The following full text is a publisher’s version.

For additional information about this publication click this link.
http://hdl.handle.net/2066/197936

Please be advised that this information was generated on 2021-01-11 and may be subject to change.
Disentangling common from specific processing across tasks using task potency

Roselyne J. Chauvin, Maarten Mennes, Alberto Llera, Jan K. Buitelaar, Christian F. Beckmann

Abstract

When an individual engages in a task, the associated evoked activities build upon already ongoing activity, shaped by an underlying functional connectivity baseline (Fox et al., 2009; Smith et al., 2009; Tavor et al., 2016). Building on the idea that rest represents the brain's full functional repertoire, we here incorporate the idea that task-induced functional connectivity modulations ought to be task-specific with respect to their underlying resting state functional connectivity. Various metrics such as clustering coefficient or average path length have been proposed to index processing efficiency, typically from single fMRI session data. We introduce a framework incorporating task potency, which provides direct access to task-specificity by enabling direct comparison between task paradigms. In particular, to study functional connectivity modulations related to cognitive involvement in a task we define task potency as the amplitude of a connectivity modulation away from its baseline functional connectivity architecture as observed during a resting state acquisition. We demonstrate the use of our framework by comparing three tasks (visuo-spatial working memory, reward processing, and stop signal task) available within a large cohort. Using task potency, we demonstrate that cognitive operations are supported by a set of common within-network interactions, supplemented by connections between large-scale networks in order to solve a specific task.

1. Introduction

Advances in functional brain imaging have provided tremendous insight into the neural correlates of cognition by relating behavioural descriptions to local changes in brain activity via oxygen metabolism using functional Magnetic Resonance Imaging (fMRI). Typical experimental studies probe specific cognitive functions (Geyer et al., 2011; Zilles and Amunts, 2010), and thereby inform about the sensitivity of brain areas to the experimental manipulation of interest (Aguirre et al., 2002; Harley, 2004). Yet, single neuroimaging studies do not allow making inferences about whether an observed area exclusively responds to cognitive function A or whether it is also sensitive to manipulation of function B. As such, single studies cannot inform about the specificity of a brain area for the tested cognitive function.

To be informative about specificity rather than mere sensitivity and thus allow for reverse inference (Aguirre et al., 2002; Poldrack, 2006), study participants would need to be probed for various cognitive functions across a broad repertoire of domains. Such multi-paradigm investigations are technically and logistically challenging and therefore remain rare. Accordingly, in order to indirectly infer specific behavioural relevance for the neural responses they observed, authors typically resort to previously reported results via literature meta-analysis or alternative initiatives that validate study results (Hutzler, 2014; Poldrack, 2011; Schwartz et al., 2013; Varoquaux and Thirion, 2014). Yet, such literature-based techniques are troubled by typical biases associated with the publication process including article imprecision, 'File Drawer' issues, and potential p-value tweaking (Button et al., 2013; Poldrack, 2006).

In all of these approaches, the collection of alternative tasks directly measured or inferred upon through meta-analysis provides a functional baseline that allows defining specificity in light of estimates relative to this baseline state. Moving from a focus on localised brain activity towards covariation patterns across distributed areas (by means of functional connectivity) various metrics have been proposed that index...
processing efficiency in the brain. Quantities like clustering coefficient or average path length (Rubinov and Sporns, 2010; Wang et al., 2010) are thought to reflect inter-individual differences in the degree to which processing in the brain builds upon differences in the baseline infrastructure for information integration. Typically, however, such quantities are being derived from individual scans and do not integrate information across different cognitive tasks and/or resting sessions.

Corroborating the idea of such a functional baseline, we here propose to take advantage of the fact that resting state functional MRI (rFMRI) data exhibits dynamics that correspond to major functional activation patterns that can be observed across a vast repertoire of tasks (Smith et al., 2009). The existence of the brain’s ‘functional repertoire’ during rest supports the idea that specific cognitive states are produced by specific modulation upon a baseline of common, ongoing network activity (Bolt et al., 2017; Davison et al., 2015; Ito et al., 2017; Krienen et al., 2014), rather than being orchestrated by independent activity in single regions. An important corollary is that cognitive function emerges through embedding unique regional activity within the context of larger network processes, as described in the ‘massive redeployment hypothesis’ (Anderson, 2007; Lloyd, 2000). This hypothesis is further supported by research showing that functional connectivity can successfully differentiate between mental states (Shirer et al., 2012), corroborating the idea that mental states can be deconstructed through embedding unique regional activity within the context of larger network processes, as described in the ‘functional repertoire’ (Cole et al., 2016; Yeo et al., 2011). In accordance with these ideas, we propose to utilise rFMRI-derived connectomes as a functional baseline, effectively representing a standard space to compare task modulation requirement across tasks using task potency, a measure that indexes task-related connectivity modulations away from this functional baseline.

Our framework fits in the context of emerging large cohort functional imaging studies that involve multiple experimental fMRI designs along with rFMRI measures, allowing for within-subject comparisons between cognitive paradigms relative to the resting condition. One prominent example is the Human Connectome Project (Glasser et al., 2016; Van Essen et al., 2012) aimed at deciphering the complex relationship between brain functions, cognition, and the functional and structural human connectome within a normal cohort. Similar projects are translating these efforts to the clinical domain (e.g., NeuroIMAGE (von Rhein et al., 2015a), PNC (Satterthwaite et al., 2016), ABCD (Bjork et al., 2017)).

We aim to capitalize on the increased statistical sensitivity to standardized modulations differences that such within-subject designs offer. In this context, the task potency metric can be used to characterize effect-size differences in light of experimental manipulation against a baseline that represents that participant’s full functional spectrum. Importantly, we utilise this concept to disentangle general from task-specific connectivity modulations by indexing the presence or absence of significant functional connectivity under different tasks and posit that general neural modulations occur across multiple tasks yielding limited differences in potency between tasks. In contrast, specific neural modulations that might be attributed to a single cognitive process, will yield large differences in potency between tasks or sets of tasks. We demonstrate the application of task potency to multi-subject fMRI data involving – along with a resting-state fMRI session – three different experimental tasks probing different aspects of cognition. Using population distributions of task potency, we assess task-specificity of edges in the connectome and compare the similarity of the ensuing task fingerprint across different cognitive domains.

2. Methods

2.1. Participants

We use MRI data from the NeuroIMAGE sample (N total > 800 participants; see von Rhein et al., 2015b). In the current analyses we included data from healthy control participants only (initial N = 385) who each performed at least one of the following tasks during fMRI scanning: response inhibition (Stop Signal Task (STOP) (Logan et al., 1984; von Rhein et al., 2015a; van Rooij et al., 2015)), reward processing (REWARD (Hoogman et al., 2011; Knutson et al., 2001; von Rhein et al., 2015a; von Rhein et al., 2015b)), spatial working memory (WM (van Ewijk et al., 2015; Klingberg et al., 2002; McNab et al., 2008; von Rhein et al., 2015a)) (see Supplementary Table 1). In addition to the task-based MRI scans, each participant completed a task-free resting state fMRI scan. All participants also completed a T1-weighted anatomical scan for registration purposes. MRI acquisition parameters are shown in Table 1.

Functional scans exhibiting limited brain coverage or excessive head motion were excluded from further processing. Limited brain coverage was defined as having less than 97% overlap with the MN152 standard brain after image registration. Applying this criterion excluded 47 subjects (details in Table 1). In addition, we excluded from each task those participants who were among top 5% in terms of head motion as quantified by RMS-FD, the root mean square of the frame-wise displacement computed using MCFLIRT (Jenkinson et al., 2002). After applying these criteria, we selected only participants that completed at least a task and a resting state scan resulting in the inclusion of data from 218 healthy controls, comprising 218 resting state acquisitions, 111 STOP acquisitions, 123 REWARD acquisitions, and 147 WM acquisitions. Participants ranged in age between 8.6 and 30.5; mean = 17.0; sd = 3.5; 45.9% were male.

Table 1

| MRI acquisition parameters and participant characteristics (demographic information is presented in Supplementary Fig. 1). The final characteristics (RMS, age, sex) are based on those participants that were entered into the task potency analyses. |
|---------------------------------|---------------------------------|---------------------------------|----------------|
| T1-weighted structural MRI parameters | rFMRI | STOP | REWARD | WM |
|---------------------------------|---------------------------------|---------------------------------|----------------|
| General parameters | TE = 40 ms, FOV = 224 mm, 37 axial slices, flip angle = 80°, matrix size = 64 × 64, in-plane resolution = 3.5 mm, slice thickness/gap = 3.0 mm/0.5 mm | | | |
| N volumes | >260 | >300 | 107 * 4 blocks | |
| TR in ms | 1960 | 2340 | 2340 | 2340 |
| N first volumes rejected | 5 | 4 | 5 | 3 |
| Number of Participants | N initial | 302 | 239 | 256 | 266 |
| N rejected for limited brain coverage | 73 | 4 | 8 | 6 |
| N rejected for head motion coverage | 11 | 11 | 12 | 13 |
| Final N (with a resting state scan) | 218 | 111 | 123 | 144 |
| Selected Participant Characteristics | | | | |
| RMS-FD min-max | 0.026–1.930 | 0.025–0.976 | 0.027–0.976 | 0.029–1.504 |
| RMS-FD (std) | 0.171 | 0.093 | 0.145 | 0.149 |
| (std) | (0.224) | (0.098) | (0.145) | (0.191) |
| Age min-max | 8.6–30.5 | 8.6–27 | 9.1–30.1 | 8.6–30.1 |
| Age mean (std) | 17 (3.5) | 17.4 (3.6) | 17.6 (3.7) | 17.4 (3.6) |
| % male | 45.5% | 45.5% | 44.5% | 47% |

* Except 1 scan which was done at 1860 ms in rFMRI, 1 at 2150 ms in STOP, and 1 at 2280 ms in REWARD. * The number of initial volumes removed from further analyses varied to ensure comparability with earlier studies that used these data. This variation will have very limited impact on the current analyses.
2.2. fMRI preprocessing

All fMRI acquisitions were processed using tools from FSL 5.0.6. (FSL; http://www.fmrib.ox.ac.uk/fsl (Jenkinson et al., 2012; Smith et al., 2004; Woolrich et al., 2009)). We employed the following pipeline: removal of the first volumes to allow magnetization equilibration (see Table 1), head movement correction by volume-realignment to the middle volume using MCFLIRT, global 4D mean intensity normalization, spatial filtering with a 6 mm FWHM Gaussian kernel. We then denoised all preprocessed data for motion-related artefacts. We used ICA-AROMA to detect motion-related artefacts in single-subject data through classification of ICA components extracted by MELODIC. We subsequently regressed components identified as motion-related artefacts out of the data (using fsl_regfilt, see Beckmann, 2012; Pruim et al., 2015a, 2015b). Subsequently, we regressed out mean signals from CSF and white matter extracted using participant-level masks obtained by multiplying – in the

Fig. 1. ICP atlas with 179 areas represented in their corresponding top-level networks. R_attention: right attention network; L_attention: left attention network; DMN: default mode network; sub cort: subcortical regions; cereb: cerebellum. Supplementary Fig. 2 shows the higher-level networks projected to the brain surface.
participant native space – participant-level CSF and white matter segmentations with the MN152-based CSF and white matter masks provided as part of FSL. Finally, we applied a 0.01 Hz temporal high-pass filter (Gaussian-weighted least square straight line fit to the data). For each participant, all acquisitions were registered to its high-resolution T1 image using Boundary-Based Registration (BBR) available in FSL FLIRT (Jenkinson and Smith, 2001; Jenkinson et al., 2002). All high-resolution T1 images were registered to MN152 space using 12-dof linear registration available in FLIRT and further refined using non-linear registration available in FSL FNIRT (Anderson et al., 2007). Transformations were not applied. Instead we used the inverse of the obtained transformations to bring a hierarchical atlas of brain regions to the participant's native space (see below).

2.3. Connectome atlas

For each functional imaging scan we defined connectivity matrices using regions defined in a hierarchical whole-brain functional atlas (van Oort et al., 2017). This atlas contains 185 non-overlapping regions and was defined through Instantaneous Connectivity Parcellation (ICP (van Oort et al., 2017)) as applied to resting state fMRI data of 100 participants of the Human Connectome Project (HCP (Glasser et al., 2016; Van Essen et al., 2012)). In short, ICP aims to parcel larger regions into sub-regions based on signal homogeneity, where the optimal number of subregions is determined based on split-half reproducibility at the cohort level.

Fig. 1 illustrates the hierarchical brain atlas, where areas were grouped in 11 higher-level networks: 9 resting state networks (visual1, visual2, motor, right attention, left attention, auditory, default mode network (DMN), fronto-temporal and cingulum), and 2 networks based on anatomical structures, i.e., the subcortical areas, and the cerebellum. These higher-level networks respectively contained 19, 12, 22, 22, 18, 8, 18, 13, 7, 23, and 23 subregions, resulting in a total of 185 initial parcels. Connectivity matrices were calculated in each participant's native space for each of the functional scans. To this end we transformed the atlas to each participant's native space using the inverse of the anatomical to MN152 non-linear warp, and the inverse of the linear transformation of the functional image to the participant's high resolution anatomical image. Atlas areas that were on average across our population >50% outside of the brain were rejected from further analyses. As a result, we used 179 areas, shown in Fig. 1 colour coded by their associated top-level network, to compute connectivity matrices.

To assess dependence of our results on the network grouping within our ICP atlas, we replicated our analyses using the 7 networks as described in Yeo et al. (2011) as top-level networks. We included all areas overlapping at least 50% with one of the Yeo networks. This selection resulted in a total of 77 areas across the 7 networks and divided as follows: 13 (visual), 10 (somatomotor), 9 (dorsal attention), 9 (ventral attention), 9 (frontotemporal), 18 (default), 9 (limbic). Areas in the ICP atlas that did not sufficiently overlap with the Yeo networks were removed from the replication analysis. Results of these analyses are included in the supplement.

2.4. Connectivity calculation

For each participant and each task (fMRI, WM, REWARD, STOP) we calculated 179 × 179 connectivity matrices, by cross-correlating the time series of all regions in the atlas. We obtained each region's time series through multivariate spatial regression, using all 179 regions as regressors and each task's preprocessed time series as dependent variable. The resulting regional time series were demeaned. Using these time series, we calculated 179 × 179 partial correlation matrices through inverting covariance matrices estimated by the Ledoit-Wolf normalization algorithm (Brier et al.; Ledoit and Wolf, 2004) as implemented in nilearn (http://nilearn.github.io/). The Ledoit-Wolf normalization algorithm is a shrinkage algorithm that optimises estimation of the covariance matrix and ensures sparseness. We opted for partial correlations in order to avoid redundancy in the functional connectivity estimation, thereby allowing making inferences about direct connections without the influence of indirect connections in the fingerprint. Results obtained using full Pearson correlation are presented in the supplement. Finally, all pair-wise correlations were Fisher r-to-Z transformed.

To allow comparing connectivity values between acquisitions and to account for potential differences in temporal degrees-of-freedom due to scan length differences, we normalized the distribution of connectivity values within each connectivity matrix using a mixture-modelling approach (Beckmann et al., 2005; Bielezyk et al., 2018; Feinberg et al., 2010). Note, that this approach allows to correct for differences in task-specific parameters including the number of volumes or differences in TR. In this mixture modelling approach, we fit three parameterised distributions to the histogram of connectivity values: a central Gaussian distribution representing the noise and two gamma distributions on each side of the central Gaussian that represent the signal as the tails of the data distribution. We fit this mixture of distributions under the assumption that evidence for a non-zero connection is unrelated to the spatial location of the nodes and that non-zero connections are sparse. Further, we assume that there is a sufficient total number of nodes so that the distribution of values for non-significant edges (i.e., noise) in the network can be used to estimate the within-subject null distribution of non-existing connections. In practice, we modelled the obtained connectivity values per task using a Gaussian-gamma mixture model and used the main Gaussian, i.e., the one fitting the body of the distribution, to normalize our connectivity values. Note that this overcomes any issue in deciding on the appropriate degrees-of-freedom for the Fisher r-to-Z conversion of partial connectivity values. We applied mixture modelling to each connectivity matrix and subsequently normalized the connectivity values by subtracting the mean and dividing by the standard deviation of the obtained Gaussian model. As a result, and despite differential loss in temporal degrees-of-freedom due to the partial correlation calculation, the values within the normalized, Z-transformed partial correlation matrices are readily comparable across participants and tasks.

Finally, to allow interpretation of the task-based connectivity matrices in terms of their deviation from a functional baseline defined as connectivity during the resting state, we further standardized each participant's task-based connectivity matrix. Specifically, we standardized each individual-level pair-wise correlation obtained during task by subtracting the corresponding individual pair-wise correlation obtained during rest for the same participant. As such, each task-based pair-wise correlation or edge quantifies how connectivity for that edge differed from that edge's connectivity during the resting state. As a result, after standardization, we obtain for each participant an individual connectivity matrix for each of their task acquisitions. We refer to these matrices as task potency matrices, which quantify for each edge how strongly the task-based connectivity was modulated away from its resting state baseline (i.e., the amplitude of the task-based modulation). For each task, we finally create group-level task potency matrices by averaging across participant matrices and multiplying by the root mean square of the number of participants to avoid bias in between-task comparisons related to the number of observations in each task. All scripts needed to compute task potency and ensuing analyses are available via https://github.com/roscha/task_potency.

2.5. Task-based fingerprints

To compare those connections that characterize a task's functional fingerprint across different tasks we selected, for each task, those edges that showed significant task potency. Similar to the normalization procedure described above we used the mixture modelling approach to determine a significance threshold (Bielezyk et al., 2018). After applying the mixture modelling to the task potency distribution, we use the main Gaussian to estimate the false discovery rate according to the density of
the Gaussian distribution and the ratio of connections above a certain threshold (Efron, 2007). We use an overall FDR of 0.05 to detect significant connections. To be able to estimate the corresponding task potency threshold for each side of the distribution, instead of applying an FDR of 0.025 to each side, we weighted the unilateral FDR by the size of the gamma distributions estimated by the mixture model (see formula in supplement). This weighted FDR allows us to account for asymmetry of the signal and avoids over- or underestimation of positively or negatively potentiated connections. We refer to edges with significant task potency as task-based fingerprints. The task-based fingerprints are subsequently used to define the task sensitivity and task specificity of each edge in the fingerprint. Fig. 2 provides a schematic overview of the task potency framework, further details can be found in Supplementary Table 2 and the supplementary description of the methods.

To enable statistical testing on our analyses (see results section 3.2 and 3.3), we estimated the task fingerprint 10000 times using 80 percent of the population randomly selected. This bootstrap procedure allowed estimating the variance of the task fingerprints for significance testing.

2.6. Task potency, sensitivity, and specificity

Fig. 3 illustrates how we can characterize each edge within a task potency matrix in terms of its sensitivity and specificity to the different tasks included in the study. An edge was regarded sensitive to task modulation when the strength of connectivity was above the statistical threshold in at least one of the tasks. An edge can be sensitive to modulation by several tasks, yet with a different level of potency in one task compared to another. This differential potency is not considered when assessing sensitivity.

Task specific edges were those edges that were selected for one task only. Common edges of our three tasks were defined as edges selected in all three tasks. Note that specificity and commonality are determined by the collection of available tasks, accordingly conclusions regarding the specificity of edges in the current manuscript need to be interpreted in light of the tasks we included. Also note that design choices made for the included tasks, e.g. shorter task duration and/or fewer trials within a task, will influence the signal-to-noise (SNR) properties of a given task and thereby influence the degree to which edges in the connectome become up- or down-regulated relative to the resting-state scan. While our approach to connectome-wide task fingerprinting is not aimed at adjudicating between experimental design choices, we note that it is effectively possible to assess the connectome-wide impact of such choices using our potency approach.

To differentiate which task most strongly modulated which connection, we differentiated tasks based on their potency in connections regardless of a sensitivity threshold. To this end, we assessed for each edge whether we could differentiate the tasks based on the average potency across the population using a measure of anisotropy across the three tasks, calculated as: (highest potency – second highest potency)/sum of potency across the three tasks. Note that this comparison relies on the normalization of all matrices prior to standardization by the resting state connectivity (see Supplementary Fig. 3). We displayed the anisotropy measure using colour gradients where values close to 0 are light, i.e., tasks could not be differentiated based on differences in their potency. In contrast, highest anisotropy values will appear darker, and in the colour associated with the task that exhibited

Fig. 2. Analysis framework to obtain connectivity-based task fingerprints. The framework starts at the participant-level with obtaining a partial correlation matrix (Fisher-Z transformed), which is normalized, and subsequently standardized by that participant’s resting state connectivity (subtraction of baseline), resulting in individualized task potency matrices. A group task fingerprint can be obtained by averaging the individual task potency matrices and thresholding based on the z-score of the group potency. More details about the used algorithm are available in Supplementary Table 2.
the highest potency. We used the same approach to label maximum task-related up- or down regulation in connectivity in the 11-network framework. Here we used the sum of potency across selected edges for each network. Finally, we investigated maximum task potency at the level of areas (i.e., columns in our connectivity matrices). To this end, we summed potency for group-level edges across each area’s 179 connections.

2.7. Network-based summary metrics

The hierarchical ICP atlas defines 179 areas as subdivisions of 11 large-scale networks. Accordingly, next to reporting at the level of individual areas, we can average across edges within each network to summarize potency, sensitivity, and specificity at the network level. We can differentiate edges that link areas within a network (within-network edges) from edges that link areas between two different networks (between-network edges). In order to compare between networks, we corrected for the number of edges averaged over the different 11 × 11 interactions by multiplying each average by the root mean square of the number of edges within or between two networks. By comparing the within- and between-network connections we assessed whether a task was associated with specific networks or resulted in an overall, diffuse modulation of connectivity. In practice, to derive network-level scores, we calculated the percentage of selected edges included in each network. This was done for each entry in the 11 × 11 network connectivity matrix, and allowed quantifying the selection of edges at the within- (diagonal matrix entries) and between-network (off-diagonal matrix entries) level.

Fig. 4. Illustration of connectivity matrix calculation for the reward task. A: Normalized resting state Z partial correlation averaged across the population; B: Normalized Z partial correlation for the reward task averaged across the population; C: Average reward task potency across the population. Upper triangle displays the 179 × 179 connectivity fingerprints; lower triangle displays the average summary per network. R_attention: right attention network; L_attention: left attention network; DMN: default mode network; sub cort: subcortical regions; cereb: cerebellum. The normalized Z partial correlation and task potency matrices for the two other tasks are display in Supplementary Fig. 4.
2.8. Reproducibility of the edge selection procedure

Every single participant has associated within-subject differences relative to the cohort-derived group potency fingerprint. In order to assess reproducibility of our group-level fingerprint pattern, we defined individual task fingerprints, applying the same selection procedure as above, but applied to the individual task potency matrices (i.e., select those edges with a pFDR \( < 0.05 \) in the individual task fingerprints). This

Fig. 5. Radar plots of the percentage of edges showing significant task potency (i.e., sensitivity) summarized across 11 brain networks. When splitting the percentage of sensitive connections (top row) into within- (middle row) and between (bottom row) network connections, we observed a larger percentage of sensitivity for within-network connections, compared to between-network connections. As an example, 58.7% percentage of edges within the motor network exhibited sensitivity, compared to only 1.76% of its between-network connections. To allow direct comparison between both radar plots, we also show the between-network percentages on top of the within-network percentages in the middle left plot. Bar plots on the right illustrate edge sensitivity for each task. For further details regarding the percentage calculation we refer to Supplementary Fig. 5. R_attention: right attention network; L_attention: left attention network; DMN: default mode network; sub cort: subcortical regions; cereb: cerebellum.
enables us to quantify subject-specific variations in the edge selection and thereby permits quantification of reproducibility across participants. We indexed the number of times an edge was selected across participants. This proportion is interpreted as the reproducibility of an edge’s potency.

3. Results

3.1. Task-based fingerprints

To evaluate connectivity sensitivity to each task we created task-based fingerprints by standardizing the task connectivity by the resting state connectivity, resulting in a matrix quantifying each edge’s functional potency, as illustrated in Fig. 4C for the reward task (see Supplementary Fig. 4 for the other tasks). This task-based fingerprint served as the basis to identify sensitive edges and assess their task specificity.

3.2. Task sensitivity

Significance of task sensitivity was determined using mixture-modelling thresholding applied to each task’s functional fingerprint. We applied the mixture-modelling thresholding on the average potency across participants. From these we calculated both the relative proportion of selected edges (above the mixture-modelling threshold) as well as the normalized proportion of edges per individual task. Further, grouping the parcel-wise estimates into the 11 large-scale networks we quantified the percentage of sensitive connections within and between the large-scale networks (Fig. 5).

Across tasks and networks, 5.37% of all edges exhibited significant sensitivity to task modulation (Fig. 5, top row). When comparing the percentage of sensitive edges across tasks we observed that WM potentiated significantly more connections compared to REWARD (4.2% (sd = 0.66) vs. 1.9% (sd = 0.54); \( p < 0.05 \)). In turn, REWARD significantly potentiated more connections than STOP (1.9% (sd = 0.54) vs. 1.1% (sd = 0.19); \( p < 0.05 \)).

When further differentiating between edges that connect regions within networks versus edges that connect between regions in two networks, we observed a higher prevalence of sensitive edges for within-network connections compared to between-network connections. In particular, the visual and motor networks showed a high percentage of within-network sensitive edges, both networks include the primary sensory areas needed to process the task information (Fig. 5, middle row). In contrast, the number of sensitive between-network connections was considerably lower, with on average only 1.55% of edges selected across the 11 networks versus 25.44% of within-network connections. At the between-network level, sensitive edges were relatively equally distributed across the 11 networks (Fig. 5, bottom row). We highlight the result for the cingulum network which exhibited the highest percentage of between-network sensitivity and the lowest within-network sensitivity.

Across networks, the relative distribution of sensitive edges per task (task proportion) showed little variation, with the exception of the within-network DMN connections, where STOP did not yield any sensitive within-network edges (see stacked bars in Fig. 5).

At the network level, we observed that compared to the other networks, the motor, visual and cerebellum networks exhibited significantly stronger within-network potentiation, while the cingulum network exhibited lower within-network potentiation and stronger between-network potentiation compared to the other networks in our analysis (Bonferroni corrected \( p \)-values across each pair of networks are reported in Supplementary Fig. 6). The higher level of significant within-network connectivity supports the hypothesis that the brain strongly segregates information at the level of individual networks, while more weakly integrating information between networks, in line with theoretical predictions using integration and segregation to model the dynamic of brain networks (Deco et al., 2015; Jirsa et al., 2014).

3.3. Task specificity

To disentangle overlapping connectivity modulations in light of the included tasks, we defined the task specificity of edges by splitting the collection of sensitive edges into those that were modulated by one task only (i.e. are specific to a particular task), those that were sensitive to modulation by several (but not all) tasks, and those that were significantly modulated by all tasks (see also Fig. 3). Fig. 6 illustrates the percentage of sensitive edges modulated by one task only and those modulated by all tasks. We observed that overall 68.85% of the sensitive edges were specific to a particular task, compared to 12.8% that were modulated by all tasks (we refer to these as ‘common’ edges). Note that this also means that 18.35% of sensitive edges was modulated by more than one, yet not all, tasks.

We observed a difference in the level of specificity for within-vs between-network connections. Comparing the dark versus light coloured areas in the top row of Fig. 6 it is evident that the ratio of specific versus common connections was smaller \( (t = -3.82; \ p < 0.05) \) for the sensitive within-network connections (mean ratio across networks = 3.21 ± 1.96) compared to the ratio of specific versus common connections for the sensitive between-network connections (mean ratio across networks = 11.74 ± 9.59). This result shows that between-network connections are almost exclusively modulated in a specific fashion, where different tasks modulate different edges connecting networks to the rest of the brain.

We further characterised the nature of the task-specific connections to assess how specificity is distributed across tasks and networks (Fig. 6 bottom). While the between-network connections were more homogeneously distributed between tasks and across all networks, we observed greater variation in the specificity of the within-network connections, with some networks showing notable task specificity. We observed that REWARD showed more specifically potentiated connections involving subcortical regions, while STOP showed a limited amount of between-network specificity, yet strongly potentiated connections among subcortical regions. In contrast, edges specific to WM were equally distributed across networks, suggesting an extensive involvement of different networks, corroborating the observation that WM overall potentiated more connections. Indeed, in Supplementary Fig. 10 we demonstrate that placing more stringent thresholds to determine sensitivity enables to capture network specificity in WM, showing that the edges exhibiting strongest potency are within the motor network and the DMN. Importantly, the relation of specificity to the amplitude of potentiation further supports the idea that potency should be considered as a continuum instead of defining a threshold of significance (see section 3.4 for further results).

In contrast to task-specific edges, about 13% of all task-sensitive edges were modulated by all three tasks (union across all tasks in Fig. 3; dark line in Fig. 6, top row). Brain regions that yielded the highest number of common edges are represented in Fig. 7. Apart from visual and motor regions where we expected shared modulation, as all three tasks were using visual stimuli and requested motor response, all tasks modulated edges involving regions that were part of the fronto-temporal and attention networks in our atlas. This modulation included anterior cingulate cortex, left inferior frontal gyrus, areas from inferior parietal lobe, bilateral frontal orbital cortex extending into Broca’s area, the temporal pole, amygdala, and entorhinal cortex. Interestingly, no DMN or subcortical regions were represented in the top connections of areas that potentiated edges across all tasks. Brain regions that showed the highest number of edges modulated by each task specifically are shown in supplementary figure 7 (STOP), 8 (REWARD) and 9 (WM).

3.4. Reproducibility of the selection across individual fingerprints

The result in Fig. 6 shows that different tasks exhibit a specific pattern of network potentiation, which can be accessed by comparing a set of
different tasks. Nevertheless, a large proportion edges are sensitive to all tasks. In order to establish the utility of evaluating individual task fingerprints in a reproducible manner, we studied the detection rate of edges sensitive to one or all tasks across task fingerprints obtained for individual participants. Specifically, we investigated whether the group selection was reproducible at the individual level and in particular how well the task-specific connections where represented at the level of individual participants. We defined the individual fingerprint by selecting edges that showed a pFDR below 0.05 using the mixture modelling thresholding on the individual-level task fingerprints. We computed the sum of selected edges across the population for each task. As shown in Fig. 8, we observed a set of edges with high selectivity across participants for each task: 2.3% of edges within the union of individual masks were selected by minimum 13.7% of the population and in each of the three tasks. These edges mainly linked homotopic areas of each hemisphere, including bilateral motor areas, cerebellum, attention networks, visual areas, and bilateral putamen (See Fig. 9).

In contrast to the set of highly selected edges at the individual level, we note that the edge selection at the individual level showed substantial variability: 80% of sensitive edges were selected in less than 12.6% of subjects (Fig. 8 dashed line). As indicated above, the most consistently selected edges between participants involved connections sensitive to all tasks. In contrast, highest inter-individual selection variability was found for task-specific edges as edges selected in only one task at the group level showed a lower individual selection reproducibility than edges selected in all tasks at the group level (Fig. 9).
3.5. Differentiating tasks based on potency

Even if each task’s set of specific connectivity modulations reflects network enlistment by a specific experiment, this information is only accessible in comparison to other tasks. Additionally, to define edges as being task-specific brings limitations as soon as the number of tasks increases or when tasks share cognitive processes that are differentially involved in the experimental design. Therefore, we propose to move from the binary concept of specificity and sensitivity to a continuous measure of connectivity potency that quantifies the amplitude of the connectivity modulation required to enlist a connection, a network, or an area under certain task processes.

Using potency as a quantitative measure of the strength of enlistment of connections in a task, we can characterize which task potentiated an edge, a region, or a network most strongly. To this end, we assessed whether modulation of each edge could be attributed to one of the three tasks in our comparison by computing a measure of anisotropy between them. We investigated this potency anisotropy of each edge separately (Fig. 10, upper triangle of the matrix) and observed that only few connections were modulated most strongly by one task compared to both other tasks. Across all edges, the anisotropy was on average 0.25 (sd = 0.2) suggesting a relatively equal representation of tasks. At the brain region (Fig. 10 brain slices) and network-level (Fig. 8, lower triangle of the matrix), each task displayed a specific pattern of strongest potentiation across the brain (Fig. 10B). REWARD principally potentiated the fronto-temporal network as well as areas from the reward circuit (anterior cingulate cortex, prefrontal areas, thalamus). Whereas STOP most strongly potentiated connections with the visual 1 network and with areas in motor cortex. Finally, WM potentiated regions included in the DMN as shown in Fig. 10B.

4. Discussion

When an individual engages in a task, the associated evoked activities build upon the brain’s ongoing activity, itself shaped by an underlying functional connectivity baseline (Fox et al., 2009; Smith et al., 2009; Tavor et al., 2016). Here, we show how this functional baseline architecture can be used to index task-dependent modulations, providing a means for quantitatively comparing evoked effects across tasks and
cognitive domains. This model incorporates the idea that functional connectivity observed under cognitive manipulation is task-specific with respect to its underlying resting state functional connectivity (Cole et al., 2014; Geerligs et al., 2015; Shirer et al., 2012; Smith et al., 2009). To facilitate understanding the building blocks of cognition, we demonstrate that differential levels of localised sensitivity to task manipulation inform about the relative potency of a specific task.

In this regard, task potency could be interpreted as indexing the resources required to modulate away from the brain's functional baseline in order to perform a task. As such, task potency provides a novel index of efficiency. By comparing task modulation away from a common baseline acquired in the same individual, task potency bridges between a traditional seed- or ICA-based connectomic description and derived measures of efficiency provided by graph theory analysis. As such task potency provides a context to interpret brain fingerprint modulations across tasks at the whole brain level.

We calculated task potency for three tasks (working memory, response inhibition, and reward processing) in a large healthy population and showed that all tasks predominantly potentiated edges at the within-network level, i.e. connecting areas within networks, particularly those including lower-order sensory-motor regions. Such larger connectivity changes in primary sensory networks may highlight a more straightforward and automatic response to incoming stimuli, accompanied by standardized motor activity. This fits with the idea that visual and motor areas adhere to a highly constrained organization that is strongly evolutionary conserved, resulting in lower inter-individual variability (Mueller et al., 2013), but higher within-subject flexibility (Laumann et al., 2015).

Comparing tasks across multiple distinct cognitive domains allowed us to distinguish connections that were specific to each task versus those common to all manipulations.Fig. 6 illustrates how the edges that were specific to each of our three tasks were distributed across the higher-level networks. The largely similar distribution of percentages shown for WM indicates that functional connectivity modulations induced by WM did not display strong network specificity when applying a nominal threshold, illustrating the overall strong potentiation required to perform WM. When applying more stringent thresholds (see Supplementary Fig. 10) we observed that WM exhibited highest potency in DMN, in accordance with the idea that DMN areas are involved in working memory (Piccoli et al., 2015; Pyka et al., 2009). To access the areas involved in modulation, we summarized the number of selected edges per area and describe the highest 10% of them in supplementary figures 7, 8 and 9. Supplementary Fig. 9 illustrates the high specificity of ventral and dorsal pathway connectivity in WM (Baddeley, 2003). By comparison, STOP showed specific modulations involving areas typically observed in the inhibition networks (see Supplementary Fig. 7; van Rooij et al., 2015), while REWARD specifically modulated putamen connectivity (see Supplementary Fig. 8). Next to task-specific modulations, motor, visual, and higher-order cognitive regions including temporo-frontal areas showed sensitive yet unspecific involvement across multiple tasks (Fig. 7). This result suggests that while our tasks probed different cognitive domains, they did tap into similar cognitive
processes resulting in similar connections exhibiting significant task potency. This is not surprising however, given that the three tasks included in the current study all loaded relatively high in terms of the amount of cognitive control needed to solve them successfully. Fig. 7 illustrates this by showing that the commonly modulated edges loaded primarily on areas implicated in the cognitive control network as described by Cole and Schneider (2007).

While task potency can be used in a binary way, it effectively indexes the amplitude of task-induced connectivity modulations. Accordingly, task potency can be regarded as a continuum of potencies across different tasks per connection. Here, we investigated different cognitive loads by computing the ‘most potent task’ per connection. In line with literature (Fedorenko, 2014; Stiers et al., 2010) WM proved to be the most requesting task in our set in terms of absolute potency. In contrast, STOP seemed to be the least potent task in our study as less connections were selected compared to both other tasks, yet supported by specific potentiation of connections linked to the visual 1 and subcortical networks (see Fig. 10). REWARD most potentiated brain regions and networks known to be part of typical reward circuitry (see Fig. 10 and Supplementary Fig. 5). While we compared between tasks in the current study, a task potency continuum can also be obtained in relation to variation in cognitive load within a given task. Including such task designs would allow investigating the link between potentiation of connectivity and cognitive complexity.

Our task potency model is based on the idea that task activity builds on the brain’s inherent functional architecture as captured in large-scale resting state networks (Beckmann et al., 2005; Kelly et al., 2008; Mennes et al., 2010; Smith et al., 2009). Here, we showed preferential modulation of connections within those large-scale networks, while the limited number of modulated between-network connections exhibited greatest task specificity. This observation is consistent with the hypothesis that local processing is supported by out-of-network connections during task performance (Gratton et al., 2016). The idea that the resting functional architecture provides a common baseline further supports the need for a full and independent resting state acquisition to allow capturing the full acquisition-based task potency. Therefore, when investigating edge selectivity and its variability across bootstraps, we observed that the selectivity in the residual-based task potency is lower compared to the full acquisition-based time series (see Supplementary Fig. 15). This shows that regressing out the design reduces the stability of the edge selectivity by removing (part of) the functional connectivity modulations induced by the task.

These observations suggest a need to further investigate task potency around task trials to better understand connectivity modulation mechanisms and to study specific cognitive process by assessing specificity within a task design using task contrasts as is done in activation analyses. This could be approached by using beta time series that correspond to a concatenation of regression coefficients for each of the trials of a specific event type (see Mennes et al., 2013 for an application; Rissman et al., 2004). We did not perform this analysis due to the limited number of trials, which did not provide the required statistical power. Additionally, we note that, as it is the case for activation analysis, the task potency will depend on task design choices such as the number of trials in our tasks, their interval, or the task length. Therefore, in this manuscript, we explicitly interpret the specificity across the three tasks available in the NeuroImage database without extrapolation to other (not included) tasks and cognitive constructs. More investigation on the impact of such design choices on the observed modulations is required. In this regard, using task potency to compare a task’s ability to probe cognitive domains is possible as presented in Fig. 10. Yet, using the potency framework to make design optimization choices is not particularly beneficial compared to other available tools such as NeuroDesign (Durnez et al., 2018) or FILM (Woolrich et al., 2001) to estimate design efficiency as the task potency framework requires sufficient statistical power to reliably compute functional connectivity modulations. In this context, we would also like to refer again to our supplementary analyses that used the time-series with the task design regressed out. The results from this analysis show that apart from its stability, overall task potency is not affected by removing the task design, and thus largely unaffected by design choices. These results corroborate analyses by Cole et al. who used residual task time series to investigate the brain’s evoked functional architecture (Cole et al., 2014).

In the context of reverse inference investigations, the opportunity to compare tasks in a standardized space can also be a means to resolve and quantify how specific a given task-activation pattern is for a cognitive function. Current implementations typically rely on mining available literature (Hazlitt, 2014; Polack, 2011; Schwartz et al., 2015; Varoquaux and Thirion, 2014) to compare a task modulation amplitude across a multitude of tasks in order to infer task specificity for a certain region. In contrast, our approach is to compare each task against a common resting baseline that can effectively be regarded as a superposition of the brain’s full functional repertoire (Smith et al., 2009) allowing implicitly comparing tasks against each other. This allows assessing connectivity specificity while avoiding potential literature bias or strong a-priori models (study design, HRF response), albeit at the cost of being restricted to the typically smaller number of within-study cognitive domains being probed.

While offering a framework to study task fingerprints and connectivity specificity, we did observe variability in task potency across individuals. Task-related functional connectivity yields potential to understand individual variability in performance or task-related individual markers, as it has been successfully used to categorize tasks (Shirer et al., 2012), to predict performance (Cole et al., 2016), or to predict task-induced activity (Tavor et al., 2016). In our investigation, we were unable to find strong associations linking task potency to task performance. Yet, supplementary figures 11 (STOP), 12 (REWARD) and 13 (WM) illustrate that when correlating task potency and a corresponding task performance value that edges showing the strongest behaviour-potency correlations were in fact linked to task-relevant areas. Moreover, understanding task-related modulations might enable predicting functional connectivity in individuals that deviate from the norm, e.g., in a pathological response or using an alternative strategy to perform a task. Indeed, defining task potency relative to an individual resting state baseline is relevant for clinical applications where we cannot assume that individuals with pathologies have similar baseline architectures as healthy control participants. As such, task potency might prove an interesting feature for cohort stratification, e.g., within the framework of normative modelling (Marquand et al., 2016), aimed at characterizing how individual participants differ from a large normative range in regard to multiple brain-behaviour relationships. We are currently investigating the case of Attention Deficit and Hyperactivity Disorder and Autism Spectrum Disorder and the benefit of using the task potency to study these populations. As an initial example, we investigated the effect of age on the development of potency amplitude in a separate manuscript (Chauvin et al., 2017). A similarly large age range was included in the current study. Accordingly, in light of the developmental effective as described, we here verified that the current results, which focused on sensitivity and specificity, were not driven by age. To this end, we replicated our results including only participants age 16 and older. The strong reproducibility of our results in this restricted age group is evident in Supplementary Fig. 14.

Finally, two methodological considerations in light of our potency approach ought to be discussed. First, we chose to calculate task potency using partial rather than Pearson (full) correlations between regional
time courses. Pearson correlation is often used to study functional connectivity. It has the drawback of potentially including redundant information across edges as shared variance is not excluded. Using partial correlation allowed indexing direct connectivity between areas within our atlas, thereby facilitating to detect specificity. For comparison, we also report the analyses using full Pearson correlation in the supplementary material. Results are presented in Supplementary Fig. 14 and yield conclusions that are consistent with the main analysis. However, we did observe that the ensuing results exhibit decreased specificity, in light of an overall increase in sensitivity (i.e., more edges were found to be modulated by task), yet with common edges being mainly related to motor and primary visual areas. Second, any interpretation of connectivity findings is inherently dependent on the regions used to build the connectome. Here, we used 179 regions that were part of a hierarchically defined atlas. However, this atlas is not purely function-based, as the subcortical and cerebellum network masks were anatomically defined. To verify that our results were not driven by our network definition, we repeated the analyses using an alternative higher-level grouping of our areas into seven resting-state networks (Yeo et al., 2011). The results are presented in Supplementary Fig. 14 and show very similar results to our main analysis, thus demonstrating that our results did not depend on our definition of higher-level networks.

In conclusion, our task potency framework quantifies task-induced connectivity changes relative to the resting-state baseline in order to index task-specific modulations away from the brain's functional baseline. Here, we showed that while general task performance relied mainly on within-network interactions, task specificity related to network interactions involved a close exchange between functional networks in both the cortex and subcortical structures. Using the potency framework, we can address how function emerges in response to a task, as well as how the brain's baseline functional architecture influences cognitive operations. As such, the potency of our model lies in its ability to unfold the brain's fluctuations in terms of the resources that are required while performing a task.

Acknowledgments

This work was supported by a Marie Curie International Incoming Fellowship under the European Union’s Seventh Framework Programme (FP7/2007–2013), Grant No. 327340 (Brain Fingerprint) (to MM); The Netherlands Organisation for Scientific Research Grant No. NWO-Vidi 864-12-003 (to CFB); and Wellcome Trust UK Strategic Award [098369/Z/12/Z] (CFB). NWO Brain & Cognition: an Integrative Approach grant 433-09-242 (to JB), NWO National Initiative Brain & Cognition 056-13-015 (to JB), the EU FP7 grants TACTICS (278948) (to JB), IMAGEMEND (602450) (to JB), MATRICS (603016) (to JB) and AGGRESSOTYPE (602805) (to JB), and EU IMI grant EU-AIMS (115300) (to JB). The NeuroMAGE project was supported by NIH Grant R01MH62873 (to Stephen V. Faraone), NWO Large Investment Grant 1750102007010 and ZonMw Grant 60-60600-97193 (to Jan Buitelaar), and grants from Radboud University Nijmegen Medical Center, University Medical Center Groningen and Accare, and VU University Amsterdam.