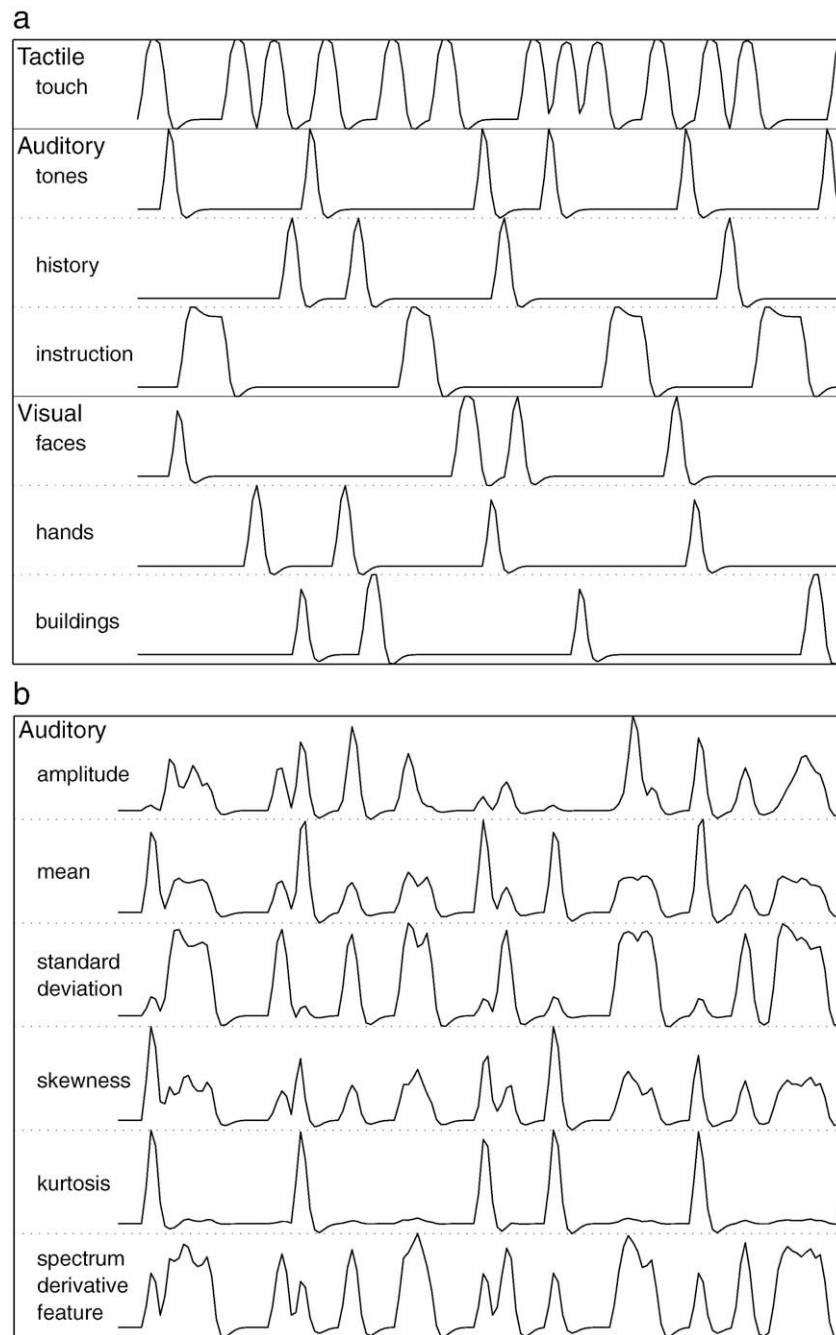For the corresponding time courses, see panel b in Fig. 2.

**Fig. 2.** Time courses of the original labels and the extracted stimulus features. In panel a, labels derived from the original stimulus sequence and, in panel b, features extracted from the spectrograms of the auditory stimuli.

match the automatically estimated value used in the original study (Malinen et al., 2007). The final results were computed in 100 runs using *tanh* nonlinearity in the symmetric mode (FastICA, 1998), producing altogether 64 different ICs. Finally, the 40 most replicable components across the three trials were retained. All the above mentioned dimensionalities resulted from criteria within the Arabica approach, explained in (Ylipaavalniemi and Vigário, 2008).

*Canonical correlation analysis*

In this setup, we have two multi-dimensional datasets, the functional patterns and the stimulus sequence data. They both have data values for the same time points, in other words, the datasets are paired through their time series. Our goal is to find the underlying

dependencies between the datasets by analyzing the common variation of the paired datasets. Naturally, both datasets contain variation that is not shared, and our aim is to distinguish between the shared and the unshared variation. A classical method for this task is the canonical correlation analysis (CCA[2], see Timm, 2002), which maximizes the mutual information for normally-distributed data (Kullback, 1997). With a sufficiently large sample size, CCA is quite robust to violations of the normality assumption (Stevens, 2001).

In brief, the method tries to find such linear projection vectors $w_A$ and $w_Y$ for the two data sets **A** and **Y** that the correlation $\rho$ between

---

[2] The abbreviation CCA can have different meanings in other context, *e.g.*, cross-correlation analysis, curvilinear component analysis.
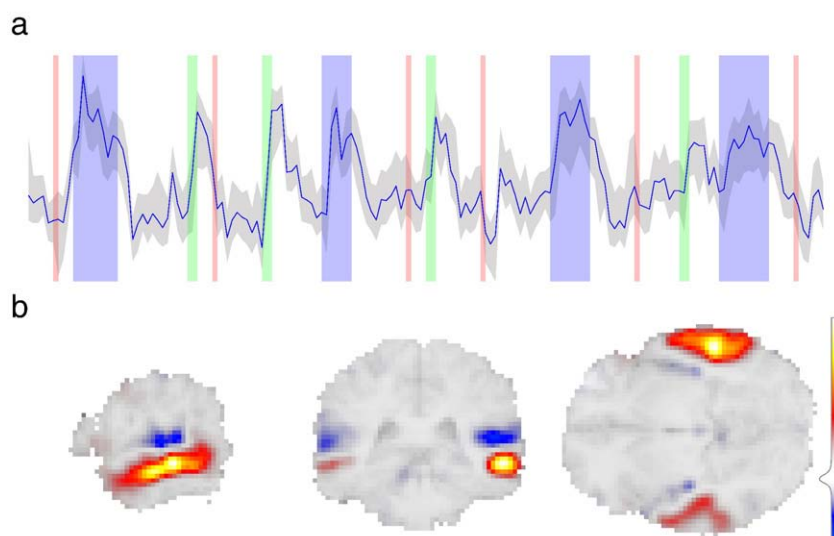
**Fig. 3.** An illustrative example of one independent component (IC). Panel a shows the mean time course of the IC, averaged across the two trials of all six subjects. The mean time courses of the IC's were used as input for CCA, together with the stimulation time courses of Fig. 2. The gray band around the trace shows the 95% confidence interval given by the Arabica approach. The three auditory stimulus labels are shown as colored blocks behind the time course (red = tones, blue = instruction, green = history). The changes in the IC do not exactly match any of the labels although the activity of the IC is correlated with them. Panel b illustrates sagittal, coronal and axial slices of the average brain, centered on the most active voxel of the superimposed IC. The bar on the right shows the used color range based on the z-score; the upper end of the scale depicts positive weights and the lower end negative weights. The left edge of the bar shows the shape of the distribution of the weight values. The more non-Gaussian the shape is, the more independent the IC is.
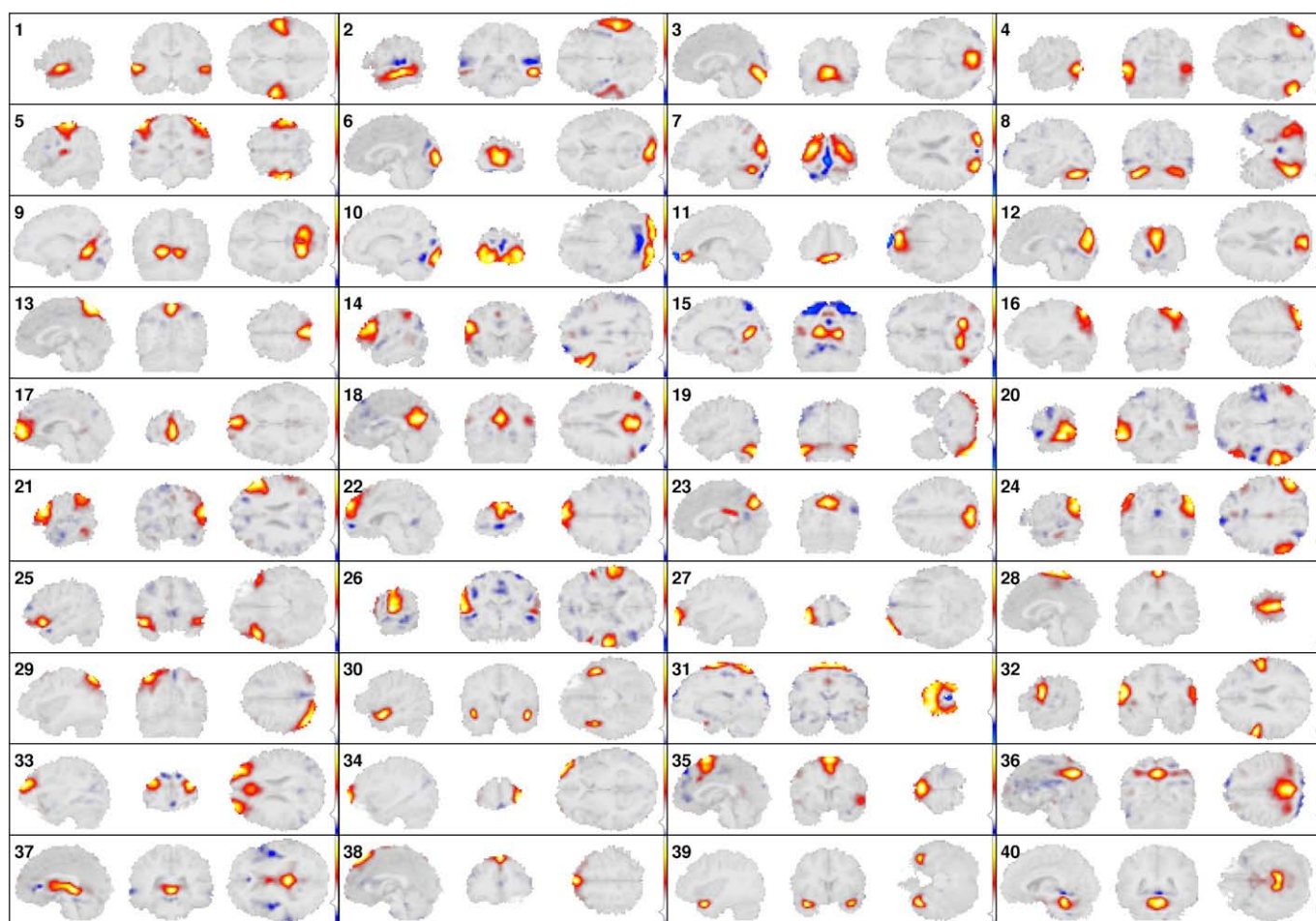


**Fig. 4.** Overview of 40 independent components representing spatially independent functional brain regions. The index number of each component is shown in the upper left corner of each panel, indicating the reliability ranking of the component. See caption of Fig. 3 for other details. For a complete version, including the corresponding time courses, see the online Supplementary material.

the projected data is maximized; the target correlation is defined as follows:

$$\rho = \frac{\mathbf{w}_A^T \left(\mathbf{A} - \mathbf{um}_A^T\right)^T \left(\mathbf{Y} - \mathbf{um}_Y^T\right) \mathbf{w}_Y}{||(\mathbf{A} - \mathbf{um}_A^T)\mathbf{w}_A|| \ ||(\mathbf{Y} - \mathbf{um}_Y^T)\mathbf{w}_Y||}. \quad (1)$$

Here $\mathbf{m}_A$ and $\mathbf{m}_Y$ denote the mean vectors of the data sets, and $\mathbf{u}$ is a column vector of ones with a matching number of elements. In our case $\mathbf{A}$ would correspond to the ICA mixing matrix and $Y$ to the stimulus time series. The maximization leads to an eigenvalue
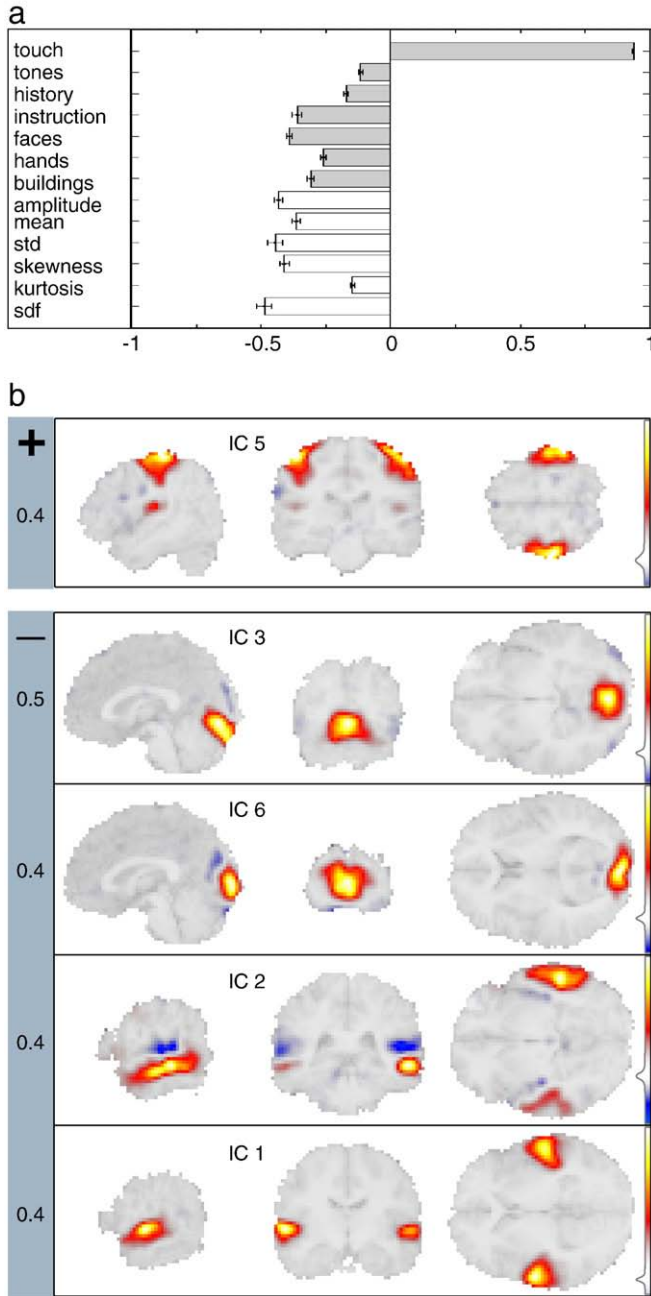


Fig. 5. The 1st functional combination (CCA component) with canonical correlation $\rho = 0.82$. Panel a shows the factor loadings for the stimuli. Panel b depicts the corresponding set of functional patterns (ICA components) with their factor loadings in decreasing order. The shown factor loading values have been rounded. Five functional patterns (ICs 5, 3, 6, 2 and 1) are the most responsible for the observed correlation in this functional combination. Positive factor loadings in panel a correspond to positive factor loadings in panel b. Analogous correspondence holds for negative factor loadings.
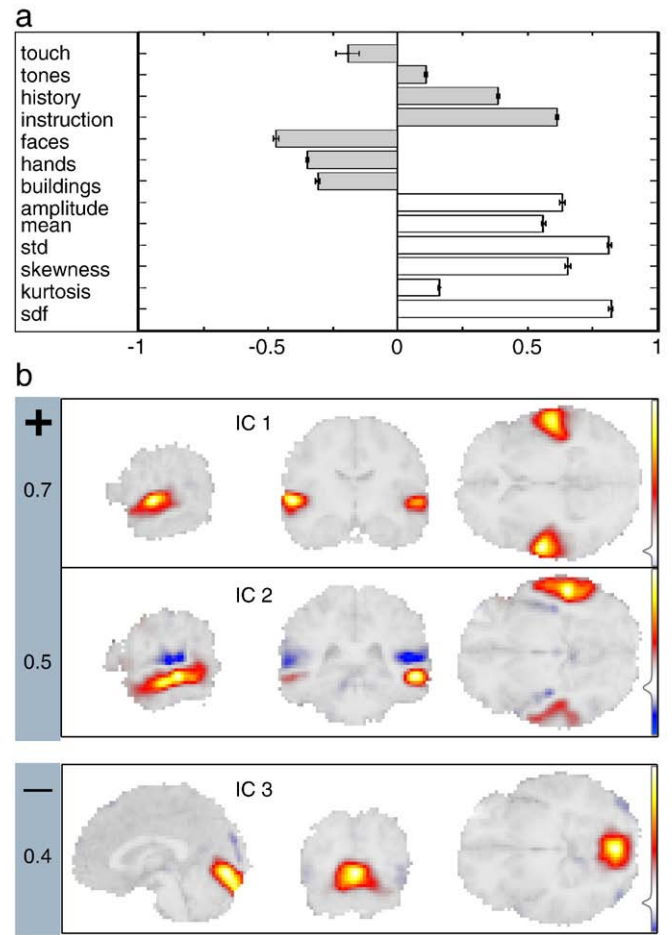


Fig. 6. The 2nd functional combination with the highest canonical correlation $\rho = 0.87$. Other details as in Fig. 5.

problem, which implies that successive results can be retrieved and they can be ordered according to how much of the between-data-set variation they explain. The resulting canonical components are orthogonal in the space where the within-data-set correlations have been removed.

The CCA components are conventionally interpreted in terms of their *factor loadings*, that is, the correlations between the CCA component and each of the original dimensions. In our analysis the original dimensions are the functional patterns and the stimulus labels and features.

### Reliability of ICA and CCA

Theoretically, the number of components found by both ICA and CCA is limited by the smallest dimensionality of the datasets, in our case 1932 for ICA and 13 for CCA. Some weak components may be indistinguishable from noise because of limited amount data. We will assess the replicability of both ICA and CCA using bootstrap techniques (Efron and Tibshirani, 1994).

For Arabica, the bootstrapping has been defined to use 100% sampling, correlation threshold of 0.85, and link power of 4; see Ylipaavalniemi and Vigário (2008) for further details on the parameters of the reliable ICA approach.

The variability of CCA was controlled with bootstrapping using 50% sampling. The result figures show the 99% confidence intervals over 1000 bootstrap iterations. Five out of all 13 CCA components (*i.e.*, functional combinations) were replicable in bootstrap, and were retained for further analysis.

## Results

Arabica decomposed the measurements into 40 components that are sorted in Fig. 4 according to their consistency across trials. The index numbers refer to the ranking by Arabica. As expected, the majority of the individual functional patterns were not directly related to the stimuli. The functional patterns agree well with previous applications of ICA, and with the results reported by Malinen et al. (2007). The components represent different spatially independent patterns of brain processing, including, *e.g.*, the temporal-lobe auditory cortex (IC 1), superior temporal sulcus (STS; IC 2), and several functional patterns on the sensorimotor regions (*e.g.*, ICs 5, 26 and 32), the visual cortices (*e.g.*, ICs 3, 4, 6, 8 and 10), and the frontal lobe (*e.g.*, ICs 14, 17, 21, 22 and 25).

CCA found five reliable functional combinations between the 40 functional patterns and the 13 stimulus labels and features. Figs. 5–9
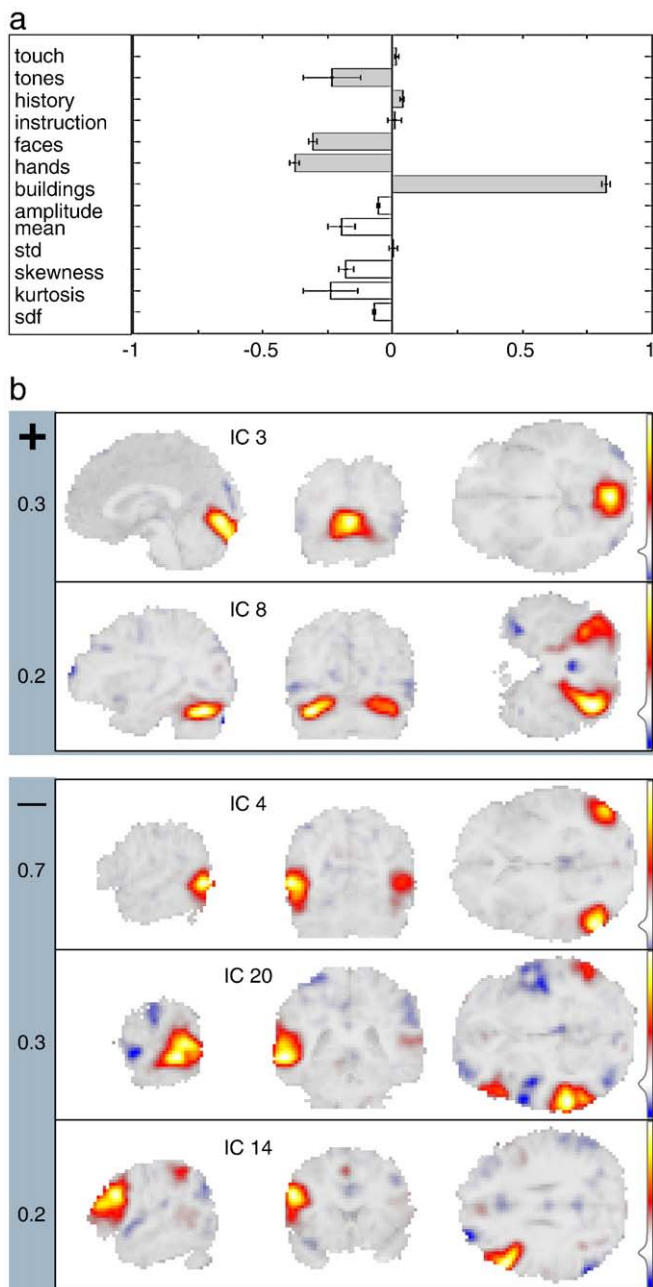


**Fig. 7.** The 3rd functional combination with canonical correlation $\rho = 0.66$. Other details as in Fig. 5.
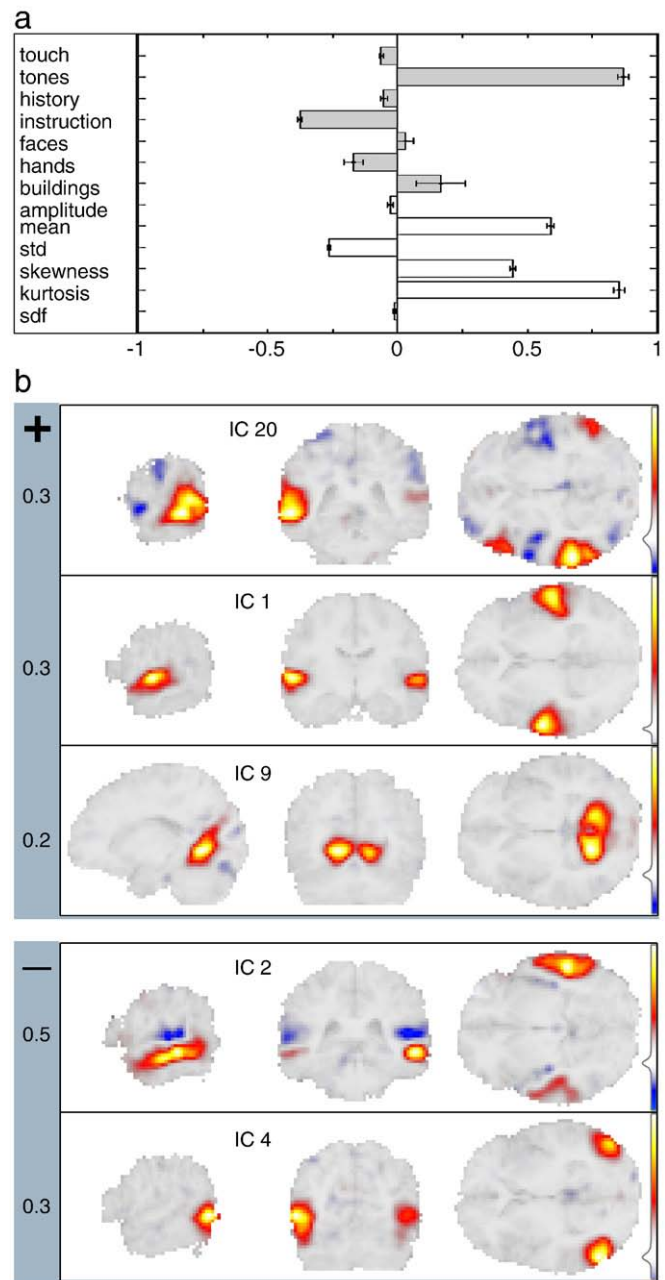


**Fig. 8.** The 4th functional combination with canonical correlation $\rho = 0.58$. Other details as in Fig. 5.

visualize the combinations. In panel a of each figure, the gray bars illustrate the experimental labels and the white bars the extracted stimulus features. The error bars depict the 99% confidence intervals from the reliability analysis. Panel b of each figure shows the set of independent components that build 85% of the total canonical correlation with the stimulus features. The ICs are grouped according to the sign of their factor loadings. In all figures, positive factor loading(s) of the stimulus features correspond to positive factor loading(s) of the ICs, and vice versa for negative factor loadings.

The factor loadings in Fig. 5 discriminate the touch label from all the other labels and features. Thus, this CCA component does not reveal a novel combination of functional patterns, but a discrimination of a single label from all the others. The negatively signed set of functional patterns (ICs 3, 6, 2 and 1) represents all auditory and visual stimuli combined together. The only positively correlated component (IC 5) shows activity in the hand representation area of the
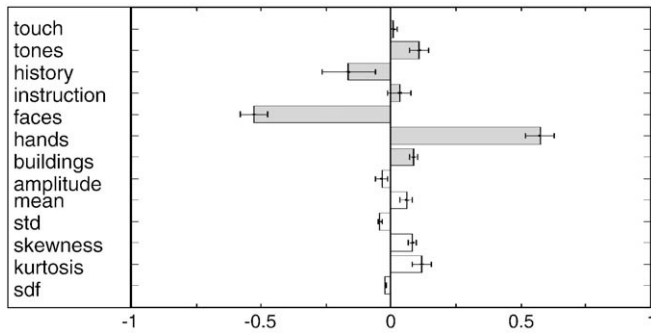
**Fig. 9.** The 5th functional combination with canonical correlation $\rho = 0.45$. The factor loadings for the stimuli. The corresponding 11 ICs are given in the online Supplementary material.

sensorimotor cortex, with weaker activation just above the Sylvian fissure, likely in the region of the second somatosensory cortex SII.

Fig. 6 shows the functional combination with the highest canonical correlation. On the stimulus side, all the auditory labels and features have become grouped together, all with positive factor loadings, in contrast to labels from tactile and visual stimuli that all have negative factor loadings. Correspondingly, there is a strong positive factor loading for the auditory cortex (IC 1) and STS (IC 2), against the negative factor loading for the visual cortex (IC 3). Additionally, the STS component (IC 2) includes negative weights for a region just above the Sylvian fissure, near the region in the sensorimotor component (IC 5) shown in Fig. 5.

The next functional combination, shown in Fig. 7, discriminates between buildings, with positive factor loading, and other visual labels, with negative factor loadings. Some of the auditory labels and features have small negative loadings, but they have large error bars. The corresponding functional patterns include two with positive factor loading in the posterior visual cortices, likely including the primary visual cortex (IC 3) and the fusiform gyri (IC 8). Strong negative loading is seen for IC 4 that coincides with the V5/MT region. Less strong negative loading exists for IC 20.

The fourth functional combination, shown in Fig. 8, discriminates between two auditory labels: tones and verbal instructions. The contributing functional patterns coincide with activation posterior of the Sylvian fissure (IC 20), near the Wernicke's region, and the supratemporal auditory cortex (IC 1) with a positive loading. The STS component (IC 2) has a negative factor loading in this combination. The combination also groups the visual images of hands together with the auditory instruction for guitar fingering. This cross-modal dependency is also supported by the functional pattern V5/MT region (IC 4) with a negative factor loading that could result from brain activity related to movement of hands.

The last functional combination, shown in Fig. 9, represents a clear discrimination between the visual labels hands and faces. Other labels and features have small factor loadings and large error bars. In this case, there were altogether 11 contributing functional patterns, six with positive factor loadings (ICs 26, 16, 10, 21, 3 and 7) and five with negative factor loading (ICs 2, 15, 9, 12 and 6). The factor loadings varied between $-0.2$ and $0.3$, which can be considered weak dependencies. Therefore, speculations of the roles of the individual ICs would not help to understand the whole combination. The contributing functional patterns are shown in the online Supplementary material.

## Discussion

We aim to develop methods for studies of brain activity in more complex experimental settings, where several stimuli overlap. In real-life-like experimental conditions and with natural stimuli, the conventional stimulus labels may need to be replaced by physical or other measurable features that can be derived from the stimuli, such

as the extracted acoustic features we used in this work. Our aim was to find patterns of the fMRI signal explainable by features of the stimuli, both by the original labels and by the less salient features extracted from the stimuli.

We proposed a two-step approach to find related combinations of functional patterns of brain activity and stimulus labels or features. We first used ICA to detect functional patterns of brain activity and then CCA to search for combinations of the individual patterns. We found combinations comprising several ICs, which could support involvement of widely distributed stimulus-related neuronal circuitries. The functional relevance of the combinations can be inferred from the constituent functional patterns (ICs). Furthermore, the functional patterns also have their individual interpretations regardless of the functional combinations.

### Interpretational considerations

As shown in the original study (Malinen et al. 2007), it is possible to find ICs that discriminate between stimulation of different senses and even find details within one sensory modality, such as speech vs. other auditory stimuli. All the ICs discussed in the original paper (1 tactile, 2 auditory, 5 visual and 7 stimulus-unrelated) have highly similar counterparts among our 40 ICs, see Table 3. The main novel contribution in our method is that it automatically identifies the functional combinations responsible for the dependencies between the stimuli and the functional patterns.

We found several different types of functional combinations, all in agreement with the findings of the original study. Some combinations segregated sensory modalities (auditory or tactile) or represented intra-modal structure (buildings vs. the other visual labels; faces vs. hands). Finally, some combinations highlighted possible cross-modal dependencies, for example auditory instruction for guitar fingering and videos of hands that both were related to hands or hand actions, either imagined or real.

One combination focused on distinguishing a single stimulus label (touch) from all the others. However, as will be discussed in more detail in the following section, a stimulus sequence that includes mutually exclusive stimuli can force all other labels to have a negative correlation with the time course of one label. With our current measurements, it is therefore impossible to rule out the possibility that the difference effect is caused, or strongly enforced, by the stimulus design.

### Methodological considerations

The 40 functional patterns agree well with previous applications of ICA, as well as with the results reported by Malinen et al. (2007). The

**Table 3**

The correspondence between originally reported ICs and our reliable components.

| Original component | Our IC |
| --- | --- |
| ACx | IC 1 |
| STS | IC 2 |
| V1/V2 central (lf) | IC 6 |
| V1/V2 central (uf) | IC 3 |
| V1/V2 peripheral | IC 9 |
| Posterior convexial | IC 10 |
| V5/MT | IC 4 |
| SI | IC 5 |
| Fig. 7a) | IC 18 |
| Fig. 7b) | IC 29 |
| Fig. 7c) | IC 32 |
| Fig. 7d) | IC 30 |
| Fig. 7e) | IC 26 |
| Fig. 7f) | IC 21 |
| Fig. 7g) | IC 14 |

The "original components" refer to Figs. 3, 4, 6, and 7 in Malinen et al. (2007) and the component names there-in, and they are grouped accordingly.

components represent different spatially independent patterns of brain activity, and the activity is concentrated in focal areas. As expected, the majority of the individual functional patterns were not directly related to the stimuli. Some of the ICs (*e.g.*, 2 and 5) contained a spatial activation pattern resembling a combination of several regions of activation with positive or negative weights. In our framework this is not a problem, since it only means that even ICA alone is sometimes able to combine brain activations having strong temporal correlation.

Our extracted acoustic features corroborated with the stimulus labels. Whenever the experimental auditory labels became grouped together, the extracted acoustic features became grouped with them. Moreover, when a functional combination discriminated between the different auditory labels (*e.g.*, tones vs. instruction), each original label was grouped together with the related acoustic features. Finally, when the factor loadings of the original auditory labels were close to zero, the extracted acoustic features behaved accordingly.

We noticed that the factor loadings of some of the extracted features were higher than those of the stimulus labels, suggesting that they did match the true brain activity better than the original labels. The extracted features might also help to explain the underlying causes for the observed groupings of labels. In this study, we focused on acoustic features and otherwise the original labels were taken as given. Ultimately, we would naturally want to do the same treatment to all sensory modalities. As a feasibility check we repeated the analysis without the augmented audio features for comparison, and found the same five reliable combinations. In conclusion, the results support the idea of using features extracted from the stimuli as surrogates of experimental labels in natural settings, when descriptive labels are not available.

Our primary question was whether our approach can find reasonable combinations of functional patterns and corresponding combinations of stimulus labels and features. The functional combination related to touch (Fig. 5) shows that when the observed brain activity can be represented by a single stimulus label, our method behaves similarly as the conventional analysis methods that use this single feature as a reference (*e.g.*, generalized linear model in SPM2, 2002). The other results show that whenever a combination is available that explains the observed data better than a single label, the method identifies it.

Thus our method constructs combinations of stimulus labels and features, as well as combinations of functional patterns that together are more strongly correlated than any one of them alone. In conclusion, the observed functional combinations seem to be well-justified, and encourage to develop the approach further.

### Effects of experimental design

Under strict laboratory control, it is often possible to design the experiment so that all relevant combinations of experimental variables (stimuli and tasks) are presented well balanced and in adequate quantities. However, in natural settings we cannot rely on controlled designs but instead the data analysis has to do the balancing. However, no data analysis can insert useful information to the data, and thus balancing can only be done within the given variation in brain processes triggered by the stimuli. Hence, adequate design of the experimental setups to include rich enough stimuli is important in natural settings, too.

The CCA analysis, which is invariant to linear transformations of the stimulus features, is able to compensate for unbalanced or imperfect choice and encoding of stimulus features. In principle then, if the stimuli were rich enough to include all relevant combinations of the unavailable but desirable experimental variables, the analysis would be able to compensate for imbalances. We will next discuss what happens if the stimuli are not rich enough in this sense; it may be worth emphasizing that the consequences

would be similar for any method used for analyzing brain activity during natural stimulation.

If the actual trigger for the measured brain activity cannot be represented with the stimulus labels or features, the brain correlates can be missed or misinterpreted in terms of the available features. However, if the set of stimulus features is expressive enough, our method could find a novel kind of combination of the stimuli that would give a hint of the missing label.

When the experimental design binds certain stimuli to always co-occur, it is not possible to distinguish the corresponding brain correlates from each other. This ambiguity is not a misinterpretation, but rather a straightforward consequence of the selected level of detail of the experimental design. However, true dependencies between the labels are also possible, *e.g.*, some of the extracted auditory features naturally co-occur during any auditory stimulation.

Somewhat surprisingly, similar confusion may happen with stimuli that never co-occur. When the intention is to design uncorrelated labels by creating a stimulation sequence without any overlap, the regularity of the setup can lead to unexpected negative correlations between labels. Such correlations that reflect the experimental design cannot be distinguished from true negative correlations between the stimuli. Thus one might find functional combinations that represent characteristics of the experimental design, rather than of the observed brain responses. More generally, the effects of spurious correlations seem to be an emerging topic of scientific interest (see, *e.g.*, Aguirre et al., 1998; Murphy et al., 2008).

In our data, such a dependency is visible and it naturally affects the statistical analysis. A possible example in our current results is the functional combination, where tactile stimulation had a negative correlation with both visual and auditory stimulation. In the analysis, touch-related brain activity was negatively correlated with brain signals related to other senses. The same phenomenon might be expressed in the functional combination, where all auditory features correlate negatively with both visual and touch labels.[3]

Such correlations can be avoided to some extent in the experimental design by delivering the stimuli in some blocks simultaneously and in some blocks separately. See the online Supplementary material for elaborated discussion with an example of the phenomenon of spurious correlations arising from the experimental design.

### Current limitations and future improvements

Both the strengths and limitations of our two-step approach stem from the assumptions of spatial independence of functional patterns, as well as covariation between stimuli and the functional patterns. Spatially dependent patterns of brain activity cannot be separated from each other by ICA and hence the same holds for the new approach as well. In addition, if some phenomenon does not express itself as a correlation between stimulus features and functional patterns, we cannot find it with our method. The same limitation applies to any other method relying on correlation between experimental labels and observed brain activity. Stimulus-unrelated activity patterns are a crucial part of studying resting-state networks (for a review, see, Raichle and Mintun, 2006).

The canonical correlation analysis implicitly assumes normally-distributed data. Especially if the data are very far from normally-distributed, improvements can be made by developing more advanced dependency exploration methods.

---

[3] The current experimental design is unable to differentiate negatively correlated and uncorrelated stimulus blocks of different sensory modalities, since the occurrence of stimuli of any sensory modality was fully determined by the absence of stimuli of all the other senses. See the online Supplementary material for details about the correlation between stimuli to different sensory systems.

## Acknowledgments

## Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2009.03.056.

## References

Aguirre, G.K., Zarahn, E., D'Esposito, M., 1998. The inferential impact of global signal covariates in functional neuroimaging analyses. NeuroImage 8 (3), 302–306.
Bartels, A., Zeki, S., 2005a. Brain dynamics during natural viewing conditions — a new guide for mapping connectivity in vivo. NeuroImage 24 (2), 339–349.
Bartels, A., Zeki, S., 2005b. The chronoarchitecture of the cerebral cortex. NeuroImage 360 (1456), 733–750.
Calhoun, V.D., Pekar, J.J., McGinty, V.B., Adali, T., Watson, T.D., Pearlson, G.D., 2002. Different activation dynamics in multiple neural systems during simulated driving. Hum. Brain Mapp. 16, 158–167.
Damoiseaux, J.S., Rombouts, S.A.R.B., Barkhof, F., Scheltens, P., Stam, C.J., Smith, S.M., Beckmann, C.F., 2006. Consistent resting-state networks across healthy subjects. Proc. Natl. Acad. Sci. 103 (37), 13848–13853.
Efron, B., Tibshirani, R.J., 1994. An Introduction to the Bootstrap, 1st Edition. Vol. 57 of Monographs on Statistics and Applied Probability Series. CRC Press.
FastICA, 1998. MATLAB™package. http://www.cis.hut.fi/research/ica/fastica.
Friman, O., Carlsson, J., Lundberg, P., Borga, M., Knutsson, H., 2001. Detection of neural activity in functional MRI using canonical correlation analysis. Magn. Reson. Med. 45 (2), 323–330.
Hardoon, D.R., Mourão-Miranda, J., Brammer, M., Shawe-Taylor, J., 2007. Unsupervised analysis of fMRI data using kernel canonical correlation. NeuroImage 37 (4), 1250–1259.
Hasson, U., Nir, Y., Levy, I., Fuhrmann, G., Malach, R., 2004. Intersubject synchronization of cortical activity during natural vision. Science 303 (5664), 1634–1640.
Haynes, J.D., Rees, G., 2005. Predicting the orientation of invisible stimuli from activity in human primary visual cortex. Nat. Neurosci. 8 (5), 686–691.
Hyvärinen, A., Karhunen, J., Oja, E., 2001. Independent Component Analysis, 1st Edition. Wiley-Interscience, New York, NY.

ISO, 2002. JTC1/SC29 information technology – multimedia content description interface – part 4: Audio. ISO/IEC FDIS 15938-4. International Standards Organization, Geneva, Switzerland.
Kamitani, Y., Tong, F., 2005. Decoding the visual and subjective contents of the human brain. Nat. Neurosci. 8 (5), 679–685.
Kullback, S., 1997. Information Theory and Statistics, 2nd Edition. Dover Publications, Mineola, NY.
Malinen, S., Hlushchuk, Y., Hari, R., 2007. Towards natural stimulation in fMRI — issues of data analysis. NeuroImage 35 (1), 131–139.
Mantini, D., Perrucci, M.G., Gratta, C.D., Romani, G.L., Corbetta, M., 2007. Electrophysiological signatures of resting state networks in the human brain. Proc. Natl. Acad. Sci. 104 (32), 13170–13175.
McKeown, M.J., Makeig, S., Brown, G.G., Jung, T.P., Kindermann, S.S., Bell, A.J., Sejnowski, T.J., 1998. Analysis of fMRI data by blind separation into independent spatial components. Hum. Brain Mapp. 6 (3), 160–188.
Murphy, K., Birn, R.M., Handwerker, D.A., Jones, T.B., Bandettini, P.A., 2008. The impact of global signal regression on resting state correlations: are anti-correlated networks introduced? NeuroImage In press. doi:10.1016/j.neuroimage.2008.09.036.
Raichle, M.E., Mintun, M.A., 2006. Brain work and brain imaging. Annu. Rev. Neurosci. 39, 449–476.
Rasmussen, C.E., Williams, C.K.I., 2005. Gaussian Processes for Machine Learning, 1st Edition. MIT Press, Cambridge, MA.
Schneider, W., Bartels, A., Formisano, E., Haxby, J., Goebel, R., Mitchell, T., Nichols, T., Siegle, G., 2006. Competition: inferring experience based cognition from fMRI. In: Proceedings of the 12th Meeting of the Organization for Human Brain Mapping (HBM 2006). Florence, Italy, http://www.ebc.pitt.edu/2006/competition.html.
Sona, D., Veeramachaneni, S., Olivetti, E., Avesani, P., 2007. Inferring cognition from fMRI brain images. In: Proceedings of the 17th International Conference on Artificial Neural Networks (ICANN 2007). Porto, Portugal, pp. 869–878.
SPM2, 2002. MATLAB™package. http://www.fil.ion.ucl.ac.uk/spm.
Stevens, J.P., 2001. Applied Multivariate Statistics for the Social Sciences, 4th Edition. Mahwah, NJ, Lawrence Erlbaum Associates.
Timm, N.H., 2002. Applied Multivariate Analysis, 1st Edition. Springer, New York, NY.
Vapnik, V.N., 1998. Statistical Learning Theory, 1st Edition. Wiley-Interscience, New York, NY.
Worsley, K.J., Friston, K.J., 1995. Analysis of fMRI time-series revisited — again. NeuroImage 2 (3), 173–235.
Ylipaavalniemi, J., Vigário, R., 2008. Analyzing consistency of independent components: an fMRI illustration. NeuroImage 39 (1), 169–180.
Ylipaavalniemi, J., Savia, E., Vigário, R., Kaski, S., 2007. Functional elements and networks in fMRI. Proceedings of the 15th European Symposium on Artificial Neural Networks (ESANN 2007). Bruges, Belgium, pp. 561–566.
Ylipaavalniemi, J., Soppela, J., 2009. Arabica: Robust ICA in a Pipeline. Proceedings of the 8th International Conference on Independent Component Analysis and Signal Separation (ICA 2009). Paraty, Brazil, pp. 379–386. http://launchpad.net/arabica.
Zolnay, A., Kocharov, D., Schlüter, R., Ney, H., 2007. Using multiple acoustic feature sets for speech recognition 49 (6), 514–525.