Science
AAAS

# Supplementary Materials for

## A Common Architecture for Human and Human-Like Artificial Cognition Explains Brain Activity Across Domains

Andrea Stocco, Zoe Steine-Hanson, Natalie Koh, John E. Laird,
Christian J. Lebiere, Paul S. Rosenbloom

Correspondence to: stocco@uw.edu

**This PDF file includes:**

Materials and Methods
Figs. S1 to S4
Tables S1 to S3

## Materials and Methods

The study presented herein consists of an extensive analysis of a large sample (*N*=200) of neuroimaging data from the Human Connectome Project, the largest existing repository of young adult neuroimaging data. The analysis was restricted to the task fMRI subset, thus excluding both the resting state fMRI data, the diffusion imaging data, and all of the M/EEG data. The task fMRI data consisted of two sessions of each of seven paradigms, designed to span different domains.

### Tasks fMRI Data

The HCP task-fMRI data encompasses seven different paradigms designed to capture a wide range of cognitive capabilities. Of these paradigms, six were included in our analysis. The Motor Mapping task was not included because it would have required the creation of multiple ROIs in the motor cortex, one for each effector (arm, leg, voice), thus making this model intrinsically different from the others. A full description of these tasks and the rationale for their selection can be found in the original HCP papers (*14*, *34*). This section provides a brief description of the paradigms, while Table S1 provides an overview.

| **Task (Representative Reference)** | **Relevant Conditions (for GLM analysis)** | **Included in DCM analysis?** |
|---|---|---|
| *Motor Mapping (35)* | *Hand, arm, foot, leg, voice responses* | *No* |
| Emotion Processing (*17*) | Neutral shapes vs. **Fearful and angry faces**. | Yes |
| Incentive Processing (*16*) | "Winning" vs. **"Losing"** blocks of choices | Yes |
| Language and Mathematical Processing (*15*) | Listening vs. **Answering questions** (in both Language and Math blocks) | Yes |
| Relational Reasoning (*19*) | Control Arrays vs. **Relational arrays** | Yes |
| Social Cognition (*18*) | Randomly moving shapes vs. **Socially interacting shapes** | Yes |
| Working Memory | 0-Back vs. **2-Back blocks** of faces, places, tools, and body parts. | Yes |

**Table S1**: Overview of the seven task-fMRI paradigms used in the HCP dataset. Italics indicate tasks and conditions that were not included in our analysis; bold typeface marks experimental conditions that were selected as "Critical" (as opposed to "Baseline") in the design of the experimental matrices (see below, "DCM-specific GLM analysis" section)

**Emotion Processing Task**. Participants are presented with 12 blocks of six consecutive trials. During each trial, they are asked to decide either which of two visual stimuli presented on the bottom of the screen match the stimulus at the top of the screen. In six of the blocks, all of the visual stimuli are emotional faces, with either angry or fearful expressions. In the remaining six blocks, all of the stimuli are neutral shapes. Each stimulus is presented for 2 s, with a 1 s inter-trial interval (ITI). Each block is preceded by a 3 s task cue ("shape" or "face"), so that each block is 21 s including the cue.

**Incentive Processing Task**. The task consists of four blocks of eight consecutive decision-making trials. During each trial, participants are asked to guess whether the number underneath a "mystery card" (visually represented by the question mark symbol "?") is larger or smaller than 5 by pressing one of two buttons on the response box within the allotted time. After each choice, the number is revealed; participants receive a monetary reward (+$1.00) for correctly guessed trials; a monetary loss (-$0.50) for incorrectly guessed trials; and receive no money if the number is exactly 5. Unbeknownst to participants, blocks are pre-designed to lead to either high rewards (6 reward trials, 2 neutral trials) or high losses (6 loss trials, 2 neutral trials), independent of their actual choices. Two blocks are designated as high-reward, and two as high-loss blocks. Each stimulus has a duration of up to 1.5 s, followed by a 1 s feedback, with a 1 s ITI, so that each block lasts 27 s.

**Language and Mathematical Processing Task**. The task consists of 4 "story" blocks interleaved with 4 "math" blocks. The two types of blocks are matched for duration, and adhere to the same internal structure in which a verbal stimulus is first presented auditorily, and a two-alternative question is subsequently presented. Participants need to respond to the question by pressing one of two buttons with the right hand. In the story blocks, the stimuli are brief, adapted Aesop stories (between 5 and 9 sentences), and the question concerns the story's topic (e.g., "Was the story about *revenge* or *reciprocity*?"). In the math blocks, stimuli are addition or subtraction problems (e.g., "Fourteen plus twelve") and the question provides two possible alternative answers (e.g., "*Twenty-nine* or *twenty-six*?"). The math task is adaptive to maintain a similar level of difficulty across the participants.

**Relational Processing Task**. The task consists of six "Relational" blocks alternated with six "Control" blocks. In relational blocks, stimuli consist of two pairs of figures, one displayed horizontally at the top of the screen and one pair displayed at the bottom. Figures consist of one of six possible shapes filled with one of six possible textures, for a total of 36 possible figures. Both pairs of figures differ along one dimension, either shape or texture; participants are asked to indicate through a button press if the top figures differ on the same dimension as the bottom figures (e.g., they both differ in shape). In the control blocks, the stimuli consist of one pair of figures displayed horizontally at the top of the screen, a third figure displayed centrally at the bottom of the screen, and a word displayed at the center of the screen. The central word specifies a stimulus dimension (either "shape" or "texture") and participants are asked to indicate whether the bottom figure matches either of the two top figures along the dimension specified by the word. Both relational and control blocks have a total duration of 16 s, but they vary in the number of stimuli. Specifically, relational blocks contain four stimuli, presented for 3.5 s with a 500 ms ITI, while control blocks contain five stimuli presented for 2.8 s with a 400 ms ITI.

**Social Cognition Task**. The tasks consists of 10 videoclips of moving shapes (circles, squares, and triangles). The clips were either obtained or modified from previously published studies (*18*, *36*). In five of the clips, the shapes are moving randomly, while in the other five the shapes' movement reflects a form of social interaction. After viewing each clip, participants press one of three buttons to indicate whether they believed the shapes were interacting, not interacting, or whether they were unsure. All clips have a fixed duration of 20 s with an ITI of 15 s.

**Working Memory**. The task consists of eight 2-back blocks and eight 0-back blocks, with each block containing 10 trials. Each trial presents the picture of a single object, centered on the screen, and participants have to press one of two button to indicate whether the object is a target or not. In the 2-back blocks, a target is defined as the same object that had been seen two trials before, so that participants have to maintain and update a "moving window" of the past two objects to perform the task correctly. In the 0-back blocks, a target is defined as a specific object, presented at the very beginning of the block, so that participants have to only maintain a single object in working memory throughout the block. The stimuli belong to one of four possible categories: faces, places, tools, and body parts. The category of the objects being used as stimuli changes from block to block, but is consistent within one block, so that there is an even number of face, place, tool, and body part blocks for each condition. Each block begins with a 2.5 s cue that informs the participant about the upcoming block type (2-back or 0-back). Each stimulus is presented for 2 s with a 500 ms ITI, for a total duration of 27.5 s per block.

Data Processing and Analysis

**Imaging Acquisition Parameters** As reported in (*34*), functional neuroimages were acquired with a 32-channel head coil on a 3T Siemens Skyra with TR = 720 ms, TE = 33.1 ms, FA = 52°, FOV = 208 × 180 mm. Each image consisted of 72 2.0mm oblique slices with 0-mm gap in-between. Each slice had an in-plane resolution of 2.0 x 2.0 mm. Images were acquired with a multi-band acceleration factor of 8x.

**Image Preprocessing** Images were acquired in the "minimally preprocessed" format (*14*), which includes unwarping to correct for magnetic field distortion, motion realignment, and normalization to the MNI template. The images were then smoothed with an isotropic 8.0 mm FWHM Gaussian kernel.

**Canonical GLM Analysis** Canonical GLM analysis was conducted on the smoothed minimally preprocessed data using a mass-univariate approach, as implemented in the SPM12 software package (*37*). First-level (i.e., individual-level) models were created for each participant. The model regressors were obtained by convolving a design matrix with a hemodynamic response function; the design matrix replicated the analysis of (*34*), and included regressors for the specific conditions of interests described in Table S1. Second-level (i.e., group-level) models were created using the brain-wise parametric images generated for each participant as input. The results of these second-level models replicated previous findings *(34)*, and are illustrated in Figure S1.

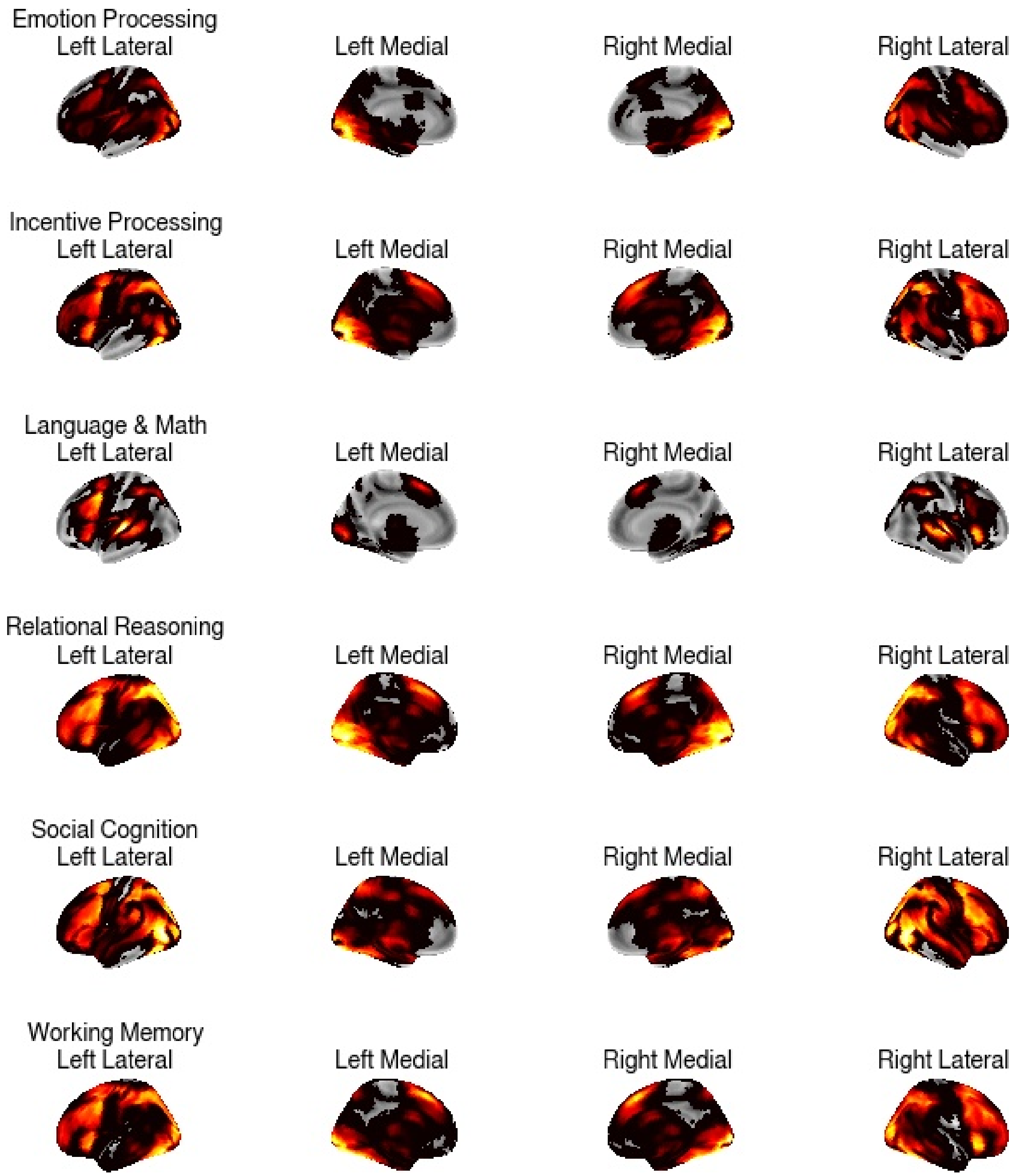


**Fig. S1.** Results of the group-level GLM analyses for each task.

**DCM-specific GLM Analysis** In parallel with the canonical GLM analysis, a second GLM analysis was carried out as a an aid to the DCM analysis. The purpose of this analysis was two-fold. First, it was needed to define the even matrix that is used in the DCM equation to measure the parameter matrix **C**. Second, it provided a way to define the omnibus *F*-test that is used in the ROI definition (see below). Because these models are not used to perform data analysis, the experimental events and conditions are allowed to be collinear. Because of the nature of cognitive neuroscience paradigms, all of our tasks include at least two different conditions under which stimuli must be processed in different ways. In all cases, the difference between conditions can be framed in terms of a more demanding, “critical” condition and an easier, “control” condition, with the more demanding events associated with greater mental elaboration of the stimuli. The critical condition of each task is highlighted in Table S1. As is common in DCM analysis, task conditions were modeled in a layered, rather than orthogonal fashion. The difference is illustrated in Figure S2: While in traditional GLM analysis the two conditions are modeled as non-overlapping events in the design matrix, in the DCM-specific definition of the matrix all trials belong to the same “baseline” condition, which represents the basic processing of the stimulus across all trials. Stimuli from the critical condition form a subset of all stimuli presented in the baseline condition. The critical condition is therefore appended to the baseline condition in the design matrix to model the additional processes that are specifically related to it. In DCM, each condition can affect one or more ROIs independently. In our analysis, the association between conditions and ROIs was kept constant across all tasks. Specifically, the baseline conditions selectively affected the perceptual ROI, while the critical condition selectively affected the WM ROI (Fig. S2C). This choice reflects the greater mental effort that is common to all critical conditions, and is confirmed by the greater PFC activity found in all of the GLM analyses of the critical conditions (*34*).

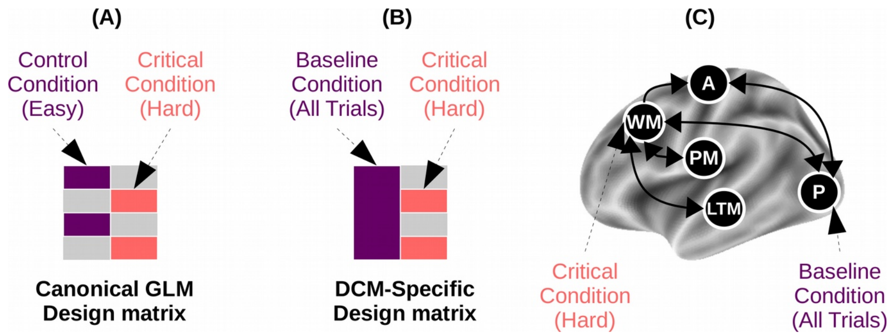


**Fig. S2.** Difference between design matrix used for canonical GLM (A) and for DCM analysis (B). In the DCMs, the Baseline condition drives neural activity in perceptual areas, while in the Critical condition drives neural activity in the Working Memory component (C).

Regions of Interests (ROI) Definition

The Regions of Interest (ROIs) were selected using a two-step procedure, designed to maximize the sensitivity of our analysis by separately accounting for two sources of variability in the spatial distribution of the ROIs.

The first step is designed to account for group-level variability due to the different tasks and stimuli used in the four datasets. This is necessary because, for example, the different complexity of the visual stimuli determine which portion of the visual cortex is most likely to be engaged, and different stimulus and task characteristics would engage different portions of the PFC. These differences were accounted for by conducting a separate group-level GLM analysis of each dataset, and identifying the coordinates of three points that have the highest statistical response within the anatomical boundaries of the primary visual areas (limited to the occipital lobe), the dorso-lateral PFC, and the basal ganglia (limited to the striatum).  Table S2 provides a detailed list of the coordinates of these regions across all tasks.

The second step was designed to account for individual-level variability in functional neuroanatomy. To do so, the task-specific group level coordinates from each task were then used as seed points to locate the closest local maximum in each subject's corresponding statistical parameter map. For maximal sensitivity, the map was derived from an omnibus *F*-test that included all the experimental conditions. In practice, this *F*-test was designed to capture any voxel that responded to any experimental condition. The same *F*-contrast was also use to adjust (i.e., mean-correct) each ROI's timeseries (*37*, *38*)

The individual coordinates, thus defined, were then visually inspected and when the coordinates were outside the predefined anatomical boundaries, manually re-adjusted. Across over 1,200 coordinates examined, only 2 required manual adjustment (~ 0.2%).

Finally, the individual coordinates were used as the center of a sphere. All voxels within the sphere whose response was significant at a minimal threshold of $p < 0.5$ were included as part of the ROI. Fig. S3 illustrates the mean number of voxels and standard deviation for each VOI in each task.

| Task | Action | LTM | Perception | Procedural | WM |
|---|---|---|---|---|---|
| Emotion Processing | -38, -26, 50 | -56, -18, 6 | -34, -80, -12 | -24, -4, 8 | -40, 6, 30 |
| Incentive Processing | -40, -22, 56 | -38, -6, -8 | -40, -80, -4 | -18, 6, 14 | -46, 2, 34 |
| Language & Math | -38, -26, 50 | -50, -10, -22 | -52, -22, 4 | -22, 10, 2 | -46, 2, 30 |
| Relational Reasoning | -38, -22, 56 | -48, -56, -14 | -12, -92, -2 | -16, 2, 18 | -44, 16, 44 |
| Social Cognition | -40, -18, 52 | -54, 0, -14 | -38, -84, -4 | -14, 12, 6 | -52, 6, 32 |
| Working Memory | -38, -22, 56 | -60, -36, -8 | -34, -50, -18 | -14, 10, 10 | -48, 22, 36 |

**Table S2.** Group-level coordinates of the centers of the five ROIs (corresponding to the five components of the CMC) across the seven tasks (Language and Math share the same paradigm and the same coordinates). Coordinates are given in *x*, *y*, *z* dimensions in MNI space. Each coordinate was used as the starting point to identify the closest peak for each individual participant's functional maps (see Fig. 2).

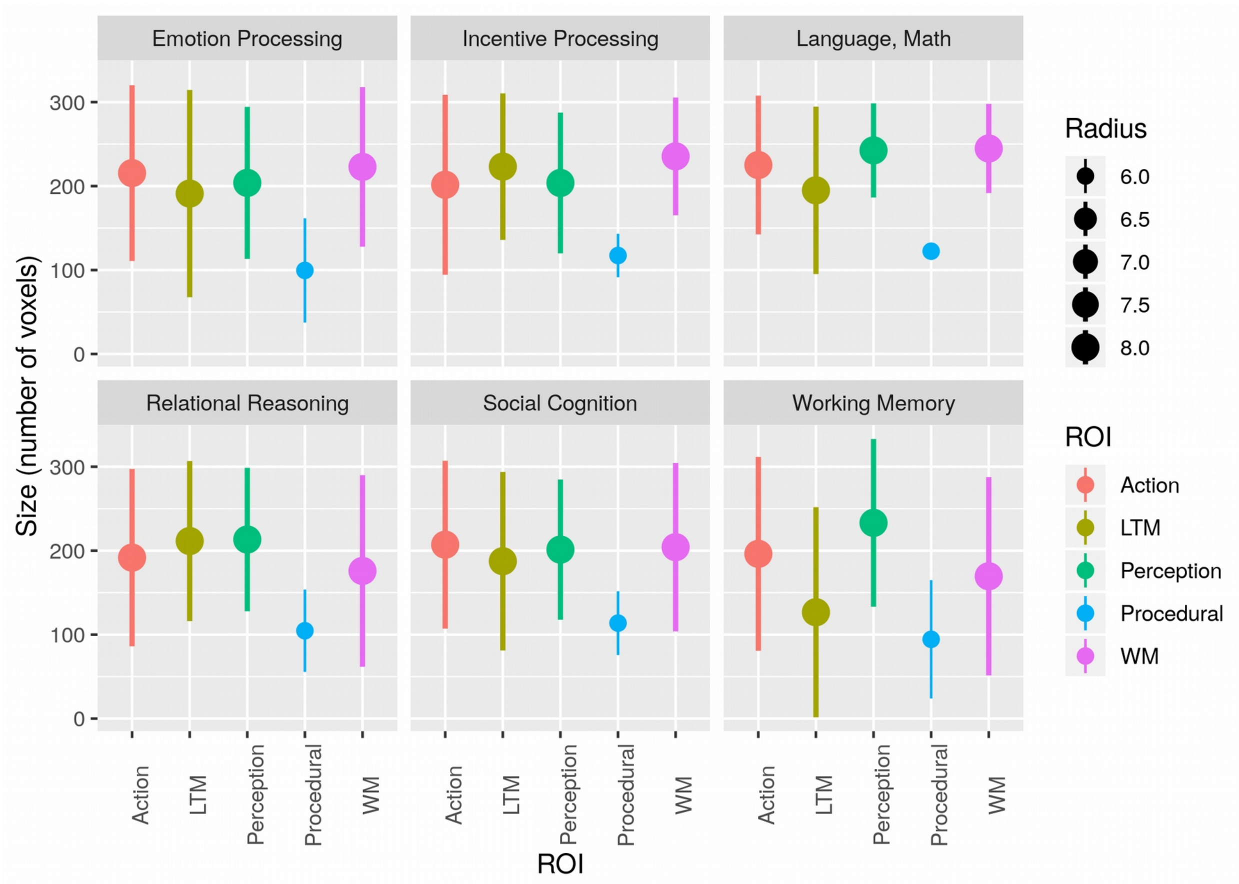


**Fig. S3**: Mean size (i.e., number of significant voxels included in each ROI spherical volume) of each ROI across participants for each task. Different colors represent different ROIs; dots represent means; dot size represents the ROI radius, and whiskers represent standard deviations.

Bayesian Model Selection

The goal of distinguishing which model provides a better fit to the data is, in turn, a problem of model comparison. In this study, a hierarchical Bayesian approach was employed, using a  random effects analysis of the model fits as outlined by (*29*). Compared to fixed-effects methods (such as log-likelihood and log-likelihood-derived methods, Bayesian Information Criterion (*39*) and Akaike Information Criterion (*30*)), a random effects analysis is less sensitive to errors introduced by outlier subjects.

In this analysis, different individuals might potentially be fit by different models (thus allowing a subject-level term), and models are compared on the basis of their relative probabilities of being the best-fitting architecture across the sample of individuals. Statistically, the prevalence of each model $k$ (i.e., the probability $r_k$ that $k$ would fit a random individual) in a sample of participants is drawn from a Dirichlet distribution Dir($\alpha_1$, $\alpha_2$,… $\alpha_k$) , and the distributions of probabilities of architectures 1, 2… $k$ across $n$ individuals are then drawn from multinomial distributions. The result of this modeling effort is a distribution of probabilities $r_k$ for each model k. These distributions can then be compared in terms of their relative *expected* and *exceedance* probability, that is, the mean probability of each model's $r_k$ across the sample and the probability that $r_k$ is larger than the competing models. Expected probability is calculated as the mean of each distribution; the properties of the Dirichlet distribution guarantee that the sum of the means of all distributions is 1. The Exceedance probability can be calculated by sampling from a multinomial distribution generated from random samples of the original distributions, thus again guaranteeing that all probabilities sum up to 1. Fig. S4, inspired by (*29*), provides a graphical illustration of the procedure. For simplicity, and following the convention of *(29)*, Fig. S4 depicts illustrates exceedance probability in the trivial case in which only two models are present, in which case it reduces to the area of the distribution to the right of $r_k > 0.5$; when more than two models are present, no simply visual interpretation is possible. Table S3 provides a detailed list of model comparison metrics, including the ones derived from the hierarchical Bayesian procedure used in this study (Dirichlet's α, expected, and exceedance probabilities) as well as the group-level log-likelihood of each model.

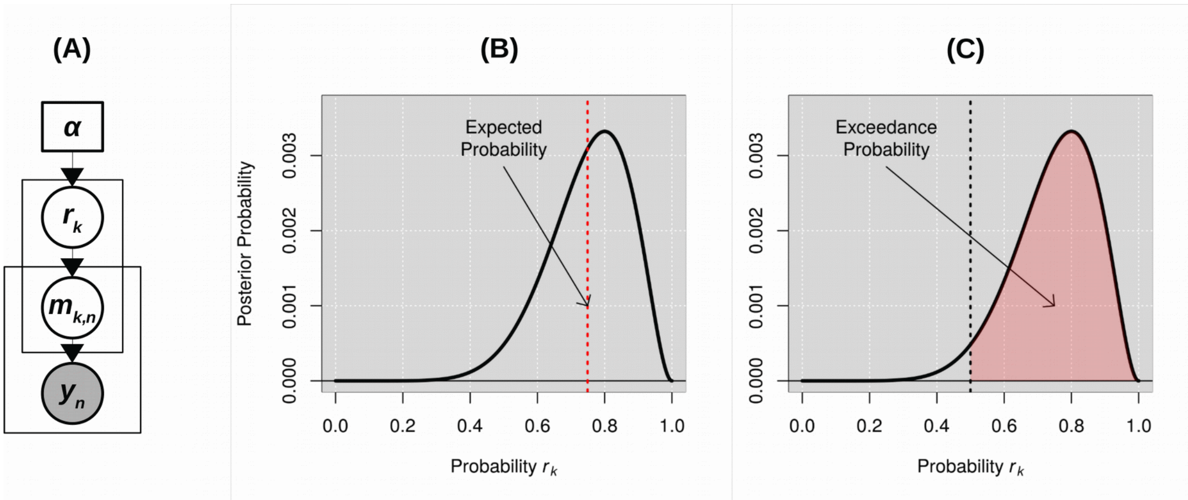


**Fig. S4.** Hierarchical Bayesian model selection estimation procedure and measures for model comparison (A): Compact visual representation of the hierarchical Bayesian modeling procedure; the procedure yields a distribution of probabilities $r_k$ that each model *k* would fit any individual *n* in the sample. (B): Visual representation of a model's posterior probability distribution over $r_k$ (black curve) and its expected probability (red dashed line); the expected probability is the mean expected value of $r_k$, i.e. $\int_0^1 r_k \times y$; (C) Visual representation of a model's posterior probability distribution over $r_k$ (black curve) and the corresponding exceedance probability (red shaded area) in the hypothetical case of two possible models (i.e., $k = 2$; the exceedance probability, in this case, is simple the area to the right of $r_k = 0.5$). Modified from (29).

| Task | Model | Dirichlet α | Expected Probability | Exceedance Probability | Log-Likelihood |
|---|---|---|---|---|---|
| All Tasks Combined | Common Model | 134.23 | 0.7759 | 1.000 | -3766837.56 |
| | Hierarchical Closed | 21.8 | 0.1260 | 0.0000 | -3797769.54 |
| | Hierarchical Open | 0.99 | 0.0057 | 0.0000 | -4292199.54 |
| | Hub-and-spoke BG | 14.99 | 0.0866 | 0.0000 | -4250392.03 |
| | Hub-and-spoke PFC | 0.99 | 0.0057 | 0.0000 | -4300912.00 |
| Emotion Processing | Common Model | 101.37 | 0.5307 | 1.0000 | -858282.97 |
| | Hierarchical Closed | 3.68 | 0.0193 | 0.0000 | -865011.12 |
| | Hierarchical Open | 37.27 | 0.1951 | 0.0000 | -860955.74 |
| | Hub-and-spoke BG | 37.66 | 0.1972 | 0.0000 | -861523.46 |
| | Hub-and-spoke PFC | 11.02 | 0.0577 | 0.0000 | -861878.56 |
| Incentive Processing | Common Model | 144.77 | 0.7131 | 1.0000 | -861026.48 |
| | Hierarchical Closed | 1.36 | 0.0067 | 1.0000 | -869976.35 |
| | Hierarchical Open | 39.82 | 0.1962 | 0.0000 | -866148.11 |
| | Hub-and-spoke BG | 14.8 | 0.0729 | 0.0000 | -869471.16 |
| | Hub-and-spoke PFC | 2.26 | 0.0111 | 0.0000 | -866795.76 |
| Language & Math | Common Model | 85.83 | 0.4494 | 0.7526 | -894992.94 |
| | Hierarchical Closed | 7.74 | 0.0405 | 0.0000 | -903067.62 |
| | Hierarchical Open | 76.84 | 0.4023 | 0.2474 | -895740.71 |
| | Hub-and-spoke BG | 13.65 | 0.0715 | 0.0000 | -918300.93 |
| | Hub-and-spoke PFC | 6.94 | 0.0363 | 0.0000 | -902568.43 |
| Relational Reasoning | Common Model | 77.8 | 0.4116 | 0.9505 | -746544.42 |
| | Hierarchical Closed | 8.35 | 0.0442 | 0.0000 | -749922.68 |
| | Hierarchical Open | 57.81 | 0.3059 | 0.0490 | -747385.57 |

| Task | Model | Dirichlet α | Expected Probability | Exceedance Probability | Log-Likelihood |
|---|---|---|---|---|---|
| | Hub-and-spoke BG | 42.56 | 0.2252 | 0.0005 | -672161.79 |
| | Hub-and-spoke PFC | 2.48 | 0.0131 | 0.0000 | -750719.3 |
| Social Cognition | Common Model | 141.47 | 0.7368 | 1.0000 | -712975.64 |
| | Hierarchical Closed | 1.37 | 0.0072 | 0.0000 | -720399.27 |
| | Hierarchical Open | 33.34 | 0.1737 | 0.0000 | -716591.85 |
| | Hub-and-spoke BG | 13.57 | 0.0707 | 0.0000 | -719374.15 |
| | Hub-and-spoke PFC | 2.24 | 0.0117 | 0.0000 | -715828.71 |
| Working Memory | Common Model | 142.87 | 0.7441 | 1.0000 | -113247.58 |
| | Hierarchical Closed | 46.13 | 0.2403 | 0.0000 | -112996.33 |
| | Hierarchical Open | 1.00 | 0.0052 | 0.0000 | -680320.81 |
| | Hub-and-spoke BG | 0.99 | 0.0051 | 0.0000 | -680396.45 |
| | Hub-and-spoke PFC | 1.01 | 0.0053 | 0.0000 | -679145.16 |

**Table S3**: Results of Bayesian model comparison across tasks and models. The table reports the dirichlet parameters estimated for each model's posterior distribution (see Fig. S4A), as well as the two statistics reported in this paper, expected and exceedance probabilities. For completeness, the last column also reports the corresponding log-likelihood for every model.