Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

HHS Public Access

Author manuscript

Nat Neurosci. Author manuscript; available in PMC 2018 July 01.

Published in final edited form as:

Nat Neurosci. 2018 February ; 21(2): 283–289. doi:10.1038/s41593-017-0051-7.

Generalizable Representations of Pain, Cognitive Control, and Negative Emotion in Medial Frontal Cortex

Philip A. Kragel[1,*], **Michiko Kano**[2,3], **Lukas Van Oudenhove**[4], **Huynh Giao Ly**[4], **Patrick Dupont**[5], **Amandine Rubio**[6,7,8], **Chantal Delon-Martin**[6,7], **Bruno L. Bonaz**[6,7,8], **Stephen B. Manuck**[9], **Peter J. Gianaros**[9], **Marta Ceko**[1], **Elizabeth A. Reynolds Losin**[10], **Choong-Wan Woo**[11,12], **Thomas E. Nichols**[13], and **Tor D. Wager**[1,*]

[1]Department of Psychology and Neuroscience, University of Colorado Boulder, Boulder, CO, USA

[2]Frontier Research Institute for Interdisciplinary Sciences, Tohoku University, Sendai, Japan

[3]Department of Behavioral Medicine, Tohoku University Graduate School of Medicine, Sendai, Japan

[4]Department of Clinical and Experimental Medicine, University of Leuven, Leuven, Belgium

[5]Department of Neurosciences, University of Leuven, Leuven, Belgium

[6]Univ. Grenoble Alpes, Grenoble Institut des Neurosciences, GIN, F-38000 Grenoble, France

[7]INSERM, U1216, F-38000 Grenoble, France

[8]CHU Grenoble Alpes, F-38000 Grenoble, France

[9]Department of Psychology, University of Pittsburgh, Pittsburgh, PA, USA

[10]Department of Psychology, University of Miami, Miami, FL, USA

[11]Center for Neuroscience Imaging Research, Institute for Basic Science, Suwon 16419, Republic of Korea

[12]Department of Biomedical Engineering, Sungkyunkwan University, Suwon 16419, Republic of Korea

[13]Oxford Big Data Institute, Li Ka Shing Centre for Health Information and Discovery, Nuffield Department of Population Health, University of Oxford, Oxford

Abstract

Users may view, print, copy, and download text and data-mine the content in such documents, for the purposes of academic research, subject always to the full Conditions of use: http://www.nature.com/authors/editorial_policies/license.html#terms

*Correspondence to: philip.kragel@colorado.edu, tor.wager@colorado.edu.

Author Contributions

P.A.K. and T.D.W. designed the experiment and drafted the manuscript. P.A.K. conducted data analysis. P.A.K., T.N., and T.D.W. developed simulated experiments for evaluating statistical procedures. A.R., B.B., M.C., C.D-M, H.G.L, E.A.R.L., L.V.O., M.K., P.D., P.J.G., S.B.M, T.D.W., and C.-W.W. contributed neuroimaging data. All authors provided feedback and revised the manuscript.

Competing Financial Interests Statement

The authors have no competing interests as defined by Nature Publishing Group, or other interests that might be perceived to influence the results and/or discussion reported in this article.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

The medial frontal cortex (MFC), including anterior midcingulate cortex, has been linked to multiple psychological domains, including cognitive control, pain, and emotion. However, it is unclear whether this region encodes representations of these domains that are generalizable across studies and subdomains. Additionally, if there are generalizable representations, do they reflect a single underlying process shared across domains, or multiple domain-specific processes? We decomposed multivariate patterns of fMRI activity from 270 participants across 18 studies into study-specific, subdomain-specific, and domain-specific components, and identified latent multivariate representations that generalized across subdomains but were specific to each domain. Pain representations were localized to anterior midcingulate cortex, negative emotion representations to ventromedial prefrontal cortex, and cognitive control representations to portions of the dorsal midcingulate. These findings provide evidence for MFC representations that generalize across studies and subdomains, but are specific to distinct psychological domains rather than reducible to a single underlying process.

Introduction

A central aim of cognitive neuroscience is to identify how different mental processes are represented in brain activity. The medial frontal cortex (MFC), which includes multiple functionally distinct cortical areas in the superior frontal and cingulate gyri[1], is one brain region which has been linked to diverse psychological domains, i.e., sets of related psychological states with different adaptive functions[2]. Clearly, different areas within MFC encode different functions, but there is a striking convergence of overlapping functions across domains in several 'hub' areas, particularly the anterior midcingulate cortex (aMCC[3]). Research across species has linked activity in aMCC with multiple functions, including cognitive control[4–5], reward-based learning and decision making[6–9], somatic pain[10, 11], and processing of emotional[12, 13] and social information[14, 15]. In fact, this area responds to such a variety of tasks, and so many underlying functions have been proposed to explain its responses, that it has been described as a "Rorschach test" and understanding it a "holy grail for many cognitive neuroscientists"[16].

Theories of aMCC function often explain the numerous signals in this area as components of an underlying process that operates across domains. Candidate processes have included conflict monitoring[4], adaptive control (i.e. control processes broadly engaged by negative affect and nociception[17]), cognitive effort[18], valuation of actions[19] and control[20], and detecting threats to survival[21], among others. These models have value because they offer integrative explanations for aMCC engagement across multiple domains. However, measuring brain activity across domains with fMRI glosses over a potential multiplicity of different local neural circuits with distinct functions[22, 23]. Electrophysiological and optogenetic studies of likely homologues of human aMCC provide evidence for distinct subpopulations of neurons with different functional properties[6, 8, 24]. Recent evidence suggests that multivariate patterns of fMRI activity can, in some cases, identify representations distributed across subpopulations of cells, including identifying functionally dissociable patterns within aMCC associated with different tasks[25, 26].

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Thus, unified accounts of aMCC function make predictions about the similarity of multivariate brain representations across domains that have not been adequately tested. If a set of domains activate representations of a single underlying process, then engaging these representations by tasks from these domain sets should produce similar patterns of brain activity in aMCC and other MFC areas. Conversely, if different domains engage an underlying pattern that is specific to each domain and not shared by other domains, this would provide evidence against a common underlying process.

Here, we test these predictions using a construct-validation approach grounded in psychometric theory. We investigated three constructs that engage MFC: 'pain', 'cognitive control' and 'negative emotion' (see Methods). We sampled human fMRI data from 18 studies (15 subjects per study, total $n = 270$) in a balanced, hierarchical structure, with three different experimental manipulations in each domain (e.g., evoked cutaneous, visceral nociceptive, and acute mechanical stimulation) and two independent studies for each of these experimental manipulations (i.e., subdomain; Fig 1a). Although it is commonplace in neuroimaging studies to equate a pattern of activity from a single study with a 'representation,' measurement theory and first principles dictate that representations of latent constructs must be *generalizable*. For instance, a representation of 'pain' must generalize across different types of painful stimuli. Our approach allowed us to develop multivariate models that localize brain representations which correspond to a single domain, rather than being driven by the particulars of a subdomain or idiosyncrasies of an individual study (Fig 1b). In this way, these models assess the generalizability of brain representations and test the validity of the theoretical constructs of 'pain,' 'cognitive control,' and 'negative emotion.'

Results

Anatomical delineation of psychological domains

Given evidence for regional specialization of cingulate function on the basis of cytoarchitecture[27], we first applied representational similarity analysis (RSA[28]) within six anatomically defined cortical regions of interest: posterior midcingulate (pMCC), anterior midcingulate (aMCC), perigenual anterior cingulate (pACC), subgenual anterior cingulate (sgACC), ventromedial prefrontal (vmPFC), and dorsal medial frontal cortex (dMFC; Fig 2). By assessing how similar patterns of brain activity are across studies, subdomains, and domains in a single model, RSA can provide evidence for *generalizable* brain representations.

This analysis revealed generalizable representations of painful stimulation in aMCC, pMCC, and dMFC that were not shared by other domains. Parameter estimates for the effect of the pain domain—across the heat, mechanical, and visceral pain subdomains and controlling for study- and subdomain-level effects—were positive within aMCC ($\hat{\beta} = .990 \pm .266$ (*se*), $z = 3.72$, $P = .0002$), pMCC ($\hat{\beta} = .470 \pm .186$ (*se*), $z = 2.55$, $P = .0107$), and dMFC ($\hat{\beta} = .294 \pm .116$ (*se*), $z = 2.59$, $P = .0097$). See Table S1 for more information. These results indicate that patterns of pain-evoked activity in these areas are qualitatively distinct from activity elicited during manipulations of cognitive control or negative emotion, independent of subdomain and study. Accordingly, in terms of aMCC activity patterns, participants in pain studies

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

across different subdomains were more similar to each other (r = .1289 ± .0039 (*se*)) than to those in studies of cognitive control (r = .0083 ± .0026 (*se*); 95% CI of difference = [.1089 .1325]) and negative emotion (r = .0111 ± .0032 (*se*), 95% CI of difference = [.1051 .1297]; Fig 2). Because these correlations are computed across subdomains, they are unlikely to be driven by similarity in any particular subdomain or study. Qualitatively similar results held for patterns of activity in pMCC and dMFC, although they were smaller in magnitude (Tables S1 and S2). These findings are concordant with theoretical models that implicate the aMCC in pain[29, 30] and studies identifying nociceptive circuits in dorsal anterior cingulate cortex[31]. Observations of aMCC activity during noxious stimulation have often been attributed to more general mechanisms, such as directing attention, response selection, or responding to salient events. However, we identified representations of evoked pain distinct from those related to cognitive and emotional domains, which are also attention-demanding, salient, and involve motor preparation, ruling out such general explanations as the primary drivers of aMCC responses during painful stimulation.

The regional analysis also revealed generalizable representations of negative emotion—across social emotion, emotional pictures, and emotional sounds— in vmPFC ($\hat{\beta}$ = .514 ± .140 (*se*), z = 3.65, P = .0003) and dMFC ($\hat{\beta}$ = .404 ± .133 (*se*), z = 3.03, P = .0024; Table S2). Within vmPFC, patterns of activation from different subdomains of negative emotion were more similar to each other (r = .0474 ± .0027 (*se*)) than they were to evoked pain (r = .0117 ± .0023 (*se*), 95% CI of difference = [.0270 .0440]) and cognitive control (r = .0072 ± .0020 (*se*), 95% CI of difference = [.0316 .0484]) studies (Table S4). These observations agree with those of recent neuroimaging studies identifying representations of cross-modal subjective value[32] and perceived emotion[33] in vmPFC. By revealing representations of negative emotion that generalize across stimulus modality and social contexts, these results further substantiate the notion that vmPFC integrates emotional value across diverse stimuli[34, 35]. Further, these data suggest that although painful and unpleasant emotional events can engage a shared negative affective component, vmPFC representations evoked by these two types of stimuli are qualitatively distinct. Recent meta-analytic work has suggested that this difference may be related to the generation of affective meaning[36], where information about environmental cues, memories of past events, and evaluations of potential outcomes are combined into an integrated representation of an organism's well-being in the current environment. This integrative processing would stand in contrast to affective representations that are not conceptually driven, such as pain. We note that these data do not directly assess the generalizability of vmPFC representations to positive emotion, or internally generated states elicited through memory retrieval, as we focused on inductions using negative stimuli.

No cingulate or other areas within MFC exhibited a generalizable representation specific to cognitive control across working memory (N-back), response selection (stop-signal), and response conflict (Flanker, Simon) subdomains (see Methods for citations to included studies). However, we did identify a generalizable representation of response selection, a subdomain of cognitive control particularly involved in motor inhibition, in vmPFC (Table S1). As we observed *deactivation* in this area during task performance (Fig 3; like others[37]), this representation may reflect a pattern of deactivation *not* shared by other domains or other cognitive control subdomains. Patterns of vmPFC activation from different response

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

selection studies were more similar to one another (r = .0828 ± .0033 (*se*)) than to those during manipulations of negative emotion (r = .0052 ± .0023 (*se*), 95% CI of difference = [.0674 .0876]), evoked pain (r = .0286 ± .0024 (*se*), 95% CI of difference = [.0443 .0647]), working memory (r = .0514 ± .0045 (*se*), 95% CI of difference = [.0185 .0442]), or response conflict (r = .0254 ± .0043 (*se*), 95% CI of difference = [.447 .0701]; Table S4). Thus, though generalizable representations of both negative emotion and response selection were observed within vmPFC, these representations appear to be distinct. It is also possible that control-related representations are highly dependent on individual study parameters, as we found strong study-specific effects in multiple regions, including aMCC (Table S2).

Analysis of activation spanning the full extent of MFC (combining the six regions of interest) produced similar results, with effects of painful stimulation and negative emotion that generalize within but not across domains (Fig S1; Table S2). Confirmatory analyses that directly contrasted the spatial similarity of brain activity within domains against spatial correlations across domains further supported these results (Fig S2, Tables S3–S5). Additional confirmatory analyses using different model parameterizations produced qualitatively similar results (Fig S3).

To quantify the weight of evidence favoring generalizable representations specific to each of the three domains, we additionally conducted model comparisons using the Bayesian Information Criterion in each region of interest (see Methods for details). Results of this analysis corroborate inferences drawn on individual parameter estimates (Table 1). aMCC representations were best explained by a model including the domain of pain (in addition to terms for study and subdomain), but not cognitive control or negative emotion. vmPFC representations, on the other hand, were best explained by modeling the domain of negative emotion but not pain or cognitive control. The best fitting models of dMFC and full MFC representations included all three domains, indicative of diverse coding in these regions. Additional model comparisons using the Brainnetome atlas, a parcellation based on functional and anatomical connectivity[38], provide evidence for generalizable representations in other brain regions outside the MFC as well (Table S6, Fig S4).

Searchlight mapping of psychological domains

As there is well-established variability in the anatomy of the cingulate sulcus[39], we additionally conducted searchlight mapping[40] to localize domain-specific representations without strongly relying on the boundaries between regions and lessen the impact of anatomical variability. In this approach, we modeled the similarity structure of spherical volumes (radius = 8 mm) centered at each voxel in MFC, identifying areas wherein local patterns of brain activity contain generalized representations of ‘pain’, ‘cognitive control’, and ‘negative emotion’. By examining patterns of activation in small spherical volumes, these searchlights provide a smooth estimate of pattern information[40] that is not constrained by fixed boundaries that may not match the anatomy of every subject (of importance here, ~40% of the population has a paracingulate gryus[39], which extends the spatial extent of MCC).

The results of the searchlight analysis were largely concordant with those based on anatomical parcellation (Fig 4a, d). Generalizable representations of painful stimulation

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

were found in aMCC within the cingulate sulcus (z_{peak} = 4.88, MNI_{xyz} = [2, 14, 24], P = 1.06e-06, q < .05 FDR-corrected), and extending into dMFC (z_{peak} = 5.58, MNI_{xyz} = [2, 8, 46], P = 2.41e-08, q < .05 FDR-corrected). Also consistent with the regional analysis, representations of negative emotion were found in dMFC, above the dorsal bank of the cingulate sulcus in pre-SMA (z_{peak} = 3.00, MNI_{xyz} = [−10, 10, 52], P = 0.0027), and vmPFC (z_{peak} = 3.78, MNI_{xyz} = [0, 48, −10], P = 1.57e-04) albeit at lower (uncorrected) thresholds.

Unlike the regional analyses, the searchlight analysis revealed domain-specific representations of cognitive control along the cingulate sulcus, extending into SMA and motor cortex (z_{peak} = 2.81, MNI_{xyz} = [−2, 2, 46], P = 0.005). This localization, which falls along the boundary between aMCC and dMFC, agrees with meta-analyses showing the epicenter of control-related activity in this area across working memory, inhibition, and attention shifting tasks[41], and with the posterior rostral cingulate zone[42], a region classically thought to be involved in response selection. Due to its proximity to and connectivity with functionally related brain regions[17], this area is a prime candidate for integrating different types of control signals from multiple sources, such as the expected value of control[20] and value-guided behavioral adaptations[19].

Integrative views of cingulate function are in part supported by observations that overlapping activation is observed in aMCC and adjacent MFC during manipulations of pain, cognitive control, social, and evaluative processing[17]. To assess whether the domain-specific representations identified in the present study similarly overlap within the broader territory of the MFC, we performed a conjunction analysis of the searchlight maps (Fig 4b). Results revealed that these representations were predominately dissociable; the three domains did not commonly overlap in any voxel. Minimal overlap was found in dMFC, with small clusters of activity coding for both pain and cognitive control (60 voxels, 0.57%), and for pain and negative emotion (93 voxels, 0.88%). A small degree of overlap was also observed for pain and negative emotion in vmPFC (17 voxels, 0.53%). The only overlapping effects in cingulate cortex were for pain and cognitive control, spanning the border between pMCC (6 voxels, 1.16%) and aMCC (2 voxels, 1.29%).

To evaluate evidence *against* overlap, we computed Bayes Factors using the minimum z-score from the three-domain conjunction analysis[43, 44]. In this analysis, if the minimum statistic from all three domains is less than or near zero, then there is little support for overlap. Conversely, if the minimum statistic is large and positive, it is more likely that there is overlap across the domains. Values < 1 reflect evidence in favor of the null hypothesis of no representation in all three domains, and values > 1 reflect evidence in favor of overlap. A Bayes Factor < 0.1 is generally considered “strong evidence” against overlap, and a factor > 10 is generally considered “strong evidence” for overlap. This analysis revealed substantial evidence against overlap in the MFC (Fig 4c), with a maximum Bayes Factor = 0.0898.

Discussion

Our results reveal generalizable representations of pain, cognitive control, and negative emotion in separable patterns of MFC activity. The limited overlap of generalizable representations across domains contrasts with conventional, univariate assessments of

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

cingulate function which show substantial overlap across domains[17, 45, 46]. Thus, the novel findings here highlight functional diversity within MFC, and they suggest that domain-specific representations exist in most parts of MFC, including aMCC and adjacent regions. Though domain-specific representations are limited to one domain (pain, negative emotion, and cognitive control), our design allows us to infer that they do generalize across multiple subdomains (e.g., somatic thermal, somatic mechanical, and visceral pain). The generalizable representations we identify provide empirical constraints on what integrative theories of aMCC function must explain. In most of the MFC, it may not be necessary to explain pain- or affect-related and cognitive error-related signals with a single mechanism. Along these lines, it has recently been suggested that the aMCC functions to monitor for conflicts related to ever-present, survival-relevant goals[30]. Consistent with this proposal, we identified generalizable representations of pain in aMCC. However, these representations were qualitatively distinct from those evoked by negative emotional stimuli, including social rejection, that were not generalizable. This distinction makes it unlikely that the aMCC is engaged in the same way to achieve different survival-relevant goals. It is also consistent with other recent work identifying dissociations in MFC activity across tasks taken from different domains[47].

On the other hand, our findings leave room for integrative theories that explain computational mechanisms in terms of a convergence of different neural populations that interact to achieve computational goals. That is, it is possible that the same computational function may be implemented in diverse neural circuits, depending on their inputs and outputs. Our data therefore do not argue against unified computational accounts of aMCC, but rather a unitary neural implementation of those computations.

The proximity of pain, negative emotion, and cognitive representations we identified provides a neural substrate for their comparison and integration[17]. It is possible that integration across domains could be identified in carefully controlled studies implementing within-subject designs across domains and subdomains. Further, although we included many studies and subdomains, our sampling was far from exhaustive, and testing specificity is an open-ended process. Future work combining a more diverse set of subdomains (e.g., incorporating studies of chronic pain, positive and negative reward prediction error, and behavioral withdrawal) with model-based approaches will help further test the claims of integrative theories of control.

In conclusion, we identified domain-specific representations for pain and negative emotion in the aMCC and vmPFC. These representations generalized across participants and diverse subdomains (three per domain). These representations could only be identified by extracting generalizable brain patterns across studies and subdomains. Currently, conventional research investigating population-level neural representation has been conducted at the level of individual studies. Even when very large, studies that sample limited numbers of tasks (i.e., a single exemplar task for a psychological domain) are not capable of identifying generalizable brain representations in this way. Although cross-validation procedures, reliability analysis, and independent replication samples are becoming increasingly common to ensure the generalizability of results, it is evident that findings are often idiosyncratic to a particular study or experimental context, rather than truly reflecting the mental operations

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

under study. Our results demonstrate how modeling the similarity structure of fMRI data drawn from multiple studies and subdomains can overcome this challenge and disambiguate latent brain representations of theoretical constructs.

Methods

Experimental Design

We adopted a construct-validation approach to examine generalizable representations of three constructs that engage the MFC: 'pain', 'cognitive control', and 'negative emotion'. By (a) identifying latent multivariate representations that are indicative of an underlying psychological domain rather than idiosyncrasies of one particular study or subdomain (i.e., task), and (b) examining the similarity of those latent representations across multiple psychological domains, this framework provides a more definitive test of shared representation in the aMCC and areas within the MFC compared to conventional single-study or single-method investigations.

This approach has been taken for decades in psychometric research to assess construct validity[48]. The idea is to define a latent *construct*— 'intelligence' and 'anxiety' are classic examples—and measure it with multiple distinct indicators. Using multiple indicators permits the extraction of common factors that underlie the construct. Inter-correlations among indicators of the same construct provide evidence for convergent validity, suggesting that the indicators measure the same construct. If different indicators uniquely load on different constructs, they provide evidence for discriminant validity. Together, establishing convergence and discrimination provides strong evidence for construct validity[49].

Applied to fMRI, indicators are patterns of brain activity evoked by different tasks, and the constructs are the functional psychological domains (e.g., 'pain' or 'working memory') that the tasks putatively measure. The vast majority of studies, even large-scale studies like the Human Connectome Project and UK Biobank[50, 51], use only one specific task variant as a single indicator for a domain. This poses a problem in inferring the similarity of psychological domains from the similarity in patterns of brain activity. For example, if a 'pain' task differs from a 'negative emotion' task in aMCC[25, 26], is it because 'pain' is represented differently from 'negative emotion', or because the *particular* variant of pain studied differs from the *particular* variant of negative emotion? A study with multiple varieties of pain and multiple varieties of negative emotion could address this question, if it showed that a brain representation common to multiple varieties of pain is distinct from a representation common to multiple varieties of negative emotion. This is the method we used in the present study, as detailed below.

Study and Contrast Selection

fMRI data were sampled from studies of acute thermal somatic stimulation[52, 53], acute visceral stimulation[54, 55], acute mechanical somatic stimulation, working memory[56, 57], response selection[58, 59], response conflict[60], induction of negative emotion using images of visual scenes[61, 62], social rejection[63], the perception of others in pain[25], and emotionally aversive vignettes from the International Affective Digital Sounds system[64]. Together these

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

data formed a balanced hierarchical sample, with six pain studies (two thermal, two visceral, and two mechanical), six cognitive control studies (two working memory, two response selection, and two response conflict), and six negative emotion studies (two visual, two social, and two auditory). Although negative emotion can be evoked through diverse methods, including the recollection of emotional events, the brain activity we analyze here is focused on exteroceptive processing – which has reliably been linked to overlapping MFC activity[17].

Due to variability in sample size across studies (range = 15–183), data were randomly subsampled by selecting 15 participants from each study (total $n = 270$). Although no statistical methods were used to pre-determine this sample size, it is similar to those reported in previous publications[7,9,11] (see also Table S7 for maximal univariate effects in the present sample). Because our focus was to generalize across studies, no attempts to replicate individual experiments were made. No participants were excluded from the analysis. A subset of these data was previously utilized to validate the use of automated meta-analysis to decode cognitive states of pain, emotion, and working memory[62].

This was not a randomized study; it was a mega-analysis of multiple studies. Participants were recruited independently for each of the 18 studies being analyzed. Data collection was conducted blind to the goals of the present study. A posteriori group assignment was based on the goals of each study and experimental manipulation being used (e.g., studies involving thermal stimulation of the forearm were considered members of the 'pain' domain). Data analysis was not performed blind to the conditions of the experiments.

Informed consent was provided by all subjects in accordance with local ethics and institutional review boards. Participants sampled from studies 5 ($n = 15$, 4 female, $M_{age} = 26.9$), 6 ($n = 15$, 8 female, $M_{age} = 24.2$), 17 ($n = 15$, 7 female, $M_{age} = 31.1$) and 18 ($n = 15$, 9 female, $M_{age} = 24.4$) provided informed consent as approved by the University of Colorado Boulder institutional review board. Participants in study 11 ($n = 21$, 9 female, $M_{age} = 30.5$) provided informed consent in accordance with the New York University institutional review board. Descriptions of ethics approval, image acquisition and analysis, and demographics are described briefly in Table S7 and in full detail in the corresponding references (see also the Life Sciences Reporting Summary).

Contrasts from thermal painful stimulation were between high and low levels of pain[53] or between high levels of painful stimulation and baseline[52]. Contrasts for visceral stimulation studies were between rectal distension trials and baseline. For mechanical stimulation studies, contrasts were made between pressure application to the thumb and baseline. Contrasts for both working memory studies were between N-back blocks and a fixation baseline. For response selection studies, contrasts were between trials in a go/no-go task engaging response selection (as defined in the Cognitive Atlas[65]) against baseline. Response conflict contrasts were made between congruent and incongruent trials in studies using the Eriksen Flanker and Simon paradigms. Studies of visual negative emotion compared negative to neutral IAPS pictures[62] or negative pictures against baseline[61]. Social negative emotion studies compared viewing pictures of ex-partners versus friends[63] and viewing

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

images of others in pain[25] versus baseline. Auditory negative emotion studies compared listening to unpleasant affective sounds to baseline.

fMRI Analysis

We employed three converging methods to isolate generalizable brain representations: 1) we modeled how dissimilar patterns of brain activity were from another, called representational similarity analysis (RSA[28, 66]); 2) we directly compared the similarity of brain activity from studies coming from the same domain (but different studies and subdomains) to studies from different domains; and 3) we used partial-least squares regression[67] to characterize the spatial profile of generalizable brain representations in MFC.

Feature Selection

Input data were defined *a priori* as voxels located within the anterior midline, defined as voxels within cingulate cortex and superior frontal gyrus (in the LONI Probabilistic Brain Atlas[68]) anterior to the plane y = −22 mm in MNI space. The primary analyses were conducted within four cingulate subregions[27] as well as ventral and dorsal aspects of medial frontal cortex. These anatomically defined regions of interest include posterior midcingulate cortex (pMCC; y = −22 mm to y = 4.5 mm, z > 5 mm, 822 voxels), anterior midcingulate cortex (aMCC; y = 4.5 mm to y = 30 mm, z > 5 mm, 815 voxels), perigenual anterior cingulate cortex (pACC; y > 30 mm, z < 5 mm, 794 voxels), and subgenual anterior cingulate cortex (sgACC; y = 4.5 mm to y = 30 mm, z < 5 mm, 302 voxels), as well as within superior frontal gyrus and ventromedial prefrontal cortex (split by the plane z = 5 mm, dMFC and vmPFC, 4311 and 10619 voxels respectively). As these divisions were originally defined in Talairach space, they were converted to MNI152 space using the Lancaster transform[69]. A secondary analysis was performed using searchlight mapping[40], where multiple analyses were conducted, each using patterns of fMRI activation within spherical regions (radius = 8 mm) centered at every voxel in the MFC as input. An additional exploratory analysis using RSA-based model comparison, described in full detail below, was conducted using a whole-brain parcellation based on structural and functional connectivity[38].

RSA: model specification

We estimated representational dissimilarity matrices (RDMs) by computing the correlation distance (1 – Pearson's *r,* excluding on-diagonal elements which have a dissimilarity value of 0) of multi-voxel patterns of brain activity. Each pattern was acquired from one of 270 subjects, which came from the full sample of 18 studies (n = 15 per study). Next, we constructed model-based RDMs to characterize different components of a psychological hierarchy (Fig 1). At the lowest level, the 18 studies were individually modeled to account for study-specific idiosyncrasies. Next, the nine subdomains (visceral stimulation, thermal stimulation, mechanical stimulation, response conflict, stop/go response selection, working memory, visual negative emotion, social negative emotion, auditory negative emotion) were modeled to account for response patterns that generalize across studies. Finally, the three psychological domains (pain, cognitive control, and negative emotion) were modeled as independent predictors, to account for response patterns that generalize not only across studies, but across subdomains as well, hence being generalizable within but not across the

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

three domains. Individual RDMs were computed from binary vectors indicating membership based on study (18 RDMs), subdomain (9 RDMs), or psychological domain (3 RDMs). The unique off-diagonal elements (inter-subject dissimilarities) of these 30 RDMs, in addition to a constant RDM, were vectorized to form regressors in a model. Linear regression was used with this model to fit the observed inter-subject brain dissimilarity matrix. On-diagonal elements were excluded for all similarity-based analyses, as they have perfect correlation and zero dissimilarity. The general linear model assumes independence while dissimilarity matrices exhibit complex dependence; as a result we use bootstrap inference to obtain *P*-values (see inferences on model parameters section below).

RSA: model properties and diagnostics

To assess the suitability of our models, we conducted simulated experiments using resting-state fMRI (rsfMRI) data ($n = 270$) from the 1,000 functional connectomes project[70]. This allowed us to test the RSA models' false positive rates, and the bias and variance of parameter estimates, using data with no true effects but real fMRI noise. This is the approach recently taken by Eklund et al. (Ref. [71]) to assess false positive rates in standard GLM analyses.

We used a Monte Carlo procedure to conduct 1,000 RSA-based analyses of rsfMRI data, each with randomly generated event-related fMRI models for the 270 participants, allowing us to estimate the distribution of RSA parameter estimates under the null hypothesis. To mirror the dependence structure (including study/site effects) in our present sample, we selected rsfMRI participants from 18 different sites, sampling 15 participants from each site (total $n = 270$). All data were subjected to standard preprocessing, including realignment to correct for motion, non-linear warping to standard MNI space, spatial smoothing (4mm FWHM), and high-pass filtering (128s cutoff). Then, for each Monte Carlo iteration, for each participant, we generated a series of 10 events with random onsets, modeled them with diraç delta (impulse response) functions and convolved them with SPM's standard hemodynamic response function to generate a single regressor of interest for each subject (a constant term was also included). Then, we estimated these models, generating a series of 270 event-related activation maps. These were subjected to RSA-based modeling, as described in the model specification section above. To estimate the false positive rates for RSA model parameter estimates in bootstrap-based inference, we conducted bootstrap resampling (b = 200 bootstrap samples) and obtained *P*-values for each RSA model parameter. We repeated this entire procedure 1,000 times, with randomly specified models on every iteration, i.e., estimating a total of 270,000 unique fMRI activation maps, and 1,000 RSA model fits and bootstrap tests.

Under the assumption that the rsfMRI data contain no consistent relationship with the randomly specified models, an unbiased model should have parameter estimates centered on zero across the 1,000 Monte Carlo iterations, indicating no systematic bias. Thus, to estimate bias, we calculated the mean deviation of each parameter estimate from zero. We also estimated the variance of parameter estimates across the 1,000 iterations; lower variance indicates greater precision and power. To estimate variance, we calculated the standard deviation in each RSA parameter estimate across the 1,000 iterations. The proportion of

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

false-positives was assessed using a $P < .05$ cutoff (two-tailed). False positives were defined as parameter estimates below the 2.5th or above the 97.5th percentile on the bootstrap distribution, and the false positive rate for each RSA model parameter estimate was defined as the proportion of the 1,000 RSA models for which that parameter estimate was significant.

The results of these analyses indicated that the modeling procedure is unbiased, i.e., average null-hypothesis values were nearly exactly zero. Of 180 parameters evaluated (6 ROIs x 30 parameters) the largest effect was not significantly different from zero ($z = 0.904$, $P = 0.366$). In addition, false-positive rates for all RSA model parameters were at or below the nominal value of 0.05 (Fig S5).

We also repeated the entire RSA model simulation (500 iterations) using synthetic null-hypothesis data generated from a Wishart distribution (Fig S6) and with a homogeneous set of task data (using 180 subjects from study 13) and found qualitatively identical results (not shown). This simulated a case in which there are no study-level effects. Overall, the RSA model procedures provide unbiased estimates under the null hypothesis, and false positive rates are appropriately controlled.

RSA: Model identifiability

Here we model inter-subject RDMs, so the data to be modeled comprise an n x n dissimilarity matrix, where n is the number of participants; this matrix has rank min(n,r) where r is the number of voxels used to compute the correlation. The lower triangle of the dissimilarity matrix is vectorized and fit with a linear model with design matrices based on equivalently vectorized dissimilarity (as illustrated in Fig 1). Thus, the outcome data is a vector of length $u = n \times (n-1)/2$. The model dimension depends on the exact parameterization used, but a saturated model for the effect of study would have dimension $k + k \times (k-1)/2$, where k is the number of studies; this is k parameters for the average intra-study relationship (for pairs of subjects) within each study, and $k \times (k-1)/2$ parameters for each possible relationship between each pair of studies. In practice, we use a much simpler model, but confirm identifiability by checking the rank of the design matrix and variance inflation factors.

In our study, we included $k = 18$ studies with $n = 270$ participants divided equally among them (15 per study). We utilized the constant term in the model to characterize the average similarity of data within studies, leaving a subspace of inter-study relationships spanned by $k + k \times (k-1)/2 - 1 = 170$ dimensions. In principle, the upper bound on the number of regressors for inter-study differences in an identifiable model is 170. However, we were primarily interested in specific inter-study relationships, particularly those common to subdomains (9 parameters for pairs of studies with the same subdomain) and those common to domains (3 parameters for three sets of 9 studies that load on the same domain construct across 3 subtypes). Our full model thus contained 31 regressors: 18 for specific studies, 9 for subdomains, 3 for domains, and one intercept term. As the number of voxels per ROI was at least 302 > 31, there was no risk of degenerate (zero-residual) models.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

RSA: Model rank and variance inflation factors

To confirm the identifiability and efficiency of our models, we computed the rank of our design matrix and variance inflation factors (VIFs) for each regressor. The variance inflation factors (VIFs) show the degree to which variance in each parameter estimate is increased due to partial collinearity with linear combinations of other regressors. The rank of our design matrix is 31, indicating that the model is full-rank and identifiable, and not overparameterized. VIFs were finite for all regressors, consistent with the fact that model was identifiable (Figure S2). VIFs varied across regressors based on the partially shared variance across study, domain and subdomain, which was unavoidable because part of the covariance common to participants in each domain is shared with the subdomains and studies that fall within it. However, the VIFs were clearly in a range that indicated reasonable ability to make inferences on the unique variance explained by each parameter, given the sample size. The regressors which were of primary interest (i.e., the three domain level terms) had VIFs = 1.66.

Statistics: inferences on RSA model parameters

Parameter estimates ($\hat{\beta}$) from the model provide estimates of generalizability, with the interpretation of a significant (non-zero) parameter estimate depending on the nature of the regressor. Study-specific regressors test generalizability across individual participants within a study, in the sense that they capture inter-subject correlations in the spatial patterns of activity. Positive values indicate similar patterns across participants for the study modeled. Subdomain-level regressors test generalizability across two studies of the same subdomain, controlling for other model parameters. Positive values indicate shared spatial patterns across studies of the modeled subdomain, and thus evidence for a coherent subdomain. Domain-level regressors, which were of primary interest here, test generalizability across three distinct subdomains (6 studies), controlling for shared patterns specific to the study and subdomain. Positive values indicate shared spatial patterns across subdomains, and thus evidence for a coherent domain-related pattern. Thus, we refer to these parameter estimates as generalization indices (e.g., Fig 2), as they reflect the extent to which patterns of brain activation generalize across subdomains, studies, or subjects. Because the regressors were tested jointly in multiple regression modeling, significant domain-level parameter estimates imply that the pattern of brain activity shared across the domain was not reducible to a subdomain-specific or study-specific pattern.

Inference on parameter estimates ($\hat{\beta}$) was made using bootstrap resampling of subjects. This procedure involved repeatedly resampling subjects with replacement for each study over 5,000 iterations for the regional analysis and over 1,000 iterations for the searchlight analysis (see below). In each resampling, a new RDM was constructed using fMRI activation for the resampled subjects and GLMs were estimated. Because samples from the same subject can be drawn multiple times in this approach, pairwise dissimilarity values from the same subject (with a dissimilarity value of 0) were excluded from the analysis. Bootstrap distributions for individual model parameters were compared against 0 using normal approximation for inference. These distributions were visually inspected and assumed to be normal, although this was not statistically tested. Unless noted otherwise, the main results reported in the manuscript (both regional and searchlight analyses) were

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

thresholded after correcting for multiple comparisons based on the false discovery rate (FDR corrected, $q < .05$). All tests are two-tailed unless otherwise specified.

Statistics: RSA model comparison

To formally compare the amount of evidence for domain-generalizable representations, model comparisons were made using the Bayesian Information Criterion (BIC). The reference model included terms for each study (18 parameters), each subdomain (9 parameters), and a constant (28 parameters in total). Next, we fit three models that each included a single additional term for one of the psychological domains (29 parameters in total). Finally, a more fully specified model which contained all three psychological domains (31 parameters in total) was fit. These five models were fit for each region of interest and the full extent of MFC. BIC values were computed using the log-likelihood of fitted models and penalizing based on the number of free parameters[72]. The number of samples was set the number of participants included in the analyses (270) as opposed to the number of unique elements in the dissimilarity matrix, because of dependence between elements of the dissimilarity matrix. BIC values were converted to weights using the formulation in Wagenmakers and Farrell[73]. These weights characterize which model is most likely to have produced the observed similarity structure in each region, given the *a priori* set of models. Finally, the adjusted R^2 was computed for the model favored by the BIC analysis in each region of interest.

This analysis was additionally conducted for regions spanning the whole brain (as delineated in the Brainnetome Atlas[38]). Because this parcellation contains regions with fewer voxels than participants included in the analysis, the number of samples was set to the minimum of 270 and the number of voxels in each parcel.

Statistics: “Model-free” analysis comparing spatial correlations within and between domains

To provide evidence for generalizable brain representations without using an explicit model of expected similarity relations, we tested for non-zero correlations within domains and between different subdomains, within domains and between studies, and within domains generally. Tests of non-zero correlations were performed by constructing confidence intervals ($\alpha = .05$) using bootstrap analyses with the bias corrected and accelerated percentile method. We also conducted a series of hierarchical tests comparing: 1) inter-subject correlations from the same domain but different subdomains vs. correlations from different domains; 2) correlations from the same domain but different studies vs. correlations from the same domain but different subdomains; 3) correlations from the same domain vs. correlations from the same domain versus different studies. These tests are constructed to sequentially identify the average effect of domain, subdomain, and study for each of the three domains. Inferences were drawn by calculating bootstrap confidence intervals based on the mean difference in correlation coefficients (for a graphical depiction, see Fig S7).

Statistics: PLS estimation of latent patterns for each domain

To estimate patterns of activity associated with each domain (Fig 2), Partial Least Squares regressions[69] were run separately in each region of interest, with contrasts from all 270

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

subjects forming the data matrix, and dummy coded variables forming the output matrix (270 subjects by 30 parameters: 18 studies, 9 subdomains, and 3 domains, with values of 1/−1 based on inclusion/exclusion for each term). Parameter estimates were bootstrapped over 5,000 iterations, and z-scores were estimated based on the mean and standard error of the bootstrap distributions. We note that this approach is not designed to ensure that representations are uniquely specific to each domain; if some studies or subdomains have especially high covariance with a domain, they could have a large influence on domain-level patterns. For this reason, PLS-based estimation of patterns is used in conjunction with RSA and direct comparisons of inter-subject correlations.

Statistics: searchlight thresholds and assessment of overlap

For the conjunction analysis and visualization of searchlight maps, an uncorrected threshold ($P < .05$) was used to ensure that an overly conservative threshold did not obscure overlapping regions. To estimate the relative evidence for and against overlap of the domains, Bayes Factors were computed[44] using the minimum statistic compared to the conjunction null[74] and a uniform distribution ranging from 0 to 10 as a prior distribution (theoretically plausible values of parameters estimates). This test evaluates whether there is more evidence in favor of overlap (i.e., that the minimum statistic of the three maps is substantially greater than zero) or against overlap (the minimum statistic of the three maps is relatively close to or less than zero). Bayes Factors greater than 3 provide evidence of overlapping representations, whereas values less than .33 provide evidence against overlap.

Correspondence of searchlight maps with existing parcellations

River plots were created to depict the correspondence (quantified as the cosine similarity) between the searchlight maps and existing anatomical[27] as well as functional[75] parcellations.

Data Availability

The fMRI data for studies 9–12 are available from OpenfMRI, https://openfmri.org/dataset/ds000008/, https://openfmri.org/dataset/ds000007/, https://openfmri.org/dataset/ds000101/, and https://openfmri.org/dataset/ds000102/. fMRI data for study 13 is available at NeuroVault, http://neurovault.org/collections/503. The fMRI data that support the findings of this study are available from https://canlabweb.colorado.edu/files/MFC_Generalizability.tar.gz

Code Availability

MATLAB code for implementing all analyses is available at https://github.com/canlab/ and in the Supplementary Software.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Shin Fukudo, Tomohiko Muratsubaki, Joe Morishita for assistance with data collection; K. Ochsner for sharing data from studies of negative emotion; Todd Braver and Jeremy Gray for sharing working memory data;

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

and R. Poldrack for sharing response selection data (available at OpenfMRI.org). This research was supported by grants R01 HL089850 to P.J.G.; P01 HL040962 to S.B.M., grants OCI-1131801; R01 DA035484; R01 MH076136 to T.D.W., JSPS-FWO grant VS.014.13N to L.V.O. and S. Fukudo, JSPS-KAKENHI grant 26460898 to MK, R01 MH076137; R01 AG043463 to K.O., from the Direction de la Recherche Clinique of the University Hospital of Grenoble Alpes, and the pharmaceutical labs Ferring and Cephalon. L.V.O. is an assistant research professor funded by the KU Leuven Special Research Fund. T.E.N. is supported by the Wellcome Trust.

References

1. Amodio DM, Frith CD. Meeting of minds: the medial frontal cortex and social cognition. Nat Rev Neurosci. 2006; 7:268–277. [PubMed: 16552413]
2. Cosmides L, Tooby J. Origins of domain specificity: The evolution of functional organization. Mapping the mind: Domain specificity in cognition and culture. 1994:85–116.
3. Vogt BA. Midcingulate cortex: Structure, connections, homologies, functions and diseases. Journal of chemical neuroanatomy. 2016; 74:28–46. [PubMed: 26993424]
4. Botvinick M, Nystrom LE, Fissell K, Carter CS, Cohen JD. Conflict monitoring versus selection-for-action in anterior cingulate cortex. Nature. 1999; 402:179–181. [PubMed: 10647008]
5. Ridderinkhof KR, Ullsperger M, Crone EA, Nieuwenhuis S. The role of the medial frontal cortex in cognitive control. Science. 2004; 306:443–447. [PubMed: 15486290]
6. Ito S, Stuphorn V, Brown JW, Schall JD. Performance monitoring by the anterior cingulate cortex during saccade countermanding. Science. 2003; 302:120–122. [PubMed: 14526085]
7. Dosenbach NU, et al. A core system for the implementation of task sets. Neuron. 2006; 50:799–812. [PubMed: 16731517]
8. Procyk E, Tanaka YL, Joseph JP. Anterior cingulate activity during routine and non-routine sequential behaviors in macaques. Nat Neurosci. 2000; 3:502–508. [PubMed: 10769392]
9. Kolling N, Behrens TE, Mars RB, Rushworth MF. Neural mechanisms of foraging. Science. 2012; 336:95–98. [PubMed: 22491854]
10. Buchel C, et al. Dissociable neural responses related to pain intensity, stimulus intensity, and stimulus awareness within the anterior cingulate cortex: A parametric single-trial laser functional magnetic resonance imaging study. Journal of Neuroscience. 2002; 22:970–976. [PubMed: 11826125]
11. Rainville P, Duncan GH, Price DD, Carrier B, Bushnell MC. Pain affect encoded in human anterior cingulate but not somatosensory cortex. Science. 1997; 277:968–971. [PubMed: 9252330]
12. Etkin A, Egner T, Peraza DM, Kandel ER, Hirsch J. Resolving emotional conflict: A role for the rostral anterior cingulate cortex in modulating activity in the amygdala (vol 51, pg 871, 2006). Neuron. 2006; 52:1121–1121.
13. Bishop S, Duncan J, Lawrence AD. Prefrontal cortical function and anxiety: controlling attention to threat-related stimuli. Nature Neuroscience. 2004; 7:184–188. [PubMed: 14703573]
14. Tomlin D, et al. Agent-specific responses in the cingulate cortex during economic exchanges. Science. 2006; 312:1047–1050. [PubMed: 16709783]
15. Rudebeck PH, Buckley MJ, Walton ME, Rushworth MFS. A role for the macaque anterior cingulate gyrus in social valuation. Science. 2006; 313:1310–1312. [PubMed: 16946075]
16. Ebitz RB, Hayden BY. Dorsal anterior cingulate: a Rorschach test for cognitive neuroscience. Nat Neurosci. 2016; 19:1278–1279. [PubMed: 27669987]
17. Shackman AJ, et al. The integration of negative affect, pain and cognitive control in the cingulate cortex. Nat Rev Neurosci. 2011; 12:154–167. [PubMed: 21331082]
18. Critchley HD, et al. Human cingulate cortex and autonomic control: converging neuroimaging and clinical evidence. Brain. 2003; 126:2139–2152. [PubMed: 12821513]
19. Behrens TE, Woolrich MW, Walton ME, Rushworth MF. Learning the value of information in an uncertain world. Nat Neurosci. 2007; 10:1214–1221. [PubMed: 17676057]
20. Shenhav A, Botvinick MM, Cohen JD. The expected value of control: an integrative theory of anterior cingulate cortex function. Neuron. 2013; 79:217–240. [PubMed: 23889930]
21. Eisenberger NI, Lieberman MD. Why rejection hurts: a common neural alarm system for physical and social pain. Trends Cogn Sci. 2004; 8:294–300. [PubMed: 15242688]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

22. Haynes JD. A Primer on Pattern-Based Approaches to fMRI: Principles, Pitfalls, and Perspectives. Neuron. 2015; 87:257–270. [PubMed: 26182413]
23. Logothetis NK. What we can do and what we cannot do with fMRI. Nature. 2008; 453:869–878. [PubMed: 18548064]
24. Kvitsiani D, et al. Distinct behavioural and network correlates of two interneuron types in prefrontal cortex. Nature. 2013; 498:363–366. [PubMed: 23708967]
25. Krishnan A, et al. Somatic and vicarious pain are represented by dissociable multivariate brain patterns. eLife. 2016; 5
26. Woo CW, et al. Separate neural representations for physical pain and social rejection. Nature communications. 2014; 5:5380.
27. Vogt BA, Berger GR, Derbyshire SW. Structural and functional dichotomy of human midcingulate cortex. The European journal of neuroscience. 2003; 18:3134–3144. [PubMed: 14656310]
28. Kriegeskorte N, Mur M, Bandettini P. Representational similarity analysis - connecting the branches of systems neuroscience. Front Syst Neurosci. 2008; 2:4. [PubMed: 19104670]
29. Peyron R, Laurent B, Garcia-Larrea L. Functional imaging of brain responses to pain. A review and meta-analysis (2000). Neurophysiologie clinique = Clinical neurophysiology. 2000; 30:263–288. [PubMed: 11126640]
30. Lieberman MD, Eisenberger NI. The dorsal anterior cingulate cortex is selective for pain: Results from large-scale reverse inference. Proc Natl Acad Sci U S A. 2015; 112:15250–15255. [PubMed: 26582792]
31. Hutchison WD, Davis KD, Lozano AM, Tasker RR, Dostrovsky JO. Pain-related neurons in the human cingulate cortex. Nat Neurosci. 1999; 2:403–405. [PubMed: 10321241]
32. McNamee D, Rangel A, O'Doherty JP. Category-dependent and category-independent goal-value codes in human ventromedial prefrontal cortex. Nat Neurosci. 2013; 16:479–485. [PubMed: 23416449]
33. Peelen MV, Atkinson AP, Vuilleumier P. Supramodal representations of perceived emotions in the human brain. J Neurosci. 2010; 30:10127–10134. [PubMed: 20668196]
34. Levy DJ, Glimcher PW. The root of all value: a neural common currency for choice. Curr Opin Neurobiol. 2012; 22:1027–1038. [PubMed: 22766486]
35. Montague PR, Berns GS. Neural economics and the biological substrates of valuation. Neuron. 2002; 36:265–284. [PubMed: 12383781]
36. Roy M, Shohamy D, Wager TD. Ventromedial prefrontal-subcortical systems and the generation of affective meaning. Trends Cogn Sci. 2012; 16:147–156. [PubMed: 22310704]
37. Laird AR, et al. Investigating the functional heterogeneity of the default mode network using coordinate-based meta-analytic modeling. J Neurosci. 2009; 29:14496–14505. [PubMed: 19923283]
38. Fan L, et al. The Human Brainnetome Atlas: A New Brain Atlas Based on Connectional Architecture. Cereb Cortex. 2016; 26:3508–3526. [PubMed: 27230218]
39. Paus T, et al. Human cingulate and paracingulate sulci: pattern, variability, asymmetry, and probabilistic map. Cereb Cortex. 1996; 6:207–214. [PubMed: 8670651]
40. Kriegeskorte N, Goebel R, Bandettini P. Information-based functional brain mapping. Proc Natl Acad Sci U S A. 2006; 103:3863–3868. [PubMed: 16537458]
41. Van Snellenberg JX, Wager TD. Cognitive and motivational functions of the human prefrontal cortex. Luria's legacy in the 21st century. 2009:30–61.
42. Amiez C, Petrides M. Neuroimaging evidence of the anatomo-functional organization of the human cingulate motor areas. Cereb Cortex. 2014; 24:563–578. [PubMed: 23131805]
43. Gallistel CR. The importance of proving the null. Psychol Rev. 2009; 116:439–453. [PubMed: 19348549]
44. Dienes Z. Using Bayes to get the most out of non-significant results. Frontiers in psychology. 2014; 5
45. de la Vega A, Chang LJ, Banich MT, Wager TD, Yarkoni T. Large-Scale Meta-Analysis of Human Medial Frontal Cortex Reveals Tripartite Functional Organization. J Neurosci. 2016; 36:6553–6562. [PubMed: 27307242]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

46. Torta DM, Cauda F. Different functions in the cingulate cortex, a meta-analytic connectivity modeling study. Neuroimage. 2011; 56:2157–2172. [PubMed: 21459151]
47. Jahn A, Nee DE, Alexander WH, Brown JW. Distinct regions within medial prefrontal cortex process pain and cognition. Journal of Neuroscience. 2016; 36:12385–12392. [PubMed: 27807031]
48. Cronbach LJ, Meehl PE. Construct validity in psychological tests. Psychol Bull. 1955; 52:281–302. [PubMed: 13245896]
49. Campbell DT, Fiske DW. Convergent and discriminant validation by the multitrait-multimethod matrix. Psychol Bull. 1959; 56:81–105. [PubMed: 13634291]
50. Barch DM, et al. Function in the human connectome: task-fMRI and individual differences in behavior. Neuroimage. 2013; 80:169–189. [PubMed: 23684877]
51. Miller KL, et al. Multimodal population brain imaging in the UK Biobank prospective epidemiological study. Nat Neurosci. 2016
52. Wager TD, et al. An fMRI-based neurologic signature of physical pain. N Engl J Med. 2013; 368:1388–1397. [PubMed: 23574118]
53. Atlas LY, Bolger N, Lindquist MA, Wager TD. Brain mediators of predictive cue effects on perceived pain. J Neurosci. 2010; 30:12964–12977. [PubMed: 20881115]
54. Rubio A, et al. Uncertainty in anticipation of uncomfortable rectal distension is modulated by the autonomic nervous system - A fMRI study in healthy volunteers. Neuroimage. 2015; 107:10–22. [PubMed: 25479021]
55. Kano M, et al. Influence of uncertain anticipation on brain responses to aversive rectal distension in patients with irritable bowel syndrome. Psychosomatic medicine. 78:A80–81. (In press).
56. DeYoung CG, Shamosh NA, Green AE, Braver TS, Gray JR. Intellect as distinct from Openness: differences revealed by fMRI of working memory. J Pers Soc Psychol. 2009; 97:883–892. [PubMed: 19857008]
57. van Ast VA, et al. Brain Mechanisms of Social Threat Effects on Working Memory. Cereb Cortex. 2016; 26:544–556. [PubMed: 25249408]
58. Xue G, Aron AR, Poldrack RA. Common neural substrates for inhibition of spoken and manual responses. Cereb Cortex. 2008; 18:1923–1932. [PubMed: 18245044]
59. Aron AR, Behrens TE, Smith S, Frank MJ, Poldrack RA. Triangulating a cognitive control network using diffusion-weighted magnetic resonance imaging (MRI) and functional MRI. J Neurosci. 2007; 27:3743–3752. [PubMed: 17409238]
60. Kelly AMC, Uddin LQ, Biswal BB, Castellanos FX, Milham MP. Competition between functional brain networks mediates behavioral variability. Neuroimage. 2008; 39:527–537. [PubMed: 17919929]
61. Gianaros PJ, et al. An inflammatory pathway links atherosclerotic cardiovascular disease risk to neural activity evoked by the cognitive regulation of emotion. Biol Psychiatry. 2014; 75:738–745. [PubMed: 24267410]
62. Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD. Large-scale automated synthesis of human functional neuroimaging data. Nat Methods. 2011; 8:665–670. [PubMed: 21706013]
63. Kross E, Berman MG, Mischel W, Smith EE, Wager TD. Social rejection shares somatosensory representations with physical pain. Proc Natl Acad Sci U S A. 2011; 108:6270–6275. [PubMed: 21444827]
64. Bradley, MM., Lang, PJ. Tech Rep B-3. University of Florida; Gainesville, FL: 2007. The International Affective Digitized Sounds (; IADS-2): Affective ratings of sounds and instruction manual.
65. Poldrack RA, et al. The cognitive atlas: toward a knowledge foundation for cognitive neuroscience. Frontiers in neuroinformatics. 2011; 5:17. [PubMed: 21922006]
66. Kriegeskorte N, Kievit RA. Representational geometry: integrating cognition, computation, and the brain. Trends Cogn Sci. 2013; 17:401–412. [PubMed: 23876494]
67. Wold S, Sjostrom M, Eriksson L. PLS-regression: a basic tool of chemometrics. Chemometr Intell Lab. 2001; 58:109–130.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

68. Shattuck DW, et al. Construction of a 3D probabilistic atlas of human cortical structures. Neuroimage. 2008; 39:1064–1080. [PubMed: 18037310]

69. Lancaster JL, et al. Bias between MNI and Talairach coordinates analyzed using the ICBM-152 brain template. Hum Brain Mapp. 2007; 28:1194–1205. [PubMed: 17266101]

70. Biswal BB, et al. Toward discovery science of human brain function. P Natl Acad Sci USA. 2010; 107:4734–4739.

71. Eklund A, Nichols TE, Knutsson H. Cluster failure: Why fMRI inferences for spatial extent have inflated false-positive rates. P Natl Acad Sci USA. 2016; 113:7900–7905.

72. Schwarz G. Estimating Dimension of a Model. Ann Stat. 1978; 6:461–464.

73. Wagenmakers EJ, Farrell S. AIC model selection using Akaike weights. Psychon Bull Rev. 2004; 11:192–196. [PubMed: 15117008]

74. Nichols T, Brett M, Andersson J, Wager T, Poline JB. Valid conjunction inference with the minimum statistic. Neuroimage. 2005; 25:653–660. [PubMed: 15808966]

75. Yeo BT, et al. The organization of the human cerebral cortex estimated by intrinsic functional connectivity. Journal of neurophysiology. 2011; 106:1125–1165. [PubMed: 21653723]

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

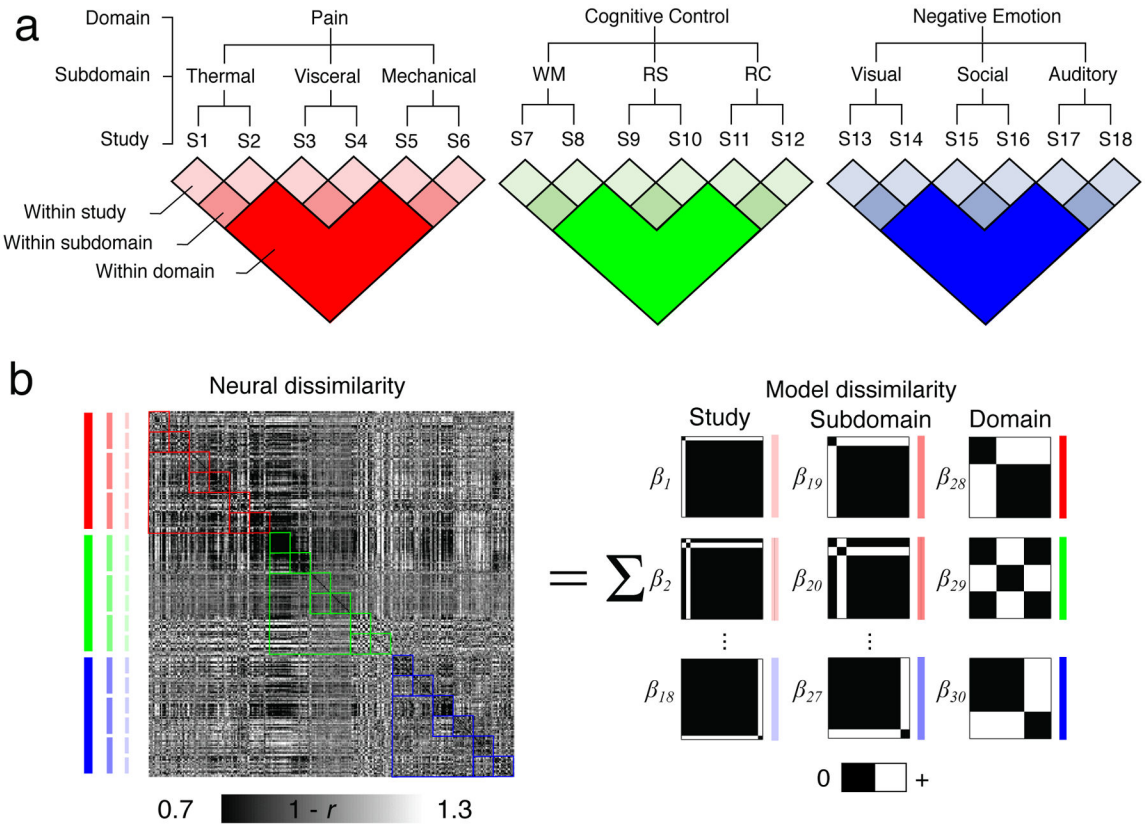

Figure 1.

Study selection and multivariate modeling. (**a**) Hierarchical structure of studies and tasks. Dendrograms convey theoretical groupings of fMRI activity at levels of study (level 1: studies S1–S18), subdomain (level 2: Thermal, Visceral, Mechanical, Working Memory (WM), Response Selection (RS), Response Conflict (RC), Visual, Social, and Auditory), and domain (level 3: Pain, Cognitive Control, and Negative Emotion). Colored regions illustrate model-based partitioning of neural similarity into components that generalize across subjects (unique to a study, top 18 squares), studies (unique to a subdomain, middle nine squares), and subdomains (unique to a domain, bottom three regions). (**b**) Decomposing multivariate pattern similarity into study-, subdomain-, and domain-specific components. The matrix in the left panel shows the dissimilarity of fMRI patterns across all subjects ($n = 270$) in the entire medial frontal cortex. Each row represents one individual participant, and each element the dissimilarity (1 – Pearson's correlation coefficient) in brain activity patterns for two individuals. Colored bars to the left indicate corresponding levels in the functional hierarchy. The right panel shows how the observed neural dissimilarity across pairs of images from the 18 studies is modeled as a weighted summation of theoretical dissimilarity matrices constructed according to study (18 parameters), subdomain (9 parameters), and domain (3 parameters) membership, in addition to a constant term (not shown).

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

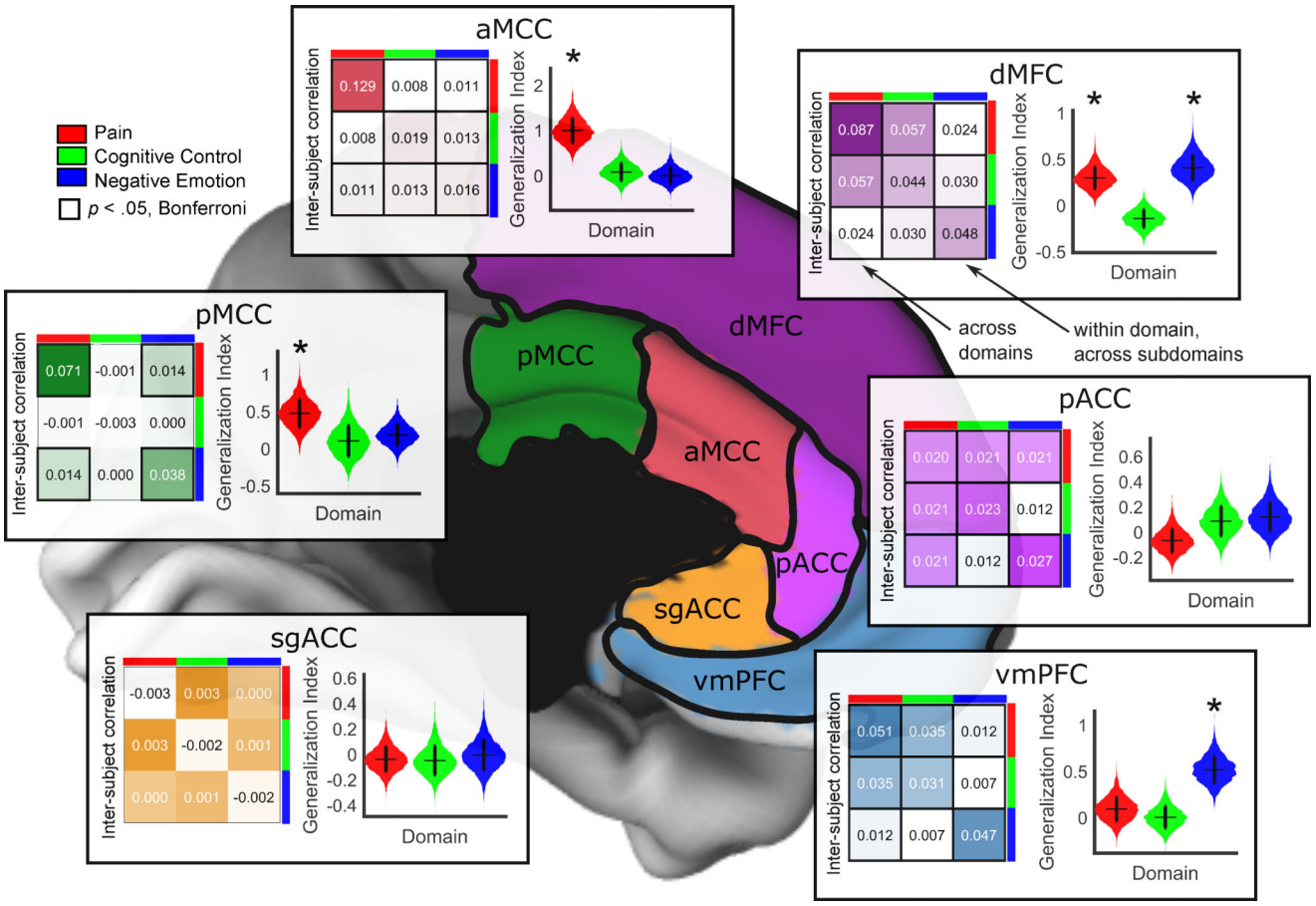

Figure 2.

Regional assessment of generalizable representations in MFC. Inset brain rendering depicts the anatomical parcellation of MFC using the four-region model of cingulate cortex[27] comprising anterior midcingulate cortex (aMCC), perigenual anterior cingulate cortex (pACC), subgenual anterior cingulate cortex (sgACC) and posterior midcingulate cortex (pMCC), in addition to ventromedial PFC (vmPFC) and dorsomedial frontal cortex (dMFC). Surrounding panels depict the average inter-subject Pearson correlations (36,315 pairwise correlations computed from $n = 270$ participants) both within domains but from different subdomains ($n = 2{,}700$ pairwise correlations) and between different domains ($n = 8{,}100$ pairwise correlations, left) in addition to bootstrap distributions of the generalization index computed from the full sample (b = 5,000 bootstrap samples, right). Correlations with significantly positive values (bootstrap test, Bonferroni corrected $P < .05$) have solid black borders. The generalization index quantifies the extent to which brain activity within a region is similar within a domain (across different subdomains) but is different across domains. *FDR $q < .05$ corrected.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

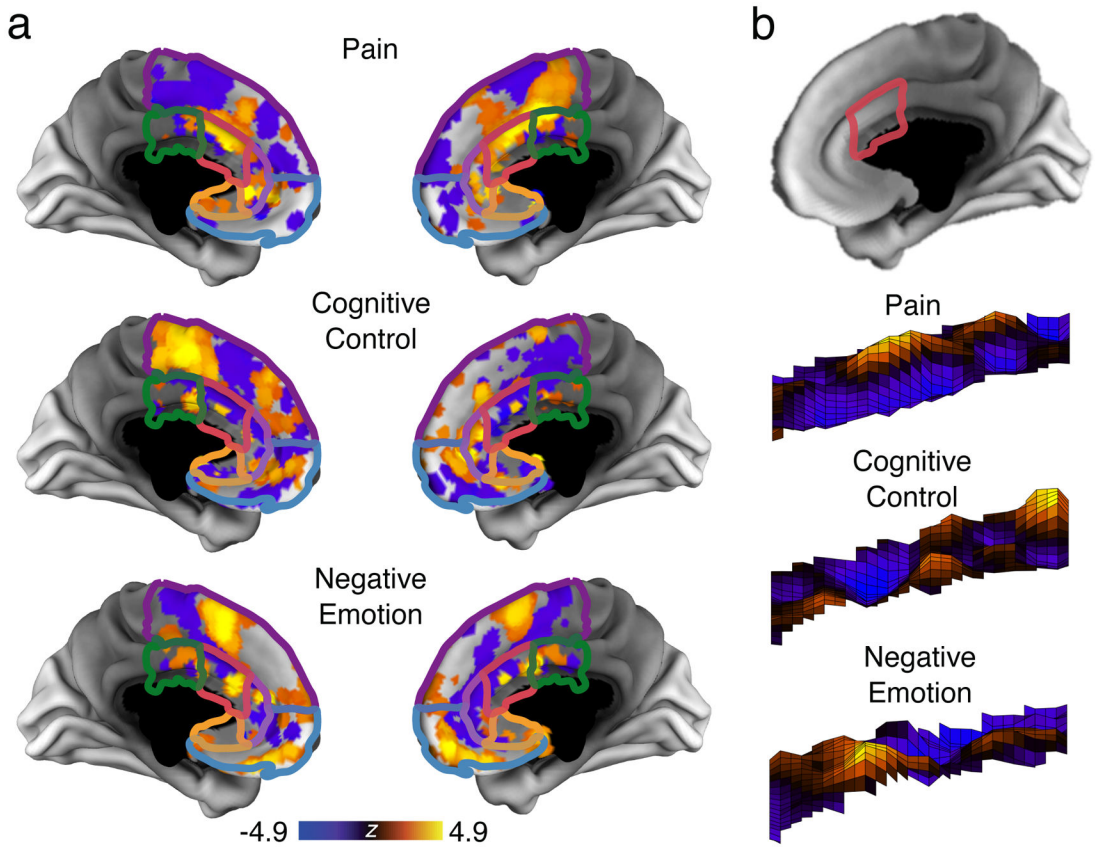

Figure 3.

Identifying latent brain representations that predict the occurrence of distinct functional domains in each region of interest. (**a**) Latent patterns of activity that generalize across studies and subdomains, but are specific for the domains of pain, cognitive control, and negative emotion, extracted using Partial Least Squares separately for each region and thresholded at $P < .05$ uncorrected for display ($n = 270$ participants). (**b**) Expanded view of latent patterns in aMCC in the left hemisphere. Images are displayed using radiological convention.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

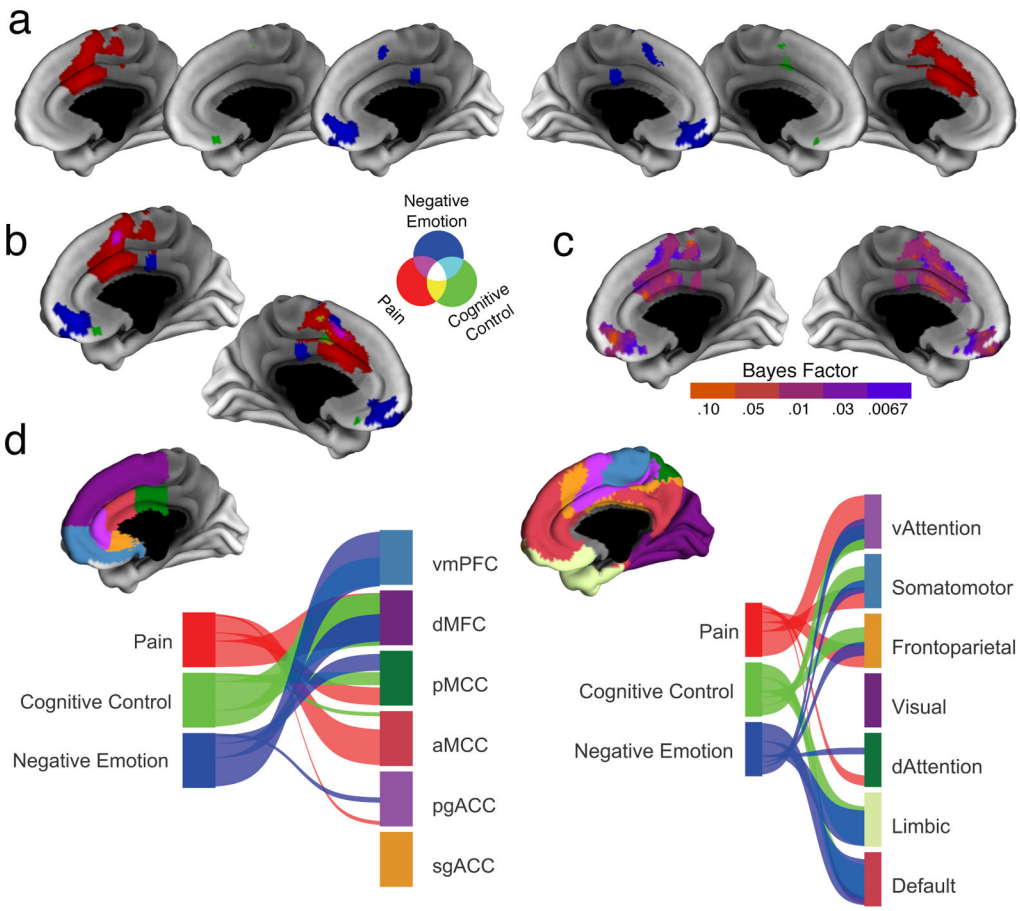

Figure 4.

Representational mapping of pain, cognitive control, and negative emotion in MFC. (**a**) Searchlight maps display where local patterns of brain activity are consistent with domain-specific representation of pain (red), cognitive control (green), and negative emotion (blue; n = 270 participants). (**b**) Additive conjunction of searchlight maps, with each domain mapped onto orthogonal dimensions in the RGB color space. Overlap between pain and cognitive control is depicted in yellow, whereas overlap between pain and negative emotion is colored in magenta. Maps are thresholded at a $P < .05$, two-tailed, uncorrected cutoff to highlight any possible overlap (n = 270 participants). (**c**) Brain maps of Bayes Factors indicating relative evidence against overlap among the three domains at each voxel. Smaller values indicate evidence against overlap; values less than .1 are considered strong evidence (n = 270 participants). (**d**) Riverplots depict the similarity between searchlight maps and anatomical parcellation of MFC (left) and functional parcellation of cortical regions from resting-state data[75] (right). The thickness of lines indicates the degree of correspondence between sets. Images are displayed using radiological convention.

Author Manuscript Author Manuscript Author Manuscript Author Manuscript

Table 1

Bayesian Information Criterion (BIC) weights and adjusted R^2 for selected models

Region	Study and Subdomain (28)	Pain (29)	Cognitive Control (29)	Negative Emotion (29)	Full Model (31)	Adjusted R^2 (Optimal Model)
pMCC	< .0001	0.0673	< .0001	< .0001	**0.9327**	.0220
aMCC	< .0001	**0.9901**	< .0001	< .0001	0.0099	.0342
pACC	**0.4686**	0.0342	0.1411	0.3484	0.0077	.0134
sgACC	**0.7986**	0.0704	0.0770	0.0536	0.0004	.0006
vmPFC	< .0001	< .0001	< .0001	**0.9669**	0.0331	.0567
dMFC	< .0001	< .0001	< .0001	< .0001	**1.0000**	.0831
MFC	< .0001	< .0001	< .0001	< .0001	**1.0000**	.0934

Bold font indicates model with highest BIC weights, and adjusted R^2 for this optimal model in each region is listed, based on the total variation in the data. BIC weights sum to 1 for each region. The number of free parameters in each model is listed in parenthesis. pMCC = posterior midcingulate cortex, aMCC = anterior midcingulate cortex, pACC = perigenual anterior cingulate cortex, sgACC = subgenual anterior cingulate cortex, vmPFC = ventromedial prefrontal cortex, dMFC = dorsal medial frontal cortex, MFC = medial frontal cortex.

Nat Neurosci. Author manuscript; available in PMC 2018 July 01.