NeuroImage 184 (2019) 993–1004

Contents lists available at ScienceDirect

NeuroImage

journal homepage: www.elsevier.com/locate/neuroimage

# Individual differences in analogical reasoning revealed by multivariate task-based functional brain imaging

Rubi Hammer [a,*], Erick J. Paul [a], Charles H. Hillman [c,d], Arthur F. Kramer [a,c,e], Neal J. Cohen [a,b], Aron K. Barbey [a,b]

[a] Beckman Institute for Advanced Science and Technology, University of Illinois, Urbana-Champaign, IL, USA
[b] Department of Psychology, University of Illinois, Urbana-Champaign, IL, USA
[c] Department of Psychology, Northeastern University, Boston, MA, USA
[d] Department of Health Sciences, Northeastern University, Boston, MA, USA
[e] Department of Mechanical and Industrial Engineering, Northeastern University, Boston, MA, USA

ARTICLE INFO

*Keywords:*
Analogical reasoning
Higher-level cognition
Large-scale brain networks
Relational similarity
Visual cognition

ABSTRACT

Although analogical reasoning (AR) plays a central role in higher-level cognition and constitutes a key source of individual differences in intellectual ability, the neural mechanisms that account for individual differences in AR remain to be well characterized. Here we investigated individual differences in AR within a large sample (n = 229), using multivariate fMRI analysis (a simple multiple kernel learning machine). The individual AR capability was positively correlated with activation level in a prefrontal executive network and a visuospatial network. Notably, the best predictors of individual differences in AR within these networks were activation in the dorsomedial prefrontal cortex (response selection) and the lingual gyrus (visual feature mapping). In contrast, AR capability was negatively correlated with activation in the default mode network. The implications of the reported findings are twofold: (i) Individual differences in AR depend on multiple executive and visuospatial brain regions, where their respective contributions are contingent upon the individuals' cognitive skills; (ii) Brain regions associated with individual differences in AR only partially overlap with brain regions sensitive to the associated task demands (i.e., brain regions sensitive to the analogy relational complexity, at the group-level). We discuss implications of such brain organization supporting AR as an example for brain architecture underlying higher-level cognitive processes.

## 1. Introduction

Analogical reasoning (AR) is a hallmark of human intelligence, enabling the capacity to reason about novel scenarios based on a comparison between objects, or systems of objects, that highlights respects in which those scenarios could be considered as similar. Therefore, by accounting to relational similarity, reasoning by analogy involves identifying higher order patterns and discovering recurrences of such patterns despite apparent variation in the elements that compose them. This capability for relational processing provides the foundation for analogical reasoning and constitutes an essential facet of human intelligence – giving rise to multifaceted categorization, problem solving, creativity and scientific discovery.

The current study had two primary objectives: (i) to investigate if individual differences in AR are explained by activity within a few prefrontal brain regions implicated by the larger part of early literature, or instead can be better explained by a distributed pattern of activity within a large-scale brain network as suggested more recently; and (ii) primarily to assess the relationship between brain regions underlying individual differences in AR and brain regions sensitive to the AR task demand (defined by the analogy complexity, evident as a group-level effect). This enabled to investigate the relation between brain regions associated with context induced cognitive load, (e.g., an increase in AR complexity) at the group level, and brain regions accounting to individual differences in AR capabilities. It may be presumed that brain regions sensitive to group-level effects, as evident when manipulating task demands, also account for individual differences in the associated tasks. However, for reasons detailed below we hypothesized that in higher-level cognitive processes,

* Corresponding author. The Neurocognitive Unit (Xtend Scientific), Segol Center for Hyperbaric Medicine and Research, Yitzhak Shamir Medical Center, Rishon-Le-Zion, 70300, Israel.
*E-mail address:* rubihammer@gmail.com (R. Hammer).

https://doi.org/10.1016/j.neuroimage.2018.09.011
Received 26 April 2017; Received in revised form 16 August 2018; Accepted 5 September 2018
Available online 12 September 2018

such as AR, this is not likely to be the case (see also Supplemental 1 for a simulation exemplifying this point).

Besides being reliant on prior knowledge and semantic retrieval, AR engages (i) feature mapping processes, in which specific features are highlighted, aligned and integrated, and inferences are projected from one analogue to others; and (ii) an evaluation process, in which the resulting analogy and its inferences are examined and judged (Gentner and Smith, 2012; Holyoak, 2012; Krawczyk et al., 2004). The diversity of contexts in which AR is engaged and the fact that multiple cognitive processes are involved in AR have motivated the perspective that AR is not a unitary construct but instead depends on multiple cognitive mechanisms that are recruited based on task demands (e.g., relational complexity) and the individual cognitive proficiencies. For example, (i) larger working memory capacity, and better attentional control capability would enable intricate feature *mapping,* which essentially enable processing a greater number of analogues and analogues' features; and (ii) better response selection capability would be associated with better capability to *evaluate* different analogues and better outcome of the reasoning process (Chuderski et al., 2012; Deary, 2001; Kyllonen and Christal, 1990; Unsworth et al., 2014).

Parallel developments in cognitive neuroscience have advanced our understanding of the neural mechanisms of AR. A seminal fMRI study of AR was conducted by Bunge et al., (2005), who demonstrated that core facets of AR engaged specific cortical regions, with anterior left inferior prefrontal cortex (PFC) underlying retrieval of relevant semantic information, left frontopolar mediating integration demands (associated with feature mapping complexity), and right dorsolateral PFC supporting response selection. Evidence from an fMRI experiment of visuospatial ('non-semantic') AR demonstrated that feature mapping complexity is associated with increased activity within the lateral PFC and lateral frontal pole in both hemispheres. When there was greater need to resolve interference (discarding irrelevant features, as part of the analogy evaluation and response selection process), activation increased only in the lateral PFC (Cho et al., 2010). On the other hand, impaired visuospatial AR was reported in individuals with left rostrolateral prefrontal injury, indicating that this brain region also plays central role in relational integration and feature matching (Urbanski et al., 2016).

The above fMRI literature has primarily addressed the role of regions within PFC, in AR. However, non-PFC regions have been also reported to play a central role in AR. In a whole brain fMRI study of visuospatial AR, Geake and Hansen (2010) found that the left ventrolateral, right dorsolateral and bilateral tempo-parietal cortices were preferentially recruited with greater feature mapping demands (associated with an increase in AR complexity). Employing a visuospatial AR task, Watson and Chatterjee (2012) observed activity within left rostrolateral PFC associated with the integration of relational knowledge during AR. Watson and Chatterjee also found that higher activity within the inferior frontal gyri, evident during visuospatial AR, is associated with greater inhibition and response selection demands (associated with an increase in the number of competing possible analogues), whereas inferior parietal activation is associated with greater respective reliance on spatial relation information. A recent meta-analysis of the fMRI literature demonstrated that a bilateral frontoparietal network is recruited across a broad spectrum of AR tasks. This network included the left rostrolateral PFC, bilateral insula, posterior parietal cortex, few clusters in the posterior region of the inferior frontal gyrus (IFG), middle frontal gyrus (MFG), superior frontal sulcus, and medial PFC. Moreover, contrasted with semantic AR, visuospatial AR was found to be associated with greater activation in the MFG, lateral IFG, angular gyrus, superior and inferior parietal lob (SPL/IPL) and the fusiform gyrus (Hobeika et al., 2016).

In contrast to the ample neuroscience literature that examines sensitivity to task-related demands in AR, at the group-level, very few studies have directly investigated neurocognitive mechanisms underlying individual differences. In a task-based fMRI study, Preusse et al. (2011) compared the visuospatial AR brain activation of a group of individuals scoring high on tests of fluid intelligence relative to that of a group with average fluid intelligence scores. The authors reported greater parietal activation in high fluid intelligence individuals compared to individuals with average fluid intelligence, and a reverse effect in the ventral anterior cingulate cortex and the ventromedial frontal cortex (medial Brodmann Area 10; part of the default mode network). Preusse and colleagues suggested that high fluid intelligence may confer attention control and working memory processes that facilitate visuospatial AR, mediated by parietal cortex. On the other hand, a voxel-based morphometry study showed that visuospatial AR (with stimuli varying in perceptual characteristics such the font type, size, color and orientation) was associated with individual differences in cortical volume within the left rostrolateral PFC (including Broca's area) and the anterior part of the inferolateral temporal cortex (including Wernicke's area). A complementary tactography of diffusion-weighted images further demonstrated that these two regions exchange information via the arcuate fasciculus (Aichelburg et al., 2016).

The limited understanding of the neural basis of individual differences in AR reflect three challenges. First, AR is a multifaceted cognitive process that engages several neural mechanisms that are likely to be difficult to characterize accounting only to a few frontal brain regions. Second, there are only few published neuroimaging studies that have directly investigated individual differences in AR. These do not provide a consistent account of mechanisms underlying individual differences in AR, and, in each case, the sample size was too small to account for multiple possible sources of individual differences in AR. Third, previous neuroimaging studies of AR have primarily employed univariate analyses, which does not enable exploring the possibility that individual differences in AR are best explained by a weighted sum of activation in multiple brain regions. Thus, previous studies could not inform us about the respective contribution of different brain regions to the individual AR capability. Moreover, most earlier AR studies identified brain regions sensitive to an experimental task manipulation, but they do not discuss the nature of interdependencies between brain regions underlying individual differences in AR and brain regions sensitive to the associated AR demands (e.g., the analogy relational complexity, determined by the number of task-relevant features).

To address the above concerns, the present study examined neural mechanisms associated with individual differences in AR (where AR capacity was measured as the participant correctness/accuracy in the AR tasks) by testing a large sample of participants using functional MRI (fMRI). Each participant performed multiple AR trials with high relational complexity (High-RC) and low relational complexity (Low-RC). Whole-brain multivariate analyses (implemented by a multiple kernel learning machine) were used to find the brain network underlying individual differences in each of the two relational complexity conditions separately, and the network of brain regions sensitive to the relational complexity manipulation (group-level, condition effect). The multiple kernel learning machine enabled the execution of a whole-brain multivariate analysis with relatively low computational demands and without relying on any prior assumptions regarding the role of key brain regions in AR (e.g., unlike a seed-based analysis). Effectively, this analysis enabled detecting a sparse brain model, which is based on a subset of brain regions with significant contribution to AR (while also weighting contributions of the individual voxels within those brain regions). To better understand the role of the brain regions identified as taking part in AR, we further assessed the degree of co-activation amongst all those brain regions, sorting them into a few networks of co-activated regions, and assessing how these networks relate to known functional networks in the human brain.

This allowed to conclude which of the following characterizes the neurocognitive mechanisms underlying individual differences in AR: (i) AR primarily relies on a few brain regions (associated with cognitive functions such as feature mapping, attention control, working memory and response selection). These regions, sensitive to group-level effects (measured as the analogy relational complexity), would also underlie evident individual differences in AR. (ii) AR relies on a distributed brain

network where multiple interdependent processes contribute to AR. Each process executes a basic cognitive function and its level of engagement depends either on the individuals' skills, on the associated task demands, or on both. In such a scenario, a whole-brain multivariate analysis would detect a network in which some brain regions are sensitive only to the AR relational complexity, whereas others would contribute to individual differences in AR. While (i) suggests that AR primarily relies on a few brain regions, regardless to task difficulty or the individual competence, (ii) points to the possibility that AR may reflect the application of alternative strategies supported by a network of multiple brain regions. Each region accounts to a more basic cognitive process, where the respective contribution of at least some of those processes to AR substantially differs between individuals. For example, two individuals may have similar AR capabilities. However, one may rely more on her feature mapping capabilities, whereas the other would rely on working memory and attention control.

## 2. Materials and methods

### 2.1. Participants

Participants were recruited from the Urbana-Champaign community as part of a cognitive training intervention study. Participants were right-handed, with normal or corrected to normal vision and no history of neural or psychiatric disorders. The pre-intervention People Pieces Analogy (PPA) experiment included a total of 296 participants. Since the primary objective here was testing individual differences, we used conservative inclusion criteria when accounting for head movements and brain image quality, minimizing irrelevant individual differences in the sample data (see details below). This resulted in exclusion of 66 participants from the analysis. An additional participant was excluded due to very low response rate. The remaining 229 participants included 118 females, and their mean age was 23.4 years (SD = 5.1; range = 18–43). Participants provided informed written consent in accordance with a protocol approved by the University of Illinois Institutional Review Board, and they were paid for participation.

### 2.2. The People Pieces Analogy (PPA) task

Participants performed a modified version of the People Pieces Analogy (PPA) task (for a detailed description see Cho et al., 2010; Sternberg, 1977) in an fMRI scanning session. The session included three rapid event-related fMRI scans, each had 54 trials and was approximately 10 min long (totaling 162 AR trails in ~30 min of functional scans). Order of the three scans and trial-type within scans were counterbalanced across participants. E-Prime® 2.0 (Psychology Software Tools, INC) running on a Windows7 PC was used for stimuli presentation and for the recording of participants' responses. Stimuli were presented on a mirror mounted on the fMRI scanner head coil. Responses were recorded using an MRI compatible response box. Prior to the scanning session, participants had a short practice session outside the scanner.

Each PPA AR problem (trial) consisted of two pairs of cartoon characters (an analogy source pair, and an analogy target pair). Characters within each pair were characterized by four binary feature: color, width, height and gender (a total of $2^4 = 16$ distinct characters). In each trial the participant was required to determine if the relational similarity (analogy) between the 2 pairs was valid (index finger key-press) or invalid (middle finger key-press), accounting for the subset of feature cued to be task-relevant. Specifically, starting at the onset of the trial a written list of the four feature was displayed in black font in the center of the screen to the right of the paired characters used as the analogy source. 1.7 s after the onset of the trial the font color of either one feature word, in low relational complexity (Low-RC) trials, or three features words, in high relational complexity (High-RC) trials, switched to red indicating the task-relevant feature/s. 0.3 s later, a pair of characters used as the analogy target appeared to the right of the features list (Fig. 1 and Supplemental 2; see Cho et al., 2010 for a detailed description of a similar procedure). The analogy source and target were presented together until the participant decided (within a time interval limited to 6 s) if the analogy was valid or not. Total trial duration was 8 s.

The total of 162 trials (from the three scans) included 60 valid analogy Low-RC trials (see example in Fig. 1B) and 60 valid High-RC trials (see example in Fig. 1C). In addition to these, there were 21 invalid analogy trials in each level of relational complexity. In these invalid analogy trials, there was a between-pair relational similarity mismatch on exactly one of the *cued/relevant* feature(s), whereas the between-pair relation in the non-cued feature(s) always matched. The introduction of invalid trials prevented participants from 'automatically' approving the validity of an analogy without paying attention to the cued features. Note that levels of reasoning difficulty in invalid trials cannot be

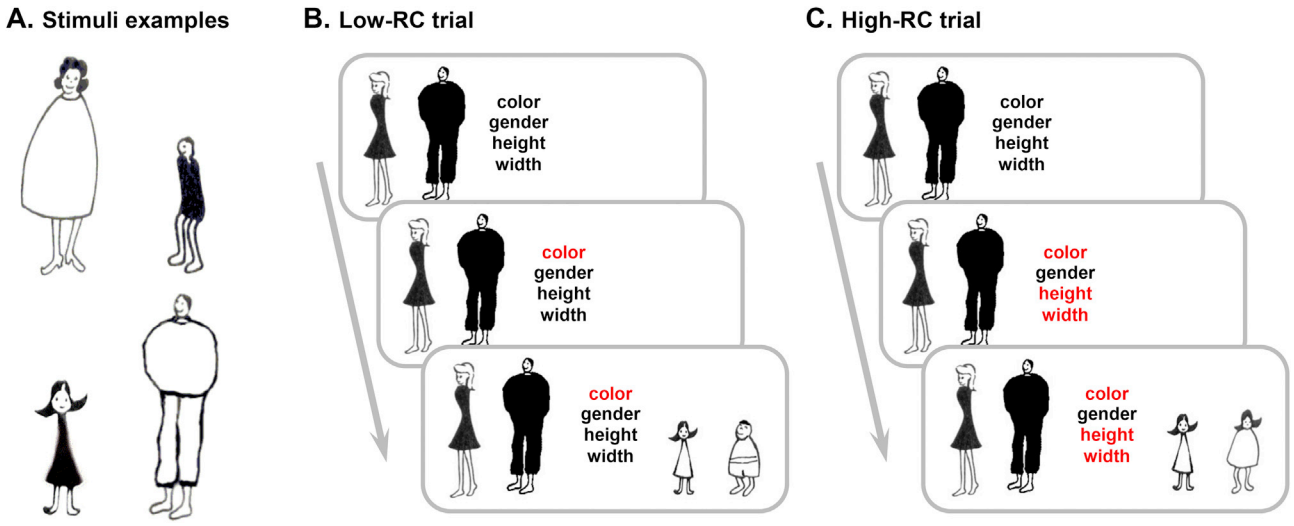


**Fig. 1.** Stimuli examples and trial composition. (A) Four character stimuli examples. Characters could have differed in color, width, height and gender. (B) Low-RC trial where only one feature (color) is cued as relevant. Note that in this example the analogy source (left pair) and target (right pair) share relational similarity in all four feature (valid analogy, with no interfering feature). (C) High-RC trial where three feature (color, height and width) are cued as relevant. Note that in this example the analogy source and target do not share relational similarity in gender (a valid analogy, with an interfering feature). See Supplemental 2 for additional trial examples.

assessed deterministically, primarily since invalid High-RC analogy trials could vary in the extent to which subjects are engaged in relational integration, as they could reject the analogy by simply searching for the nonmatching relation. For this reason, invalid trials were excluded from later analysis. The 1:3 valid-invalid trial ratio used here was like those used in previous studies (e.g., Cho et al., 2010).

In half of the valid analogy Low-RC and High-RC trials there was one interfering feature, an *irrelevant* feature with a relational inconsistency between the source and target (e.g., man/woman vs. woman/woman in Fig. 1C). 20% of the valid trials were "catch" trials, in which the analogy source character pair was omitted 1.7 s after the trial onset. The catch trials were randomly intermixed with the standard trials, and they encouraged participants attending all feature in the analogy source character pair starting at the onset of the trial, prior to the cue onset. The composition of catch trials was otherwise identical to standard trials (with an equal number of Low-RC and High-RC catch trials). Between the offset of one trial and the beginning of the next there was a fixation-only interval with random jittering between 2 and 8 s (in addition to the time left from the response interval following the participant key-press). This design allowed keeping the visuospatial complexity and the participants' engagement in the Low-RC and High-RC trials comparable. Due to the relatively small number of error trials, and to reduce the variability in the modeled brain activation associated with participants' responses, only valid trials in which the participant responded correctly were used in the later fMRI analysis.

### 2.3. Data analysis tools

Data analysis was performed using SPM-8 (Statistical Parametric Mapping, Wellcome Trust Centre for Neuroimaging, London, UK), PRoNTo-2 (Pattern Recognition for Neuroimaging Toolbox, Wellcome Trust Centre for Neuroimaging, London, UK), IBM® SPSS-22, and analysis scripts specifically developed for this project (MathWorks® Matlab-2014b/2015b).

### 2.4. MRI and fMRI data acquisition

Imaging data was acquired on a 3.0 T S 3 T Magnetom Trio scanner using a 32-channel head coil. Gradient echo localizer images were acquired to determine the placement of the functional slices. A susceptibility weighted single-shot EPI (echo planar imaging) method with BOLD (blood oxygenation level-dependent) was used for functional images acquisition with the following scan parameters: TR = 2000 ms, TE = 25 ms, flip angle = 90°, matrix size = 92 × 92, field of view = 230 × 230 × 126 mm, slice thickness = 3 mm (0.316 mm gap), number of slices = 38 (an effective functional voxel size of 2.5 × 2.5 × 3 mm). A total of 313 images (TRs) were recorded in each scan. Slices were acquired in an interleaved manner. The high resolution T1 weighted 3D image was acquired with the following parameters: 0.9 mm isotropic voxels, TR = 1900 ms, TI = 900 ms, TE = 2.32 ms, with GRAPPA and an acceleration factor of 2. The duration of the anatomical scan was 4 min and 26 s.

### 2.5. Image preprocessing

Preprocessing involved: (i) Slice timing; (ii) Realignment of all functional images to the 24th image. (iii) Co-registration of the functional and anatomical images; (iv) Normalization of the T1 image to the MNI305 template image. Linear and non-linear normalization parameters were then applied to the functional images. (v) 5 × 5 × 6 mm full width half maximum (FWHM) Gaussian kernel smoothing. (vi) We confirmed that movement was kept below 3 mm (in any of the x, y, or z dimensions) within a scan using ArtRepair (an SPM add-on; Mazaika et al., 2009). Outlier images (up to 1% of the images in a scan) were realigned by ArtRepair, interpolating between the two adjacent non-outlier images. Participants with more extensive or more frequent head movements were excluded. In the subsequent general linear model (GLM) analysis, noisy images were deweighted. To reduce variability in neural activity within each experimental condition, trials in which the participant responded correctly were modeled separately from error trials or trials in which the participant did not respond on time, with onset time-locked to the beginning of each trial and event offset determined by the participant response timing (Cho et al., 2010; Hammer et al., 2015a, 2015b). (vii) A high pass filter with a cut-off of 132 s.

### 2.6. fMRI modeling and image rescaling

Trials were modeled into seven basic event-types, based on the trial characteristics and the participant response: (i) valid Low-RC without interference; (ii) valid Low-RC with interference; (iii) valid High-RC without interference; (iv) valid High-RC with interference; (v) invalid Low-RC; (vi) invalid High-RC; (vii) all trials in which the participant made an error (or did not responded on time). For the regression and classification analyses, we computed the mean of (i) and (ii) (mean brain image, based on all the trials in which the participant correctly identified valid Low-RC analogies), and the mean of (iii) and (iv) (mean brain image, based on all the trials in which the participant correctly identified valid High-RC analogies). (v), (vi) and (vii) were omitted from later analysis. Beta values in the Low-RC and High-RC images of each participant were rescaled into z-scores (rescaling was done across all the voxels within each functional brain image). These transformed images are comparable to t-maps where each voxel is contrasted with the entire brain image global mean (see Supplemental 5 for univariate analyses that were based on the subjects' Beta maps without rescaling). The rescaled images were used as the input of the Simple-MKL machine learning algorithms (see below).

### 2.7. MRI/fMRI image quality control and head movements tolerance

In addition to the automated artifact detection and correction (using ArtRepair), brain images were visually inspected. Participants with evident image artifacts (e.g. brain image cutoff or orbital and temporal signal loss) were excluded from the analysis. We also excluded participants with head displacement larger than the size of a voxel, during a single scan, in any of the translational axes (i.e. 3 mm), or participants for which more than 1% of the TRs had to be replaced by the ArtRepair motion correction algorithm.

### 2.8. Simple multiple kernel learning (MKL) regression and classification machines

To identify the brain network associated with AR we used the simple multiple kernel learning (MKL) regression and classification algorithms. These machine learning algorithms are based on the gradient decent of the support vector machine (SVM) objective value, and they share with the SVM primary characteristics of maximum margin algorithms (Chapelle et al., 1999; Lanckriet et al., 2004): (i) usage of kernels; (ii) absence of local minima; (iii) a capacity to control sparseness, obtained by acting on the decision hyperplane margin (via a regulating hyperparamater); (iv) learning a non-linear function by a linear learning machine mapping into high dimensional kernel induced feature space. The primary principle underlying MKL is that an optimized model that is based on a linear combination of multiple basis kernels, where each basis kernel is based on a subset of features (i.e., voxels within a single brain parcel), can approximate the optimized model based on a single kernel computed by accounting for all features (i.e., all the voxels in the brain). The optimized MKL model can be a convex combination of only part of the basis kernels, and thus it may account to only a subset of the input features. Specifically, the Simple-MKL uses weighted L2-norm regularization for computing each basis kernel (i.e., weighting the contribution of voxels within each parcel), where the model sparsity at the whole brain level is controlled by a L1-norm constraint on the basis kernels sum

of weights (i.e., the sum of weights of brain parcels). The algorithm iteratively determines the optimal combination of kernels by a gradient descent (convex optimization) wrapping a standard simple-SVM solver (see details in Rakotomamonjy et al., 2008; Xu et al., 2010). These properties make the Simple-MKL an efficient large-scale regularization algorithm in scenarios where the initial number of features is similar to, or greater than, the number of the handled data points (Tuia et al., 2010).

In whole brain fMRI data analysis, the typical initial number of features (tens of or even hundreds of thousands voxels) is a few orders of magnitude greater than the typical number of the data points (i.e., the number of participants' brain images that had to be classified; few hundreds at best). Here we used a brain atlas with 400 brain parcels, predetermined based on functional connectivity (Craddock et al., 2012), to parcellate the participants' brains into parcels of functionally and spatially associated voxels (subsets of likely highly correlated features). Applying a gray-matter mask, excluding ventricles and white matter voxels (Hammer et al., 2015b; Qureshi et al., 2017), resulted in reducing the number of effective parcels (with primarily gray-matter voxels) to 291, out of which 241 had a volume of at least 50 functional voxels. The Simple MKL regression and classification machines were then used to compute the basis kernel for each parcel using L2-norm regularization (weighting voxels within each parcel). The weight of each parcel in the regression/classification models was computed using a L1-norm regularization, constraining the sum of the absolute weights to 1. This resulted in sparse regression/classification models based on a relatively few parcels (due to the L1-norm regularization), where the respective weight of each parcel in each model was determined based on a relatively large portion of its voxels (due to the L2-norm regularization). This made the discovered models more interpretable and substantially reduced the odds for discovering overfitted models, as the independent variables impacting the model fit are the non-zero weight base-kernels, instead of the lower-level features (that is, the non-zero weight parcels, instead of all the voxels in the brain).

The Simple-MKL requires setting of the SVM hyperparameter $C$, balancing between minimal error on the training sample and the size of the margin of the SVM decision hyperplane (greater margin reduces the likelihood for overfitting). The selection of the hyperparameter $C$ impacts the sparsity of the discovered model, where larger $C$ values are associated with less sparsity, lower error rate on the training sample, but smaller margin and higher likelihood for overfitting. Here hyperparameter optimization involved a first pass search using hyperparameter values between $10^{-5}$–$10^{3}$ with logarithmic steps $\{10^{-5}, 10^{-4}, \ldots 10^{2}, 10^{3}\}$, followed by a second pass of random search around the optimal value discovered in the first pass (within ±50% of the first pass optimal value; Wainer and Cawley, 2017). Performances of the reported below regression and classification models were similar across a broad range of hyperparameter values (all three models performed best with hyperparameter values between $10^{-3}$–1). Model performance for each hyperparameter value were assessed using 5-fold cross-validation (in each fold the algorithm was trained using 80% of the sample and tested on the remaining 20%). The model with minimal cross-validation error was then selected. Note that the model learning and hyperparameter optimization were independent (for each tested model the hyperparameter optimization was done independently using an additional 5-fold cross validation in a nested inner loop; Wainer and Cawley, 2017).

We used the Simple-MKL regression machine to find models associated with individual differences in AR, in the High-RC and Low-RC conditions separately. Input for these two regression analyses were the rescaled High-RC and Low-RC brain images (mean activation images, based on all the valid trials in which the participant answered correctly, in each of the two relational complexity conditions) as the independent variables (predictors), and the respective participants' performance as the dependent variable (predicted). We used the Simple-MKL classification machine to find a model differentiating between the High-RC and Low-RC conditions (differentiating between the two types of brain images by identifying brain regions sensitive to relational complexity; group-level effect). Here the dependent (predicted) variable was the image label (High-RC or Low-RC). That is, both the regression and classification implementations of the Simple-MKL were based on the gradient decent of the support vector machine (SVM) objective value. However, for the regression machine the predicted variable was on a continues ratio-scale (participants' AR capability), whereas for the classification machine it was nominal/categorical (High-RC and Low-RC images labeled as +1 and −1).

### 2.9. Multidimensional scaling (MDS) and k-means clustering

The Simple-MKL machines enabled identifying the network of brain parcels associated with individual differences in AR (the regression models; individual AR capacity measured as task accuracy), and the network of parcels showing sensitivity to relational complexity (the classification model). Furthermore, the parcels weights indicated the respective contribution of each parcel to the regression/classification model. However, these models do not show, directly, how activation in one parcel relates to activation in others. Accordingly, we conducted additional analysis, considering between-parcel similarities in patterns of activation. From each parcel, we extracted the mean activation (as rescaled Beta value) for each participant in the High-RC and Low-RC conditions (separately), accounting only for activation in voxels with a weight $\geq 0.01$ in at least one of the regression/classification models discovered by the Simple-MKL machines (see Supplemental 3). This resulted in each parcel being characterized by 458 data points (229 participants; two conditions). To simplify the presentation of similarities in activation patterns between parcels, we used non-metric multidimensional scaling (MDS), minimizing the stress cost function while keeping the number of the MDS dimensions small (see Results). Unlike metric MDS, the non-metric MDS finds a non-parametric monotonic relationship between the dissimilarities in the item-item matrix (here it is a parcel-to-parcel similarity matrix). This is done in addition to the computation of the Euclidean distances between items and the location of each item in the low-dimensional space. That is, the scaling into lower dimension space was determined by the data, instead of a few parameters of a hypothesized model. $k$-means clustering was then applied to classify parcels, in the MDS space, into few networks of co-activated parcels.

## 3. Results

### 3.1. Behavioral performance

We found a significant (Pearson) correlation between accuracy in the High-RC and Low-RC conditions, $r(229) = 0.61$, $p < 0.0001$ (two-tailed; Fig. 2A). An analysis of variance shows that the mean AR accuracy in the Low-RC condition was significantly higher than the mean accuracy in the High-RC condition, $F(1, 228) = 152.26$, $p < 0.0001$, $\eta_p^2 = 0.40$ (Fig. 2B). In the Low-RC condition there was significant negative correlation between accuracy and reaction time $r(229) = -0.25$, $p < 0.0002$ (two-tailed), whereas in the High-RC condition there was significant positive correlation between accuracy and reaction time $r(229) = 0.13$, $p < 0.05$ (two-tailed). That is, higher competence in Low-RC trials was evident as better accuracy and shorter reaction time, whereas in High-RC trials there was a speed-accuracy tradeoff.

### 3.2. Brain-behavior High-RC Simple-MKL regression model

Using the Simple-MKL machine for finding the brain-behavior regression model for the High-RC data resulted in a model with $r(229) = 0.37$ ($R^2 = 0.137$; predictive/cross-validated variance explained), with mean squared error (MSE) = 0.94, estimated based on 5-fold cross-validation (i.e., in each fold the machine was trained using 80% of the data and was tested on the remaining 20%; Fig. 3A). A permutation test with 200 permutations (randomly pairing the brain image

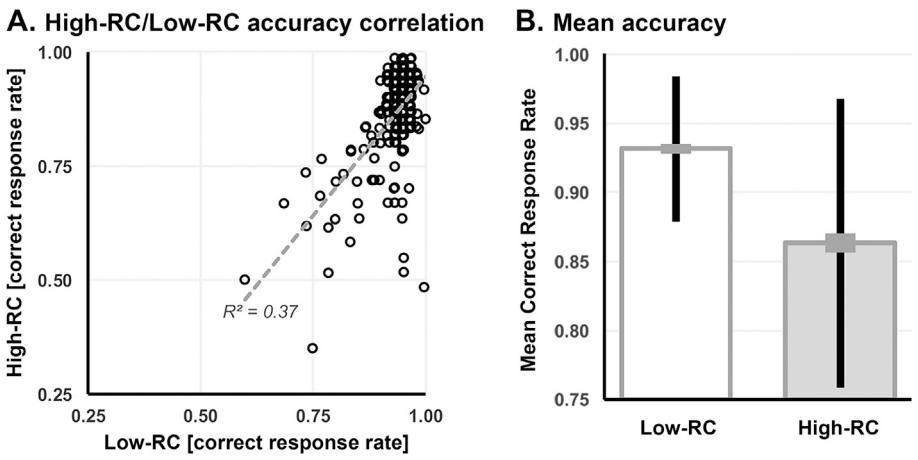


**Fig. 2.** Participants' behavioral performance (AR accuracy). (A) Correlation between AR accuracy in the Low-RC and High-RC conditions. Each data point represents a single participant (for this presentation, overlapping data points were slightly jittered). (B) Mean AR accuracy in the Low-RC and High-RC conditions. Thick gray bars represent SE and thin black bars represent SD.

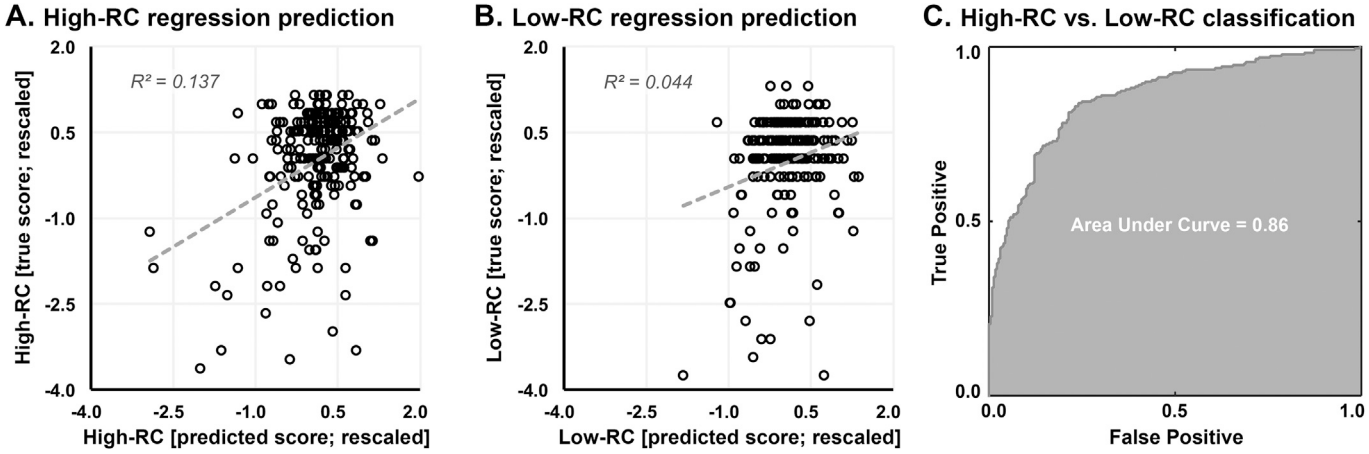


**Fig. 3.** Predictive power of the regression and classification models, each estimated based on 5-fold cross-validation (see also Supplemental 3). (A) Predictions in the High-RC regression model. (B) Predictions in the Low-RC regression model. In the two regression models' scatter plots (Panels A and B), each data point represents a single participant. The vertical axis represents the participant's actual AR accuracy and the horizontal axis represents the participant's AR accuracy as it was predicted by the Simple-MKL machine based on the participant's pattern of brain activation in the respective relational complexity level. (C) Receiver Operating Curve (ROC) presenting the overall cross-validation accuracy in the High-RC versus Low-RC classification model (the capability of the Simple-MKL machine to determine if a brain image represent activation in the High-RC or Low-RC tasks).

of each participant with the AR accuracy score of another) shows that a model with such predictive power is unlikely to be discovered by chance, with $p < 0.005$ (reflecting the probability finding by chance a model with $R^2 \geq 0.137$ or MSE $\leq 0.94$).

The significance of each brain parcel discovered in the previous step was further assessed by computing the parcel mean absolute weight and the standard error of the mean, based on the parcel's five weights computed in the five iterations (folds) of the cross validation. One-tailed $z$-test was used to decide if a parcel is with greater than chance-level weight, with a chance-level threshold $= 0.004 = 1/241$ (241 was the number of parcels with more than 50 functional voxels). False discovery rate (FDR) correction was computed by accounting for parcels with a mean absolute weight $\geq 0.004$, and by assuming positive dependence (Benjamini, and Yekutieli, 2001). Significance of parcel contribution was determined based on $p < 0.01$ (FDR corrected). The 15 parcels with significant weights in the High-RC regression model are listed in Table 1 (see weight map in Fig. 4A and absolute respective weights in Supplemental 3A). This analysis indicates the left lingual/-cuneus, the bilateral medial superior frontal gyrus (BA8) and the left lingual parcels as part of the primary contributors to the High-RC individual differences model.

### 3.3. Brain-behavior Low-RC Simple-MKL regression model

Using the Simple-MKL machine for finding the brain-behavior regression model for the Low-RC data resulted in a model with $r(229) = 0.21$ ($R^2 = 0.044$; predictive/cross-validated variance explained), $p = 0.03$; MSE $= 0.94$, $p = 0.015$ (significance determined based on 200 permutations). Nine brain parcels were found to have significant weights in the Low-RC regression model with $p < 0.01$ (FDR corrected). See Figs. 3B and 4B, Table 1, and Supplemental 3B. This analysis indicates that the right lingual/cuneus, the bilateral medial superior frontal gyrus (BA8), and the left lingual parcels to be part of the primary contributors to the Low-RC individual differences model.

### 3.4. High-RC versus Low-RC Simple-MKL classification model

Cross-validation accuracy in the Simple-MKL High-RC versus Low-RC classification model was 79.5%, $p < 0.005$ (significance determined based on 200 permutations). The classification accuracy was balanced, with 78.6% accuracy in classifying Low-RC images, and 80.3% accuracy in classifying High-RC images. Eighteen brain parcels were found to have significant weights in the High-RC versus Low-RC classification model at

**Table 1**
Listing the brain parcels associated with AR. parcels sorted by the rightmost column, indicating the parcel's network (VS = Visuospatial, Ex = Executive, Def = Default, Sal = Salience; in brackets R/G/B/P indicate color labels as appeared in Fig. 5). (+) indicates either positive activation-behavior correlation or High-RC > Low-RC, whereas (−) indicates the reveres. See also Supplemental 7 for a model summary.

| Brain parcel | High-RC | Low-RC | High–Low | Network |
|---|---|---|---|---|
| B-Cuneus | | | (+) | VS (R) |
| R-Cuneus | | | (+) | VS (R) |
| L-Lingual/Cuneus | (+) | (+) | (+) | VS (R) |
| R-Lingual/Cuneus | (+) | (+) | (+) | VS (R) |
| L-Lingual | (+) | (+) | (+) | VS (R) |
| R-Lingual | (+) | (+) | (+) | VS (R) |
| B-Precuneus | | (+) | (+) | VS (R) |
| R-Precuneus | (+) | | | VS (R) |
| L-MidOccipital | (+) | | | VS (R) |
| R-FFG | (+) | | | VS (R) |
| L-FFG | (+) | | | Ex (G) |
| L-SPL | (+) | | (+) | Ex (G) |
| L-IFG(BA44) | (+) | | (+) | Ex (G) |
| R-IFG | | (+) | | Ex (G) |
| L-IFG(BA46) | (+) | | | Ex (G) |
| L-MFG | (+) | | | Ex (G) |
| B-MedFG/ACC(BA8) | (+) | (+) | (+) | Ex (G) |
| B-CaudateHead | | | (+) | Def (B) |
| R-STG | | | (−) | Def (B) |
| B-Paracentral | | | (−) | Def (B) |
| B-PCC | | | (−) | Def (B) |
| L-ACC | (−) | | | Def (B) |
| L-MedFG/ACC(BA10) | (−) | | | Def (B) |
| B-MedFG/ACC(BA10) | | | (−) | Def (B) |
| L-Insula | | | (−) | Sal (P) |
| L-PostInsula | | (+) | | Sal (P) |
| R-Insula | | (−) | | Sal (P) |
| L-Postcentral | | | (−) | Sal (P) |
| R-MidLOC | | | (−) | Sal (P) |

$p < 0.01$ (FDR corrected). See Figs. 3C and 4C, Table 1, and Supplemental 3C. This analysis indicates the left postcentral, the right superior temporal gyrus, and the left insula parcels to be part of the primary contributors to the High-RC vs. Low-RC classification model. Conducting the classification analysis using the non-rescaled Beta maps yielded similar results, with cross validation accuracy of 82.1% (classification accuracy of 79.5% for Low-RC and 84.7% for High-RC; AUC = 0.88) and with a similar weight map.

### 3.5. Networks of co-activated brain parcels

To understand the function of a brain region it is useful to identify other brain regions functionally associated with it, besides understanding the conditions that induce changes in activation within this region. Accordingly, we show how brain parcels relate to one another by accounting for activation levels (as mean Beta values; see Methods for more details) in both relational complexity conditions, and across all participants. First, in Fig. 5A we present the similarities in patterns of activation between the brain parcels in a three-dimensional space such that it preserves the between-parcel similarities in beta values as much as possible (computed using MDS ALSCAL method; data for each brain parcel were a 458-long vector, based on mean activation of each participant, in both Low-RC and High-RC). This multidimensional scaling (MDS) model preserves 71.7% of the variance in the original data, with a stress cost function value = 0.25. Reducing the number of dimensions from three to two resulted in a 25% drop in the explained variance, whereas increasing the number of dimensions to four resulted in only marginal increase of less than 5% in the explained variance (as it also compromises the clarity of the model representation). With few exceptions, on the left side of d1 (horizontal axis) are dorsal brain regions whereas ventral brain regions are on the right; lower-level brain functions are at the bottom of d2 (vertical axis) and higher-level functions are at the top; and in d3 (color) brain regions associated with the default network are with shades of blue (low values), executive brain regions are with shades of green (intermediate values), and visual brain regions are with shades of red (high values).

In Fig. 5B (and 5C) we present the results of a *k*-means clustering (with $k = 4$) of the brain parcels based on their respective proximities in the MDS space. This clustering analysis resulted in one cluster comprising lower-level visuospatial parcels (color-labeled red), a cluster comprising executive parcels (green), a cluster comprising default network parcels (blue), and a cluster comprising salience attention network parcels (purple). The total number of brain regions to be clustered (29) constrained the maximal number of clusters ($k$). On the other hand, using less than four clusters compromised the interpretability of the observed networks. Using $k = 2$ resulted in one cluster that was positively associated with AR (included all of the visuospatial and executive brain regions) and another that was negatively associated with AR (included most of the default and salience brain regions, to exclude the caudate head and the posterior insula). Using $k = 3$ resulted in part of the salience network brain regions clustered together with the executive and visuospatial brain regions, and part clustered with the default network (see also Table 1).

## 4. Discussion

We employed task-based fMRI to study neurocognitive mechanisms associated with individual differences in analogical reasoning (AR). A whole brain multivariate analysis resulted in a novel finding, showing that only part of the brain regions associated with individual differences in AR (measured as the accuracy in AR tasks) capabilities were also sensitive to the analogy relational complexity. Overall, we identified 29 brain parcels that were significantly associated with AR. Using a clustering algorithm, we sorted these into four networks, each with a handful of co-activated parcels: One network comprised primarily dorsomedial/ventrolateral prefrontal brain regions; a second network comprised visuospatial brain regions including the cuneus, lingual, precuneus, middle occipital and right FFG; a third network comprised a default mode network parcels including the ventromedial MedFG/ACC (BA10), PCC, paracentral and right STG; a fourth network comprised salience attention network parcels including the left and right insula. Furthermore, the current findings demonstrate how individual differences in higher-level cognition can be better explained by accounting to the weighed contribution of multiple brain regions, where respective contributions of those regions rely on the associated task demands.

We demonstrated how the Simple-MKL machine (but possibly other whole-brain multivariate analysis procedures with large-scale regularization, at the brain parcel level) can be effective in identifying and characterizing neurocognitive mechanisms underlying higher-level cognitive processes. To better interpret the Simple-MKL machine results, we identified brain regions with significant contribution to AR accounting to both the mean weight and the weight variance across all folds in the Simple-MKL machine cross-validation procedures (see Supplemental 3). This enabled to discard from the reported AR model brain regions with respective large weights that nevertheless likely reflect noise, and on the other hand to identify brain regions with apparent small weights that are in fact significantly associated with AR. These supplementary analyses may be affected by the number of folds (where large number of folds may increase the odds for overfitted model, as well as the odds for a type-1 error when identifying brain regions with significant weights). Therefore, we employed a conservative approach, keeping a small number of folds in the Simple-MKL machine cross-validation procedures and by using a conservative statistical threshold ($p < 0.01$, FDR corrected, Cohen's $d > 0.5$). For a related discussion regarding methods for improving the interoperability of multivariate (backward) models see Haufe et al., 2014).

### 4.1. AR and dorsomedial/ventrolateral prefrontal (central executive) activation

Higher relational complexity (group-level effect) and better

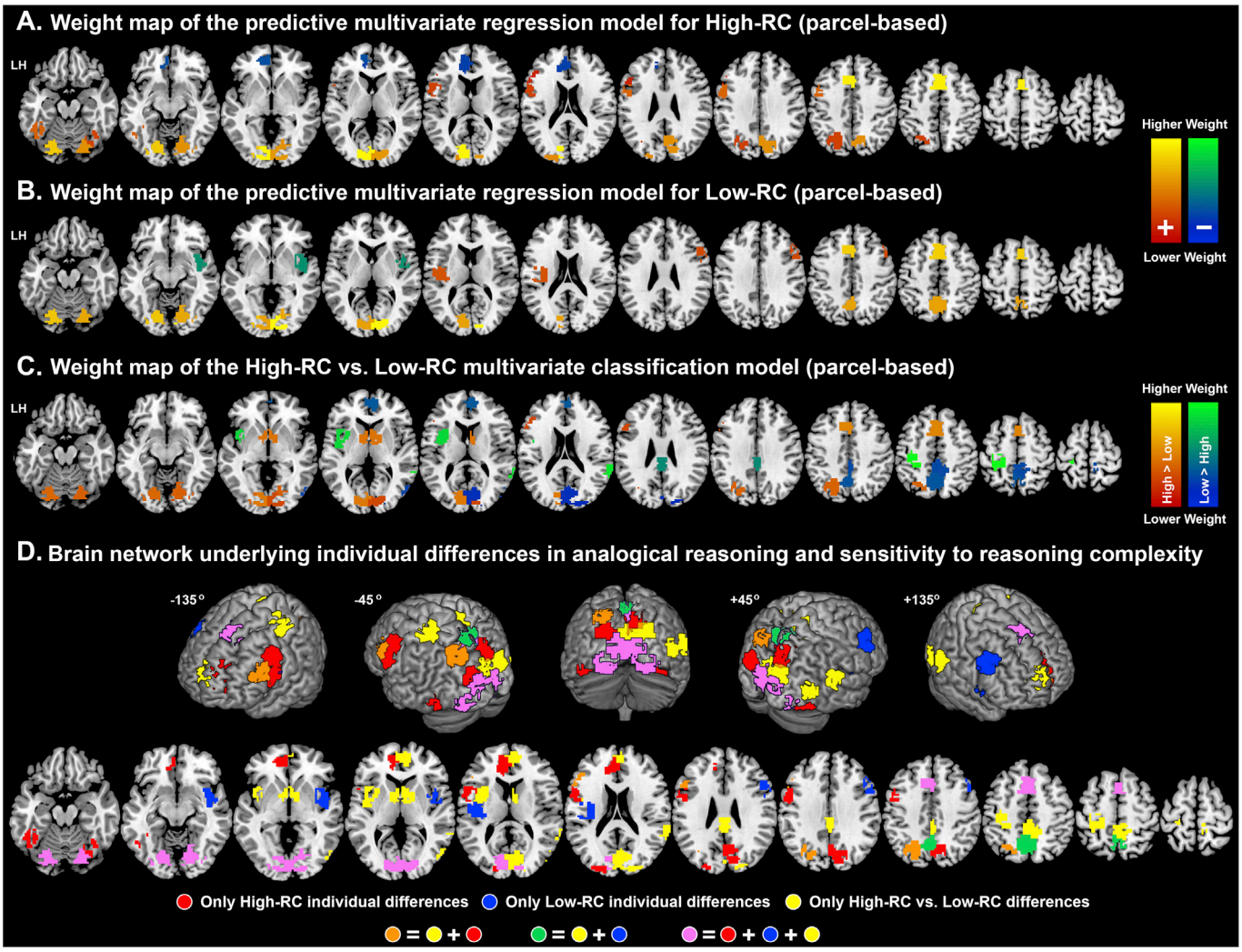


**Fig. 4.** Brain maps. (A) High-RC regression model weight map (brain parcel level). Yellow indicates higher weight with positive correlation with AR accuracy. Red indicates lower weight with positive correlation. Green indicates higher weight with negative correlation. Blue indicates lower weight with negative correlation. (B) Low-RC regression model weight map (parcel level). (C) High-RC versus Low-RC classification model weight map (parcel level). Warm colors indicate greater activation in High-RC compared to Low-RC. Cold colors indicate greater activation in Low-RC compared to High-RC. (D) Color-labeled parcels, indicating parcels taking part in either regression model or the classification model, or in a combination of few models. See Supplemental 1 for examples of respective simulated scenarios. See Supplemental 4 for details about the Simple-MKL voxel weighting within each parcel. See Supplemental 5 for corresponding univariate t-maps.

individual AR capability were both associated with higher dorsomedial and ventrolateral prefrontal activation. In all these brain parcels (see Table 1), activation level was positively correlated with individual differences in the High-RC condition. However, only the dorsomedial B-MedFG/ACC (BA8) and left IFG (BA44) also showed significant sensitivity to relational complexity (High-RC > Low-RC), where the dorsomedial B-MedFG/ACC was the only executive brain region associated with the individual AR capability in both relational complexity levels (Fig. 4, Table 1 and Supplemental 3). Early studies showed that better response selection, specifically in demanding cognitive tasks, is associated with higher neural activation in the dorsomedial frontal and anterior cingulate cortices (de la Vega et al., 2016; Platt and Huettel, 2008). Higher activation in the dorsomedial B-MedFG/ACC was associated with either better AR (individual differences) or with greater relational complexity (task demands). We suggest that this brain region plays a central role in evaluating the analogy validity towards final stages of response selection.

In the left IFG (BA44) we found higher High-RC than Low-RC activation, and a significant positive correlation between activation level and AR capability, only in the High-RC task. The left IFG BA44 comprises Broca's area, a brain region playing a central role in semantic processing and serial working memory (Nee et al., 2013). The manipulation of relational complexity in the current study involved an increase in the number of visual features that could have been processed successively (Halford et al., 1998). Specifically, in the High-RC tasks the number of cued features was larger, increasing working memory load. Similarly, in the left SPL (part of the working memory network) there was greater High-RC than Low-RC activation, and positive correlation between activation and High-RC capability. This left hemisphere High-RC lateralization is consistent with theories signifying that the left hemisphere is primarily associated with analytic/serial processing, whereas the right hemisphere is primarily associated with holistic/parallel processing (Aichelburg et al., 2016; Vendetti et al., 2015).

Unlike the left IFG BA44 (and left SPL), the left IFG BA46 and the adjoined left MFG (also part of BA46) were only positively correlated with AR capability in the High-RC task, with no significant sensitivity to relational complexity. The dorsolateral PFC (lateral BA46 and BA9) has been reported to play a central role in sustained attention and in the manipulation of both serial and spatial information in working memory (Nee et al., 2013). Our findings suggest that while levels of activation in the left IFG (BA44) are associated with the objective amount of information being processed, as well as with individual proficiencies in using

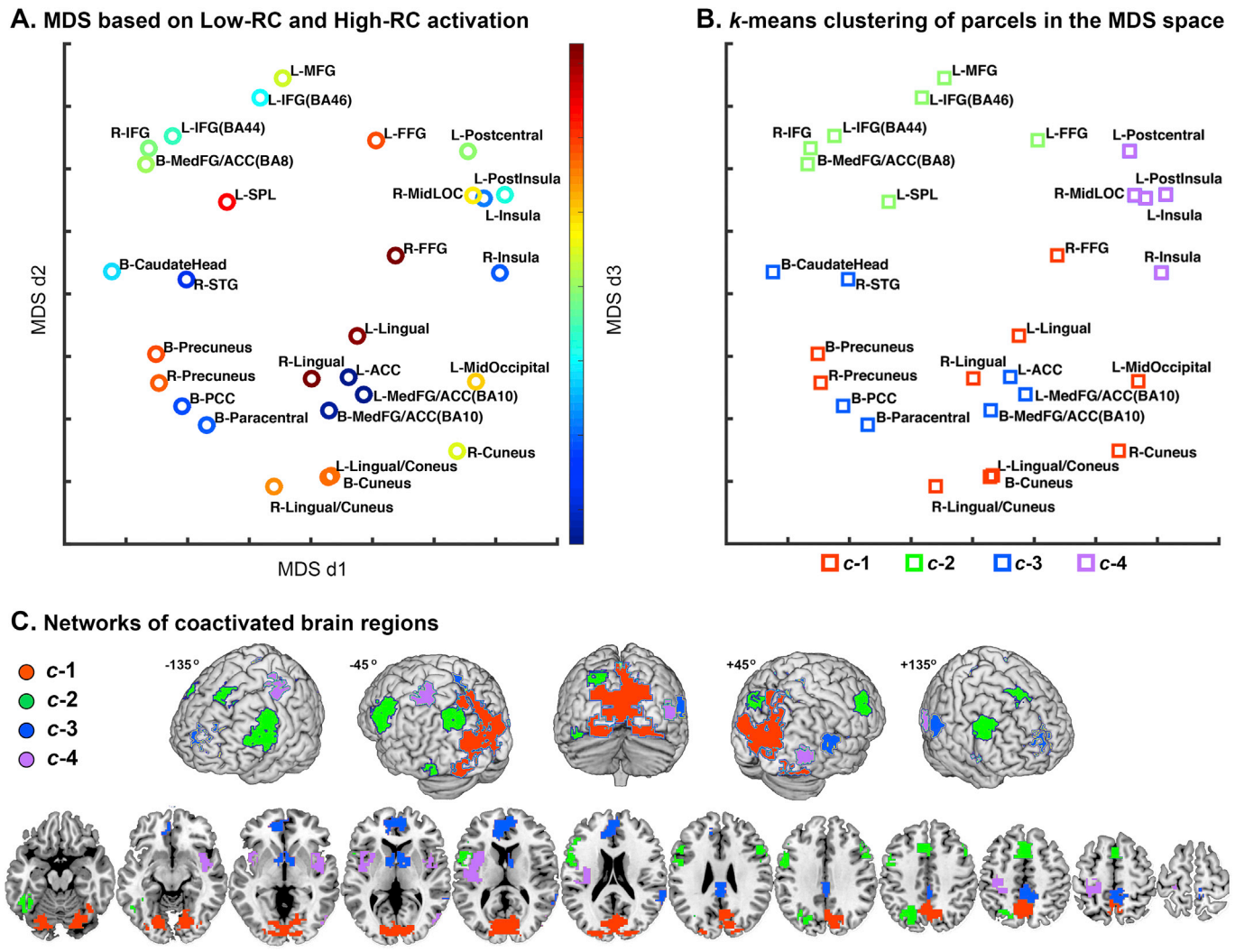


**Fig. 5.** Networks of co-activated brain parcels (see Supplemental 6 for between-parcels correlations). BA# in a parcel label indicates the respective Brodmann Area. (A) MDS based on the Low-RC and High-RC activation, combined, representing similarities in activation level between the parcels in a three-dimensional space (colorbar signify 'z-axis' values in the MDS space, where blue being the lowest). (B) *k*-means clustering ($k = 4$) of the parcels based on their proximities in the MDS space. For simplification, parcels are presented in the first two dimensions of the MDS space where colors (red, green, blue and purple) are the clusters labels. (C) A brain map with networks of co-activated brain regions, color labeled based on the *k*-means clustering.

this information, dorsolateral PFC (BA46) activation is mostly associated with the individual information processing proficiencies, when the cognitive load is high.

### 4.2. AR and occipito-parietal (visuospatial network) activation

Neural activation in occipital and parietal brain parcels was also positively associated with AR (Fig. 4; Fig. 5). These visuospatial parcels can be sorted into a few subgroups. First, the cuneus (BA17) serves as the primary visual cortex, it executes lower level visual processing and projects to extrastriate cortices (DeYoe et al., 1996). Here we found that activation in the cuneus (bilaterally) was significantly affected by relational complexity, without being associated with individual AR capability. Since the visual stimuli in the High-RC and Low-RC tasks were identical, such differences in activation can only be attributed to top-down attentional modulation, where the High-RC task required accounting for more visual information.

The second subgroup of visual brain parcels was in the lingual gyrus, which receives information from the cuneus (two of the lingual gyrus parcels partially overlapped with the cuneus). The primary functions of the lingual gyrus include complex feature processing, stimuli comparison and feature mapping (which are key to visuospatial AR and category learning), low-level featural working memory and imagery (Albers et al., 2013; Hammer et al., 2009, 2010; Harrison and Tong, 2009; Xu et al., 2015). Unlike the cuneus, activation in the lingual gyrus (bilaterally) was not only significantly associated with relational complexity, but also with the individual AR capability, in both relational complexity conditions. As with the cuneus, sensitivity to relational complexity in the lingual gyrus can only be attributed to top-down attentional modulation. However, positive correlation between AR capability and lingual gyrus activation can either indicate that better AR capability requires better top-down access and retrieval of sensory information available in the lingual gyrus; or it may indicate that AR capability depends on individual expertise associated with the capacity to process complex visual features within the lingual gyrus (thus may provide a bottom-up contribution to AR).

Beyond the cuneus and lingual gyrus, higher-level parcels within the visuospatial network (red network in Fig. 5) significantly (positively) contributed only to performance in the High-RC task (Fig. 4, Supplemental 3). These included the left middle occipital cortex, the right precuneus and the right FFG, which are respectively associated with regularity detection and feature grouping (Cardin et al., 2011),

integrative multimodal spatial attention (Shomstein and Yantis, 2006) and object recognition (Pinel et al., 2014). As reported above, activation levels in the left FFG and left SPL were also positively correlated with the individual AR capability in the High-RC task. However, activation patterns in these two brain regions were more like activation patterns in the central executive network (green network in Fig. 5) than to activation patterns in lower-level visual cortices. This left hemisphere visuospatial lateralization is consistent with the reported above left prefrontal dominance (see Macaluso et al., 2000 for a discussion on back-projections lateralization).

### 4.3. AR and ventromedial (default network) activation

We found three ventromedial brain parcels to be associated with AR. Two adjoined, left ACC and left MedFG/ACC (both BA10) parcels significantly contributed to the High-RC individual differences regression model, with negative correlation between activation level and AR capability (Figs. 4 and 5, Supplemental 3). A third, bilateral MedFG/ACC (BA10) parcel showed greater activation in Low-RC than in High-RC (note that the correlation between neural activation in this parcel and AR capability was also negative, but with no significant contribution to either one of the individual differences regression models; Figs. 5A and 4, Supplemental 3). Activation pattern in the ventromedial cortices was also strongly associated with activation in the PCC, paracentral and right STG parcels, all of which were more activated in the Low-RC task than in the High-RC task (Figs. 4C and 5).

The ventral MedFG/ACC, PCC and the superior temporal cortex (including the right-STG) are core modules of the default-mode brain network, which is involved in both spontaneous and goal-directed internal mentation. Amongst typical young adults, this network is characterized by higher activation when not being engaged in a specific behavioral task, and lower activation (deactivation in respect to resting state levels) as task demands increases. When being engaged in a demanding visual task, activation in the default network is negatively correlated with activation in frontoparietal brain regions. It has been suggested that default network deactivation is associated with the suppression of internally induced irrelevant information. (Andrews-Hanna, 2012; Anticevic et al., 2012; Christoff et al., 2016).

Consistent with earlier findings, we found lower activation in the bilateral ventral MedFG/ACC, PCC, paracentral and right STG brain parcels, in the High-RC task compared to the Low-RC task. That is, default network activation was negatively associated with relational complexity. Moreover, greater deactivation in the left ventral MedFG/ACC was associated with better capability in High-RC tasks, indicating that an effective suppression of the default network is crucial as AR demands increases. This deactivation likely reduced processing load in the central executive network, leaving more resources for the AR task. Consistent with this interpretation, we found greater negative correlations between default network parcels and executive and visuospatial parcels in the High-RC task than in the Low-RC task (Supplemental 6). Previous findings (based on Granger causality analysis) suggest that the default network has a greater impact on its anticorrelated networks than the other way around, implying that it may alter information processes in task-positive networks, and compromise task performance (Uddin et al., 2009; see Preusse et al., 2011 for an alternative interpretation of ventral MedFG/ACC deactivation in AR).

### 4.4. AR and insular ('salience network') activation

The anterior insula in both hemispheres and the left posterior insula were associated with AR. Although correlated with one another (Fig. 5, Supplemental 6), each of these three brain parcels played a distinct role in AR. The left anterior insula was amongst the brain regions most sensitive to relational complexity, with greater activation in the Low-RC task than in the High-RC task. While activation in the right anterior insula was negatively correlated with AR capability in the Low-RC task, activation in the left posterior insula was positively correlated with Low-RC capability (Fig. 4B). Insular pattern of activation was associated with activation in the left postcentral and right middle lateral occipital cortex, both were more activated in the Low-RC task than in the High-RC task (Figs. 4C and 5).

Current accounts indicate that the insula plays a central role in attentional control. The anterior insula has excessive functional connections with both the central executive network and the default network, suggesting that it mediates between these two brain networks. The posterior insula plays greater role in modulating responses to external stimuli. Together with the dorsal ACC, the insula forms a 'salience network', which identifies potentially relevant 'neural events' (internally or externally originated) in the brain networks with which it interacts. This may result in allocation of more cognitive resources to events with potential significance (Goulden et al., 2014; Menon and Uddin, 2010).

Being a hub controlling information exchange between other networks, levels of activation in the insula correlate with levels of information exchange. Thus, lower High-RC activation than Low-RC activation in the left anterior insula may indicate uncoupling between the central executive network and the default network when AR load increases (see Supplemental 6 for supporting evidences; see Hammer et al., 2015a for related findings). Patterns of activation in the left posterior insula and right anterior insula may indicate that better capability in the Low-RC task required greater exchange of information between left hemisphere visuospatial brain regions and the central executive network, while suppressing information exchange in the right hemisphere.

### 4.5. Summary and conclusions

The analysis employed here revealed a noticeable pattern – lowest-level visuospatial brain regions (cuneus) were affected only by relational complexity whereas intermediate-level visuospatial brain regions (lingual; stimuli comparison and feature mapping) were sensitive both to relational complexity and to the individual AR capability. This suggests that processing capacity and sustained representation of visual information in the lingual gyrus is the lowest-level processing bottleneck affecting the individual visuospatial AR capability. Higher-level, functionally more specialized visuospatial brain regions (FFG, SPL) were associated with individual differences in AR only in the High-RC task.

Amongst the central executive brain regions associated with AR we also found diverse sensitivity to individual differences versus sensitivity to relational complexity. The left IFG BA46 and the adjoined left MFG showed sensitivity only to individual differences in the High-RC task. As discussed above, the lateral BA46 is involved in sustained attention and the manipulation of information available in working memory (Barbey et al., 2013; Nee et al., 2013). However, not being significantly sensitive to the relational complexity manipulation indicates that activation in the lateral BA46 is a poor indicator of the objective amount of task-relevant information. Similarly, the left and right FFG showed sensitivity only to individual differences in the High-RC task. This indicates that better AR capability in scenarios that involve complex mapping of multiple features relies more on these higher-level visual processing brain regions, likely due to an increase in specialized visual processing demands in High-RC tasks. The right lateral BA46 (right IFG) may play a role in attentional control similar to that of the left lateral BA46. However, the right lateral BA46 may be more limited in its processing capacity, and thus individual differences in it were evident in the Low-RC task, but not in the High-RC task.

The left IFG (BA44) and left SPL both play central role in working memory, with a likely greater role in serial working memory (Newman et al., 2010). These brain regions were more activated in the High-RC task, than in the Low-RC task where less visual features had to be accounted for. Furthermore, the left IFG (BA44) and left SPL significantly contributed to the individual AR capability in the High-RC task. These may indicate that a better capability accessing information available in

these brain regions contribute to AR. Notably, we found the respective weight of the left IFG to be with very small variability in the 5-fold cross-validation, both in the High-RC regression model and in the High-RC versus Low-RC classification model (Supplemental 3). Such highly consistent mid-size weight indicates that the left IFG plays a necessary (but not sufficient) role in High-RC AR (in the healthy young adult population), likely by managing information available in serial working memory (Barbey et al., 2013). However, its overall contribution to the individual AR capability seem to be limited, as we found other brain regions to be with greater contribution to AR (substantially larger mean weight).

Finally, the dorsomedial B-MedFG/ACC (BA8) likely integrates information relevant for determining the validity of the analogy. Besides the lingual gyrus, the dorsomedial B-MedFG/ACC was the only brain region in which level of activation was significantly positively correlated with individual differences in both relational complexity conditions, and where activation in the High-RC task was greater than in the Low-RC task (pink parcels in Fig. 4D). Level of neural activation in the dorsomedial B-MedFG/ACC was also amongst the best predictors (together with the lingual gyrus) in the High-RC and Low-RC individual differences regression models (Fig. 4; Supplemental 3). However, levels of activation in the dorsomedial B-MedFG/ACC and lingual gyrus were not correlated (Supplemental 6).

The above indicate that AR relies on a combination of visuospatial expertise and executive skills (see Hammer and Sloutsky, 2016, Hummel and Holyoak, 1997 and Nikolaidis et al., 2016 for related discussions). That is, AR does not manifest solely in a few key brain regions at the 'top of a hierarchical processing pipeline'. Instead, visuospatial AR capability emerge from a heterarchical brain network of quasi specialized brain regions incorporating several partially independent and more basic visuospatial and executive processes, contingent on the individual skills and on task demands. Together, these brain regions form a flexible yet robust and efficient network which is in part hierarchical and in part lateral/horizontal. Such brain system organization enables multiple potential distinct resolutions for multiple computationally complex problems (Cumming, 2016; McCulloch, 1945; Medaglia et al., 2015), and thus may be characterized by substantial individual differences in patterns of neural activation, even amongst individuals with comparable cognitive skills. Consequently, it is expected that different individuals may improve their reasoning skills by improving different subsets of basic cognitive proficiencies (see Supplemental 7 for a model summary and illustrative examples).

## Authors contribution

Conceptualization of Research Objectives: R.H. and A.B.; Data Acquisition Oversight: E.P.; Data Curation: R.H. and E.P.; Software: R.H.; Formal Analysis: R.H.; Visualization: R.H.; Writing – Original Draft: R.H.; Writing – Review & Editing: R.H., A.B., C.H. and A.K.; Resources and Funding Acquisition: A.B., N.C., A.K. and C.H.

## Acknowledgments

This research is based upon work supported by the Office of the Director of National Intelligence, Intelligence Advanced Research Projects Activity (IARPA), via Contract 2014–13121700004 to the University of Illinois at Urbana-Champaign (PI: A. Barbey). The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of the ODNI, IARPA, or the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes not-withstanding any copyright annotation thereon.

We thank John Hummel for insightful discussions, Joachim Operskalski for his comments on this manuscript, and the members in the Decision Neuroscience Laboratory for their assistance in data collection.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.neuroimage.2018.09.011.

## References

Aichelburg, C., Urbanski, M., de Schotten, M.T., Humbert, F., Levy, R., Volle, E., 2016. Morphometry of left frontal and temporal poles predicts analogical reasoning abilities. Cerebr. Cortex 26 (3), 915–932.
Albers, A.M., Kok, P., Toni, I., Dijkerman, H.C., de Lange, F.P., 2013. Shared representations for working memory and mental imagery in early visual cortex. Curr. Biol. 23 (15), 1427–1431.
Andrews-Hanna, J.R., 2012. The brain's default network and its adaptive role in internal mentation. Neuroscientist 18 (3), 251–270.
Anticevic, A., Cole, M.W., Murray, J.D., Corlett, P.R., Wang, X.J., Krystal, J.H., 2012. The role of default network deactivation in cognition and disease. Trends Cognit. Sci. 16 (12), 584–592.
Barbey, A.K., Koenigs, M., Grafman, J., 2013. Dorsolateral prefrontal contributions to human working memory. Cortex 49 (5), 1195–1205.
Benjamini, Y., Yekutieli, D., 2001. The control of the false discovery rate in multiple testing under dependency. Ann. Stat. 1165–1188.
Bunge, S.A., Wendelken, C., Badre, D., Wagner, A.D., 2005. Analogical reasoning and prefrontal cortex: evidence for separable retrieval and integration mechanisms. Cerebr. Cortex 15 (3), 239–249.
Cardin, V., Friston, K.J., Zeki, S., 2011. Top-down modulations in the visual form pathway revealed with dynamic causal modeling. Cerebr. Cortex 21 (3), 550–562.
Chapelle, O., Haffner, P., Vapnik, V.N., 1999. Support vector machines for histogram-based image classification. IEEE Trans. Neural Network. 10 (5), 1055–1064.
Cho, S., Moody, T.D., Fernandino, L., Mumford, J.A., Poldrack, R.A., Cannon, T.D., et al., 2010. Common and dissociable prefrontal loci associated with component mechanisms of analogical reasoning. Cerebr. Cortex 20 (3), 524–533.
Christoff, K., Irving, Z.C., Fox, K.C., Spreng, R.N., Andrews-Hanna, J.R., 2016. Mind-wandering as spontaneous thought: a dynamic framework. Nat. Rev. Neurosci. 17 (11), 718–731.
Chuderski, A., Taraday, M., Nęcka, E., Smoleń, T., 2012. Storage capacity explains fluid intelligence but executive control does not. Intelligence 40 (3), 278–295.
Craddock, R.C., James, G.A., Holtzheimer, P.E., Hu, X.P., Mayberg, H.S., 2012. A whole brain fMRI atlas generated via spatially constrained spectral clustering. Hum. Brain Mapp. 33 (8), 1914–1928.
Cumming, G.S., 2016. Heterarchies: reconciling networks and hierarchies. Trends Ecol. Evol. 31 (8), 622–632.
de la Vega, A., Chang, L.J., Banich, M.T., Wager, T.D., Yarkoni, T., 2016. Large-scale meta-analysis of human medial frontal cortex reveals tripartite functional organization. J. Neurosci. 36 (24), 6553–6562.
Deary, I.J., 2001. Human intelligence differences: towards a combined experimental–differential approach. Trends Cognit. Sci. 5 (4), 164–170.
DeYoe, E.A., Carman, G.J., Bandettini, P., Glickman, S., Wieser, J., Cox, R., et al., 1996. Mapping striate and extrastriate visual areas in human cerebral cortex. Proc. Natl. Acad. Sci. Unit. States Am. 93 (6), 2382–2386.
Geake, J.G., Hansen, P.C., 2010. Functional neural correlates of fluid and crystallized analogizing. Neuroimage 49 (4), 3489–3497.
Gentner, D., Smith, L., 2012. Analogical reasoning. Encyclopedia of Human Behavior 130–136.
Goulden, N., Khusnulina, A., Davis, N.J., Bracewell, R.M., Bokde, A.L., McNulty, J.P., Mullins, P.G., 2014. The salience network is responsible for switching between the default mode network and the central executive network: replication from DCM. Neuroimage 99, 180–190.
Halford, G.S., Wilson, W.H., Phillips, S., 1998. Processing capacity defined by relational complexity: implications for comparative, developmental, and cognitive psychology. Behav. Brain Sci. 21 (06), 803–831.
Hammer, R., Sloutsky, V., 2016. Visual category learning results in rapid changes in brain activation reflecting sensitivity to the category relation between perceived objects and to decision correctness. J. Cognit. Neurosci. 28 (11), 1804–1819.
Hammer, R., Brechmann, A., Ohl, F., Weinshall, D., Hochstein, S., 2010. Differential category learning processes: the neural basis of comparison-based learning and induction. Neuroimage 52 (2), 699–709.
Hammer, R., Cooke, G.E., Stein, M.A., Booth, J.R., 2015b. Functional neuroimaging of visuospatial working memory tasks enables accurate detection of attention deficit and hyperactivity disorder. Neuroimage: Clinic 9, 244–252.
Hammer, R., Diesendruck, G., Weinshall, D., Hochstein, S., 2009. The development of category learning strategies: what makes the difference? Cognition 112 (1), 105–119.
Hammer, R., Tennekoon, M., Cooke, G.E., Gayda, J., Stein, M.A., Booth, J.R., 2015a. Feedback associated with expectation for larger-reward improves visuospatial working memory performances in children with ADHD. Developmental Cognitive Neuroscience 14, 38–49.
Harrison, S.A., Tong, F., 2009. Decoding reveals the contents of visual working memory in early visual areas. Nature 458 (7238), 632–635.
Haufe, S., Meinecke, F., Görgen, K., Dähne, S., Haynes, J.D., Blankertz, B., Bießmann, F., 2014. On the interpretation of weight vectors of linear models in multivariate neuroimaging. Neuroimage 87, 96–110.
Hobeika, L., Diard-Detoeuf, C., Garcin, B., Levy, R., Volle, E., 2016. General and specialized brain correlates for analogical reasoning: a meta-analysis of functional imaging studies. Hum. Brain Mapp. 37 (5), 1953–1969.

Holyoak, K.J., 2012. Analogy and relational reasoning. The Oxford Handbook of Thinking and Reasoning 234–259.
Hummel, J.E., Holyoak, K.J., 1997. Distributed representations of structure: a theory of analogical access and mapping. Psychol. Rev. 104 (3), 427.
Krawczyk, D.C., Holyoak, K.J., Hummel, J.E., 2004. Structural constraints and object similarity in analogical mapping and inference. Think. Reas. 10 (1), 85–104.
Kyllonen, P.C., Christal, R.E., 1990. Reasoning ability is (little more than) working-memory capacity?! Intelligence 14 (4), 389–433.
Lanckriet, G.R., De Bie, T., Cristianini, N., Jordan, M.I., Noble, W.S., 2004. A statistical framework for genomic data fusion. Bioinformatics 20 (16), 2626–2635.
Macaluso, E., Frith, C.D., Driver, J., 2000. Modulation of human visual cortex by crossmodal spatial attention. Science 289 (5482), 1206–1208.
Mazaika, P., Hoeft, F., Glover, G.H., Reiss, A.L., June, 2009. Methods and Software for FMRI Analysis for Clinical Subjects. Organization for Human Brain Mapping (OHBM), San-Francisco, California.
McCulloch, W.S., 1945. A heterarchy of values determined by the topology of nervous nets. Bull. Math. Biol. 7 (2), 89–93.
Medaglia, J.D., Lynall, M.E., Bassett, D.S., 2015. Cognitive network neuroscience. J. Cognit. Neurosci. 27 (8), 1471–1491.
Menon, V., Uddin, L.Q., 2010. Saliency, switching, attention and control: a network model of insula function. Brain Struct. Funct. 214 (5–6), 655–667.
Nee, D.E., Brown, J.W., Askren, M.K., Berman, M.G., Demiralp, E., Krawitz, A., Jonides, J., 2013. A meta-analysis of executive components of working memory. Cerebr. Cortex 23 (2), 264–282.
Newman, A.J., Supalla, T., Hauser, P., Newport, E.L., Bavelier, D., 2010. Dissociating neural subsystems for grammar by contrasting word order and inflection. Proc. Natl. Acad. Sci. Unit. States Am. 107 (16), 7539–7544.
Nikolaidis, A., Baniqued, P.L., Kranz, M.B., Scavuzzo, C.J., Barbey, A.K., Kramer, A.F., Larsen, R.J., 2016. Multivariate associations of fluid intelligence and NAA. Cerebr. Cortex bhw070.
Pinel, P., Lalanne, C., Bourgeron, T., Fauchereau, F., Poupon, C., Artiges, E., et al., 2014. Genetic and environmental influences on the visual word form and fusiform face areas. Cerebr. Cortex bhu048.
Platt, M.L., Huettel, S.A., 2008. Risky business: the neuroeconomics of decision making under uncertainty. Nat. Neurosci. 11 (4), 398–403.
Preusse, F., Van Der Meer, E., Deshpande, G., Krueger, F., Wartenburger, I., 2011. Fluid intelligence allows flexible recruitment of the parieto-frontal network in analogical reasoning. Front. Hum. Neurosci. 5, 22.
Qureshi, M.N.I., Oh, J., Min, B., Jo, H.J., Lee, B., 2017. Multi-modal, multi-measure, and multi-class discrimination of ADHD with hierarchical feature extraction and extreme learning machine using structural and functional brain MRI. Front. Hum. Neurosci. 11.
Rakotomamonjy, A., Bach, F.R., Canu, S., Grandvalet, Y., 2008. Simple-mkl. J. Mach. Learn. Res. 9, 2491–2521.
Shomstein, S., Yantis, S., 2006. Parietal cortex mediates voluntary control of spatial and nonspatial auditory attention. J. Neurosci. 26 (2), 435–439.
Sternberg, R.J., 1977. Component processes in analogical reasoning. Psychol. Rev. 84 (4), 353.
Tuia, D., Camps-Valls, G., Matasci, G., Kanevski, M., 2010. Learning relevant image features with multiple-kernel classification. IEEE Trans. Geosci. Rem. Sens. 48 (10), 3780–3791.
Uddin, L.Q., Clare Kelly, A.M., Biswal, B.B., Xavier Castellanos, F., Milham, M.P., 2009. Functional connectivity of default mode network components: correlation, anticorrelation, and causality. Hum. Brain Mapp. 30 (2), 625–637.
Unsworth, N., Fukuda, K., Awh, E., Vogel, E.K., 2014. Working memory and fluid intelligence: capacity, attention control, and secondary memory retrieval. Cognit. Psychol. 71, 1–26.
Urbanski, M., Bréchemier, M.L., Garcin, B., Bendetowicz, D., de Schotten, M.T., Foulon, C., et al., 2016. Reasoning by analogy requires the left frontal pole: lesion-deficit mapping and clinical implications. Brain 139 (6), 1783–1799.
Vendetti, M.S., Johnson, E.L., Lemos, C.J., Bunge, S.A., 2015. Hemispheric differences in relational reasoning: novel insights based on an old technique. Front. Hum. Neurosci. 9.
Wainer, J., Cawley, G., 2017. Empirical evaluation of resampling procedures for optimizing SVM hyperparameters. J. Mach. Learn. Res. 18 (15), 1–35.
Watson, C.E., Chatterjee, A., 2012. A bilateral frontoparietal network underlies visuospatial analogical reasoning. Neuroimage 59 (3), 2831–2838.
Xu, M., Wang, T., Chen, S., Fox, P.T., Tan, L.H., 2015. Effective connectivity of brain regions related to visual word recognition: an fMRI study of Chinese reading. Hum. Brain Mapp. 36 (7), 2580–2591.
Xu, Z., Jin, R., Yang, H., King, I., Lyu, M.R., 2010. Simple and efficient multiple kernel learning by group lasso. In: Proceedings of the 27th International Conference on Machine Learning (ICML-10), pp. 1175–1182.