

## ARTICLE

Received 8 May 2014 | Accepted 25 Sep 2014 | Published 17 Nov 2014

DOI: 10.1038/ncomms6380

# Separate neural representations for physical pain and social rejection

Choong-Wan Woo<sup>1,2</sup>, Leonie Koban<sup>1,2</sup>, Ethan Kross<sup>3</sup>, Martin A. Lindquist<sup>4</sup>, Marie T. Banich<sup>1,2</sup>, Luka Ruzic<sup>1,2</sup>, Jessica R. Andrews-Hanna<sup>2</sup> & Tor D. Wager<sup>1,2</sup>

Current theories suggest that physical pain and social rejection share common neural mechanisms, largely by virtue of overlapping functional magnetic resonance imaging (fMRI) activity. Here we challenge this notion by identifying distinct multivariate fMRI patterns unique to pain and rejection. Sixty participants experience painful heat and warmth and view photos of ex-partners and friends on separate trials. FMRI pattern classifiers discriminate pain and rejection from their respective control conditions in out-of-sample individuals with 92% and 80% accuracy. The rejection classifier performs at chance on pain, and vice versa. Pain- and rejection-related representations are uncorrelated within regions thought to encode pain affect (for example, dorsal anterior cingulate) and show distinct functional connectivity with other regions in a separate resting-state data set ( $N=91$ ). These findings demonstrate that separate representations underlie pain and rejection despite common fMRI activity at the gross anatomical level. Rather than co-opting pain circuitry, rejection involves distinct affective representations in humans.

<sup>1</sup>Department of Psychology and Neuroscience, University of Colorado, Boulder, Colorado 80309, USA. <sup>2</sup>Institute of Cognitive Science, University of Colorado, Boulder, Colorado 80309, USA. <sup>3</sup>Department of Psychology, University of Michigan, Ann Arbor, Michigan 48109, USA. <sup>4</sup>Department of Biostatistics, Johns Hopkins University, Maryland 21205, USA. Correspondence and requests for materials should be addressed to T.D.W. (email: Tor.Wager@Colorado.Edu).

The neural and psychological mechanisms of social rejection are the subject of intense study<sup>1–5</sup>; rejection is experienced as *painful*, and is an important risk factor for physical and mental illnesses<sup>6,7</sup> and other functional impairments<sup>8</sup>. How rejection experiences are represented in the brain has great theoretical significance for understanding social and emotion processes<sup>9</sup> and translational implications for assessing and treating medical conditions influenced by social factors<sup>10,11</sup>.

The ‘shared representation’ theory of social pain, the most influential theory of adverse social experience, suggests that rejection and related experiences piggyback on brain systems evolved to represent physical pain<sup>1–3</sup>. Experiences of social rejection and physical pain indeed have many common psychological and biological attributes. For example, both are often described using very similar words<sup>1</sup> and involve similar biological regulatory systems, such as endogenous opioids and oxytocin<sup>12,13</sup>. Critically, recent functional magnetic resonance imaging (fMRI) findings suggest that physical pain and social rejection activate common brain regions. These findings have been interpreted as evidence that pain and rejection share common brain representations<sup>2,4,5</sup>. Although the ‘shared representation’ theory does not claim that the neural systems underlying pain and rejection are exactly the same, it claims that shared representations exist in the brain regions thought to represent affective and sensory information in a modality-independent way. These regions include those important for representing affective distress across many conditions (for example, the dorsal anterior cingulate [dACC] and anterior insula [aINS])<sup>2,14–16</sup> and those that represent somatosensory information more specifically (for example, dorsal-posterior insula [dpINS] and S2)<sup>4,5</sup>.

However, overlapping fMRI activity within these regions does not necessarily imply shared representations at all levels of analysis (for example, neuronal population codes)<sup>17–19</sup>. For example, dACC and aINS show similar fMRI activity during the experience of positive and negative emotions<sup>20</sup> and also during diverse sensory and cognitive processes that have little to do with pain or rejection<sup>21,22</sup>. Thus, fMRI-based summaries of these regions often focus on potential *common* functions, such as surprise<sup>23</sup> or salience<sup>24</sup>. However, the dACC contains multiple, functionally specific subpopulations of neurons<sup>25</sup>, including nociceptive-specific ones<sup>26</sup> and those that code for various motivationally relevant events, such as reward expectancy<sup>27</sup> or predicted cognitive demands<sup>28</sup>. Even if the fMRI activity evoked by pain, rejection and other states is overlapping, their underlying neural representations may be non-overlapping. Thus, the question of whether representations of pain and rejection within dACC, aINS, dpINS, and S2 are common or distinct is unresolved.

In this paper, we take a fresh look at the question of shared neural representations for pain and rejection, using multivariate pattern analysis (MVPA)<sup>29</sup> combined with a rigorous inferential logic, the ‘separate modifiability’ criterion<sup>30</sup>, for assessing shared versus distinct representations at the fMRI multivariate pattern level. The ‘separate modifiability’ criterion requires two different measures (here, two fMRI multivariate patterns) to be separately modulated by two manipulations without cross-influences. For example, if two fMRI multivariate patterns can be identified such that one pattern responds to manipulations of pain but *not* rejection, and the other pattern responds to rejection but *not* pain, the patterns are *separately modifiable*<sup>30</sup>. This criterion is well suited for ruling out a single-process account of function, such as intensity or salience. The rationale for this is laid out in detail in the Discussion and ref. 30, but briefly, if two brain measures are separately modified by the two manipulations, then no single process is sufficient to account for the results. Therefore,

separate modifiability can provide strong evidence for the existence of distinct brain representations for pain and rejection.

We examined the following three specific questions in this paper: (i) whether global and local multivariate fMRI patterns that are separately modifiable by pain and rejection can be identified (*aim 1*), (ii) whether or not those multivariate patterns for pain and rejection are correlated within and across regions (*aim 2*) and (iii) whether the multivariate patterns for pain and rejection engage distinct functional brain networks (*aim 3*). Even if standard fMRI activation is insufficient to capture activity in distinct neural circuits, multivariate patterns of fMRI activity may be able to capture neuronal population codes specific to distinct sensory and affective experiences<sup>17,31,32</sup>. Thus, it may still be possible to identify multivariate patterns that respond only to manipulations of pain or rejection, and thus demonstrate separate modifiability of the two processes within and across brain regions.

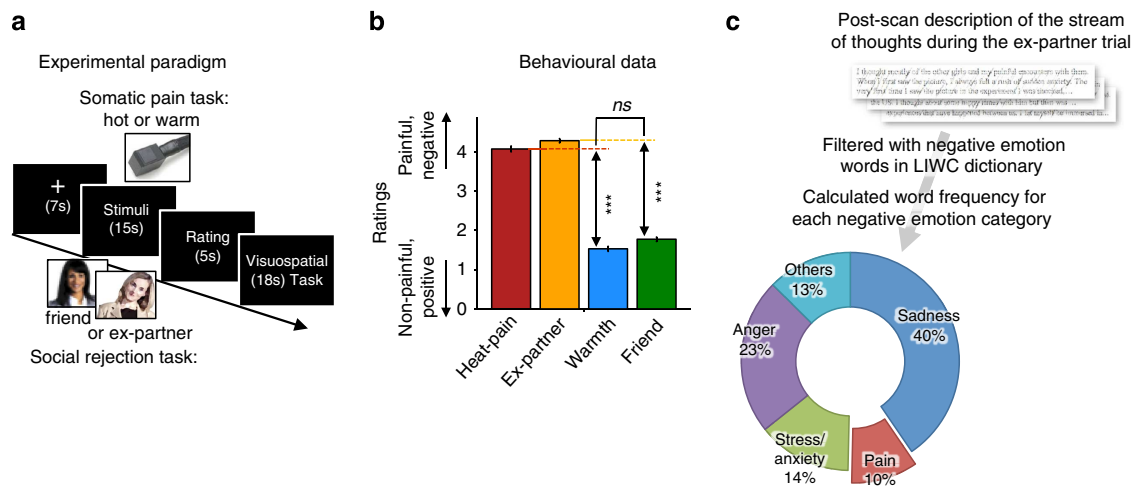
In Study 1, 60 participants were scanned with fMRI while they experienced physical pain and social rejection stimuli on separate trials (Fig. 1a). We recruited individuals who recently experienced an unwanted break-up with their romantic partners and felt intensely rejected. In the social rejection task, in order to elicit emotions of social rejection, participants viewed a headshot photograph of their ex-partner (‘Ex-partner’ condition) or a close friend (‘Friend’ condition). In the somatic pain task, we delivered painful heat (‘Heat-pain’ condition) or warm heat (‘Warmth’ condition) to the left volar forearm. Data from part of this sample ( $N=40$ ) were previously published<sup>4,32</sup>, but the analyses we performed here are qualitatively different from those in previous reports; here we focus on identifying an fMRI multivariate pattern for rejection and assessing its relationship with somatic pain representations (see Discussion for additional details).

Through a series of analyses, we challenge the ‘shared representation’ theory by identifying whole-brain fMRI multivariate patterns separately modifiable by pain and rejection. In addition to separate modifiability at the whole-brain level, multivariate patterns coding for pain and rejection within the dACC, aINS, dpINS, and S2 are uncorrelated and separately modifiable. Resting-state connectivity analyses on a separate sample (Study 2,  $N=91$ ) reveal that two multivariate patterns for pain and rejection encoded in the same set of dACC voxels have distinct functional connectivity patterns with the rest of the brain. These findings suggest that pain and rejection are distinct types of affect, with independent representations co-localized in similar gross anatomical regions.

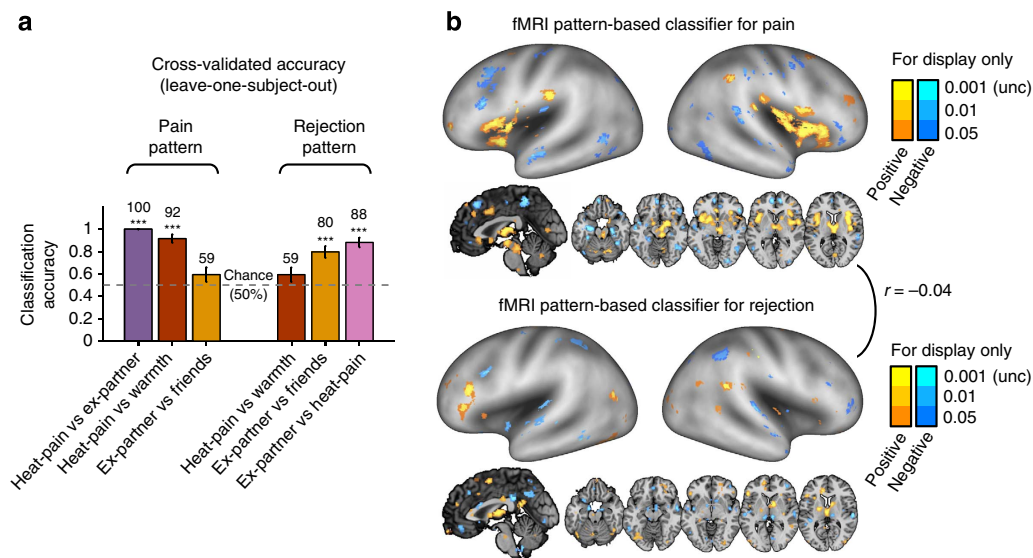
## Results

**Behavioural results.** In Study 1, both the Heat-pain and Ex-partner conditions elicited substantial negative affect on numerical rating scales ( $M_{\text{Heat-pain}} = 4.07 \pm 0.55$  (s.d.);  $M_{\text{Ex-partner}} = 4.28 \pm 0.39$  on a five-point scale), compared with their respective control (Friend and Warmth) conditions,  $t_{\text{Heat-pain versus Warmth}}(59) = 29.07$ ,  $P < 0.001$ ,  $t_{\text{Ex-partner versus Friend}}(59) = 28.37$ ,  $P < 0.001$ . Pain and rejection manipulations elicited equally strong increases in negative affect (for Heat-pain versus Ex-partner,  $\beta = 0.066$ ,  $t(118) = 0.26$ ,  $P = 0.80$ ).

Participants provided verbal descriptions of their experiences while viewing their ex-partner’s photo after the experiment. A word frequency analysis of these descriptions using the LIWC2007 software<sup>33</sup> revealed that, as expected, the photos elicited painful emotions: 9.7% of all negative emotion words provided by participants were pain-related words (for example, painful, hurt and so on), with 23% of participants using at least one pain-related word. The photos also evoked other negative emotions, including sadness (for example, sad, cry; 40.3% of negative emotion words used, and 85% of participants), anger



**Figure 1 | Behavioural results.** (a) Experimental paradigm. The social rejection and somatic pain tasks each consisted of two consecutively administered runs of eight trials (that is, 16 total trials). The order of the two tasks was counterbalanced across participants. (i) Social rejection task: each trial in the social rejection task lasted 45 s and began with a 7-s fixation cross. Subsequently, participants saw a headshot photograph of their ex-partner ('Ex-partner' condition) or a close friend ('Friend' condition) for 15 s. A cue-phrase beneath each photo directed participants to think about how they felt during their break-up experience with their ex-partner or a specific positive experience with their friend. Subsequently, participants rated how they felt using a five-point scale. To reduce carryover effects between trials, participants then performed an 18-s visuo-spatial control task in which they saw an arrow pointing left or right and were asked to indicate which direction the arrow was pointing. Ex-partner versus Friend trials were randomly presented with the constraint that no trial repeated consecutively more than twice. (ii) Somatic pain task: the structure of somatic pain trials was identical to rejection trials with the following exceptions. During the 15-s thermal stimulation period, participants viewed a fixation cross and focused on the sensations they experienced during a hot (painful) or warm (non-painful) stimulus that was delivered (1.5-s temperature ramp up/down, 12 s at peak temperature) to their left volar forearm at temperatures calibrated for each person (for details, see Methods). They then rated the pain they experienced using a five-point scale. (b) Behavioural data from trial-by-trial pain and emotion rating ( $n = 60$ , eight trials for each condition). Error bars represent within-subject standard errors of the mean (s.e.m.). (c) Relative word frequency for negative emotion word categories among negative emotional words that participants used to describe the stream of thoughts while they were viewing the ex-partner's photo in the scanner after the fMRI scanning. We used the Linguistic Inquiry and Word Count dictionary<sup>33</sup> (LIWC) to categorize emotional words. \*\*\* $P < 0.001$ , multi-level generalized linear model.



**Figure 2 | Separate modifiability of fMRI pattern-based classifiers for pain and rejection.** (a) Cross-validated (leave-one-subject-out) accuracy in two-choice classification tests ( $n = 59$ ). The results demonstrated separate modifiability (each can be changed independent of the other) of the fMRI pattern-based classifiers. The dashed line indicates the chance level (50%), and the error bars represent standard error of the mean across subjects. (b) The distributed fMRI pattern maps in which voxel activity reliably contributes to the discrimination of pain (top panel) and rejection (bottom panel) from other conditions. The maps show thresholded voxel weights based on bootstrapping (10,000 samples) of SVMs for display only; all weights were used in classification.  $r$  between two pattern maps denotes Pearson's correlation of voxel weights. \*\*\* $P < 0.001$ , binomial test.

(for example, angry, mad; 23.3% of negative emotion words, and 97% of participants) and stress (for example, upset, confused; 14.2% of negative emotion words, and 77% of participants). Thus, rejection-related stimuli did not only evoke 'painful' experiences, but they also elicited multiple intense, negative emotions.

**Separate modifiability of fMRI pattern-based classifiers.** To identify multivariate patterns of fMRI activity that are separately modifiable by pain and rejection (*aim 1*), we trained linear support vector machines (SVMs)<sup>34</sup> to discriminate Heat-pain versus other conditions and Ex-partner versus other conditions with a leave-one-subject-out cross-validation<sup>35</sup>. We used whole-brain activation parametric maps that were masked by an a priori meta-analytic map associated with 'pain', 'emotion' and 'social'<sup>22</sup> (see Methods and Supplementary Fig. 1). In order to identify pain- and rejection-specific activation patterns if they exist, it is desirable to train the classifier on Heat-pain versus all other conditions (including Ex-partner) or Ex-partner versus others (including Heat-pain) rather than Heat-pain versus Warmth and Ex-partner versus Friend because the latter contrasts could capture signal from a single process, such as salience or general aversiveness. With the identified patterns for pain and rejection, we tested for separate modifiability by testing them on *out-of-sample* participants for Heat-Pain versus Ex-partner or Ex-Partner versus Heat-Pain, Heat-Pain versus Warmth and Ex-partner versus Friend contrasts.

As shown in Fig. 2a and Supplementary Fig. 2, the pain classifier discriminated Heat-pain versus Ex-partner with 100% accuracy in the test participants (95% confidence interval (CI): 100–100%,  $P < 0.00001$ ) and Heat-pain versus Warmth with 92% accuracy (CI: 84–99%,  $P < 0.00001$ ). However, it performed at chance for Ex-partner versus Friend (accuracy = 59%, with 95% CI: 47–72%,  $P = 0.193$ ), demonstrating responsivity only to somatic pain. The rejection classifier discriminated Ex-partner versus Heat-pain with 88% accuracy (95% CI: 80–96%,  $P < 0.00001$ ) and Ex-partner versus Friend with 80% accuracy (95% CI: 69–90%,  $P < 0.00001$ ). However, it performed at chance for Heat-pain versus Warmth (accuracy = 59%, with 95% CI: 47–72%,  $P = 0.193$ ), demonstrating responsivity only to rejection. These findings provide the first evidence for separate modifiability of whole-brain fMRI multivariate patterns for pain and rejection, suggesting the existence of functionally independent neural representations for pain and rejection.

Figure 2b displays the classifier voxel weights that reliably contribute to the classification of pain and rejection, respectively, based on bootstrap tests with 10,000 iterations (thresholded at  $P < 0.001$  uncorrected for display only; all voxel weights were used in classification). Heat-pain versus Warmth was predicted by activation in many regions associated with nociceptive processing and endogenous pain control<sup>36,37</sup>, such as right (contralateral) dpINS, bilateral S2, medial thalamus and periaqueductal gray (PAG). Ex-partner versus Friend was predicted by increased activity in dorsomedial prefrontal cortex (dmPFC), right temporal parietal junction (TPJ) and precuneus, which are reliably associated with mentalizing about others<sup>38,39</sup>, and several other regions associated with negative emotion and its regulation<sup>20,40,41</sup>, including thalamus, supplementary motor area and inferior frontal gyrus.

To statistically compare the two pattern maps for pain and rejection (*aims 2, 3*), we conducted a bootstrap test with 10,000 iterations: for each iteration, we trained linear SVMs for pain and rejection and calculated (i) their spatial correlation across voxel weights and (ii) the difference weight values at each voxel. The two pattern classifiers were uncorrelated with each other,  $r = -0.04$ ,  $P = 0.28$ , and showed different weights in the following brain regions: Activation was more strongly predictive of pain in the supramarginal gyrus, middle insula, dpINS and ventral insula, PAG, amygdala and thalamus (*warm* colours in Fig. 3a,  $q < 0.05$ , false discovery rate [FDR] corrected; for the full list of the regions, see Supplementary Table 1). Activation was more strongly predictive of social rejection in dmPFC, right inferior frontal gyrus, ventromedial prefrontal cortex, perigenual

anterior cingulate cortex, TPJ and precuneus (*cool* colours in Fig. 3a). The significant weight differences in these regions were mainly driven by significant positive weights for pain or rejection (Fig. 3b), but in some regions by significant negative weights; in particular, by negative weights for rejection in amygdala and thalamus.

**Multi-voxel pattern similarity analysis.** Pain and rejection-related patterns included significantly different weights in some known targets of ascending nociceptive pathways<sup>42</sup>, including dpINS, thalamus and PAG, but not others (for example, dACC, aINS and S2). All of these regions were activated in standard general linear model (GLM) analyses by both Heat-pain versus Warmth and Ex-partner versus Friend contrasts (Supplementary Fig. 3). Common GLM activations such as these results have been interpreted in terms of shared neural representations between pain and rejection in previous papers<sup>2–4</sup>.

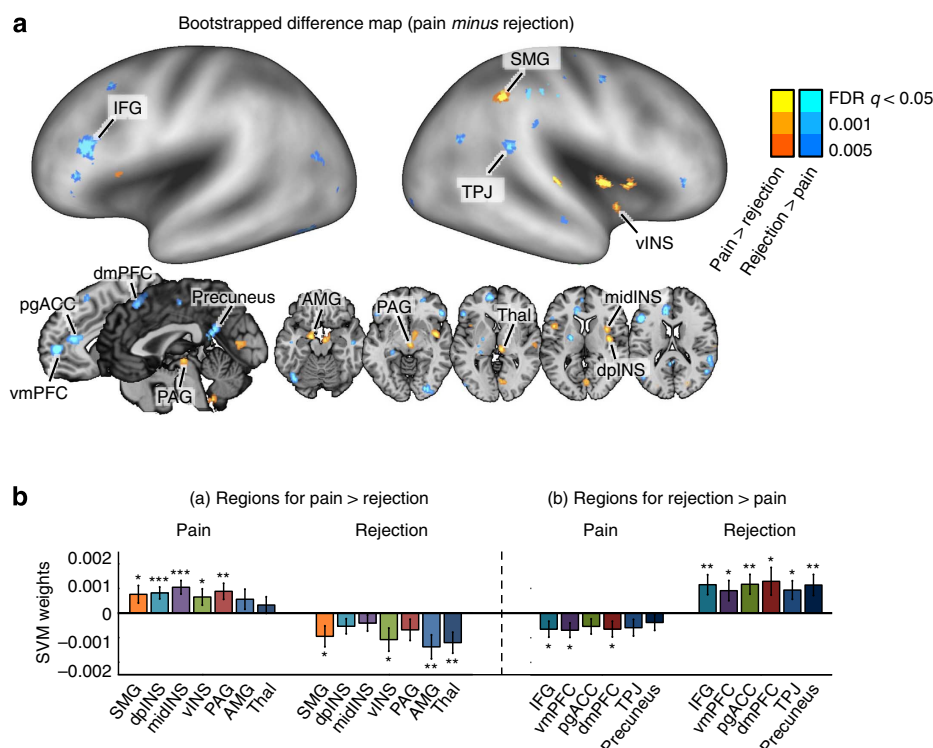
For this reason, and to identify shared *local* representations if they exist (*aim 2*), we examined whether the individual patterns of fMRI activity within these nociceptive pain-processing regions were similar or different for each participant with multi-voxel pattern similarity analysis<sup>17,43</sup>. The pattern correlations within dACC, aINS, S2 and dpINS were calculated for each individual participant (i) between the SVM weights for pain and rejection across voxels, and (ii) between the first-level contrast activation maps for Heat-pain versus Warmth and Ex-partner versus Friend (see Supplementary Fig. 4 for the detailed analysis procedure). 95% CIs and  $P$ -values were calculated based on bootstrap tests with 10,000 iterations (for the correlations between SVM weights) or based on group statistics (for the correlation between contrast images).

Pattern similarity analysis results supported the conclusion that pain- and rejection-related representations within pain-processing regions were distinct (Fig. 4a). No core pain-processing regions showed a significant correlation between (i) SVM classifier weights for pain and rejection (Fig. 4b, left) or (ii) fMRI contrast images for pain and rejection (Fig. 4b, right). Use of rank correlations yielded the same results.

**Cross-classification test results with local classifiers.** Although pain- and rejection-related fMRI patterns were uncorrelated within core pain-processing regions, this does not imply that there are no shared representations anywhere in the brain. To identify regions with shared and non-shared local patterns between pain and rejection (*aims 1, 2*), we conducted 'cross-classification' tests using a local pattern-based classification approach (*cf.* ref. 44). The cross-classification tests consisted of three steps. (i) Local regions were defined using spherical searchlight regions around centre voxels<sup>45</sup> or using functional parcellations<sup>46</sup>. (ii) Multivariate pattern classifiers using a defined local region were trained separately for Heat-Pain versus Warmth and Ex-partner versus Friend with leave-one-subject-out cross-validation. In this analysis, training classifiers on one condition versus its control condition is more desirable than the one condition versus all other conditions approach because here we sought to identify 'shared' local representations across pain and rejection, not specific and unique representations of pain and rejection. (iii) With the identified local patterns for pain and rejection, we tested for separate modifiability by applying them to *out-of-sample* participants for the Heat-Pain versus Warmth and Ex-partner versus Friend contrasts. Steps ii and iii were repeated for each local region across the whole brain.

If neural representations for pain and rejection are shared within a local region, the cross-validated accuracy for each classifier should be significant for both within- and cross-





**Figure 3 | The difference map between the fMRI pattern-based classifiers for pain and rejection.** (a) The difference map in which values represent reliable differences between two discriminant weights (SVM weights for pain minus SVM weights for rejection) based on bootstrapping of SVMs (10,000 samples). The difference map is thresholded at  $q < 0.05$ , FDR corrected. AMG, amygdala; dpINS, dorsal posterior insula; IFG, inferior frontal gyrus; midINS, middle insula; PAG, periaqueductal gray; pgACC, pregenual anterior cingulate cortex; preSMA, pre-supplementary motor area; SMG, supramarginal gyrus; Thal, thalamus; TPJ, temporal parietal junction; vINS, ventral insula; vmPFC, ventromedial prefrontal cortex. (b) Mean SVM weights for contiguous regions within the difference map (10,000 samples). The error bars represent standard error of the mean based on bootstrapping. \* $P < 0.05$ , \*\* $P < 0.01$ , \*\*\* $P < 0.001$ , one-sample  $t$ -test.

modality test contrasts. For example, a pattern trained on Heat-Pain versus Warmth would classify above chance for both Heat-Pain versus Warmth (within-modality) and Ex-partner versus Friend (cross-modality) in *out-of-sample* participants. To assess the robustness of the procedure across variations in the method, we conducted the cross-classification tests with four different local region definitions, including 6- and 10-mm radius spherical searchlights<sup>45</sup>, and 200- and 400-region functional parcellations<sup>46</sup>. As the results were comparable across all local region definitions (Supplementary Fig. 5), here we focus on the results of 6-mm radius spherical searchlights (Fig. 5 and Supplementary Fig. 6).

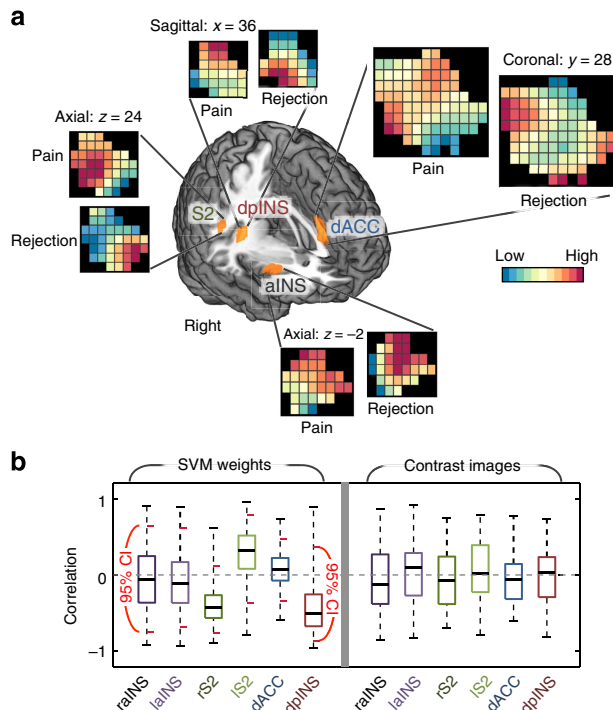
As shown in Fig. 5 and Supplementary Table 2, a number of regions showed significant cross-classification between pain and rejection (yellow in Fig. 5), implying shared representation at the local pattern level. These included the left parahippocampal and fusiform gyri, retrosplenial cortex, right TPJ, posterior cingulate cortex and striatum ( $q < 0.05$ , FDR corrected). Critically, none of these regions are primary 'pain-processing' regions<sup>36,42</sup>. Many other regions showed significant classification accuracy for both pain and rejection (orange in Fig. 5), but without significant cross-classification (that is, separate modifiability). Such regions did include those targeted by primary nociceptive afferents (for example, dACC, aINS and dpINS/S2).

**Distinct functional connectivity with dACC patterns.** In Study 2, we tested whether the distinct multivariate patterns for pain and rejection in the same local region of interest within the dACC are functionally connected to similar or different brain networks

in a separate resting-state fMRI data set ( $N = 91$ ; aim 3). In previous sections, we showed that the dACC was commonly activated by both pain and rejection (Fig. 4a), but we identified distinct multivariate fMRI patterns within the dACC, which were separately modifiable by pain and rejection (Fig. 5). However, it is unknown whether these multivariate patterns for pain and rejection within the dACC connected to *distinct* brain circuits that correspond to distinct affective processes and large-scale brain networks.

To examine this question, we used the Study 1 data to obtain dACC local-region patterns—defining the region as the portion of dACC that showed strong separate modifiability in the previous analyses—separately for Heat-Pain versus Warmth and Ex-partner versus Friend. This yielded dACC<sub>pain</sub> and dACC<sub>rejection</sub> multivariate patterns optimized to specifically detect each experience (Fig. 6a). Then, we calculated the expression of each pattern at each time point in the Study 2 resting-state fMRI data, yielding two 'seed' time courses based on the two patterns. In order to calculate the strength of expression of each pattern, we used the dot-product of a vectorized activation map within the dACC region of interest with the pattern classifier weights. Finally, we performed random-effects GLM analyses to estimate functional connectivity with each seed's time course across the brain and the connectivity difference maps.

Figure 6a shows the multivariate pattern classifiers within dACC for pain and rejection and the cross-validated classification accuracy results, indicating that the dACC patterns for pain and rejection were distinct,  $r = -0.04$ ,  $P = 0.23$  and separately modifiable by pain and rejection, as we expected. The whole-brain functional connectivity patterns for dACC<sub>pain</sub> and



**Figure 4 | Multi-voxel pattern similarity analysis for pain-processing regions.** (a) SVM classifier weight patterns within regions-of-interest (ROIs) from group-level GLM results. The regions were activated in both contrasts, Heat-pain versus Warmth and Ex-partner versus Friend, and have been implicated in both pain and rejection. Here, a liberal threshold ( $P < 0.05$ , uncorrected) was applied to use large enough regions for multi-voxel pattern similarity analyses (the averaged number of voxels across six ROIs = 213). For GLM results corrected for multiple comparisons, see Supplementary Fig. 3. The patterns presented here are the averaged SVM classifier weights from bootstrap tests (10,000 samples). aINS, anterior insula; dACC, dorsal anterior cingulate cortex; dpINS, dorsal posterior insula; S2, secondary somatosensory cortex. (b) Left: the bootstrap test results for SVM classifier weight correlations. Right: the group-level correlations between fMRI activations of contrast images for pain and rejection ( $n = 59$ ). The short red lines in the left panel indicate 95% confidence intervals obtained from bootstrap tests (10,000 samples). No regions showed significant correlations between SVM classifier weights, and no regions showed significant average correlations between patterns of contrast values across participants.

dACC<sub>rejection</sub> (Fig. 6b) and their difference map (Fig. 6c) showed that dACC<sub>pain</sub> and dACC<sub>rejection</sub> engage distinct functional connectivity patterns. For example, dACC<sub>pain</sub> showed stronger functional connectivity with thalamus, posterior insula, midbrain regions, ventral and anterior part of medial frontal regions, posterior cingulate cortex and cerebellum (yellow in Fig. 6c), whereas dACC<sub>rejection</sub> was more strongly associated with ventral and dorsal lateral prefrontal cortex, parietal cortex, temporal pole and middle temporal lobe, dmPFC and TPJ (blue in Fig. 6c). These networks correspond well to the regions with differential predictive weights for pain and rejection in Study 1, but they were identified using distinct methodology (resting-state fMRI connectivity) in an independent sample in the absence of overt pain or social rejection.

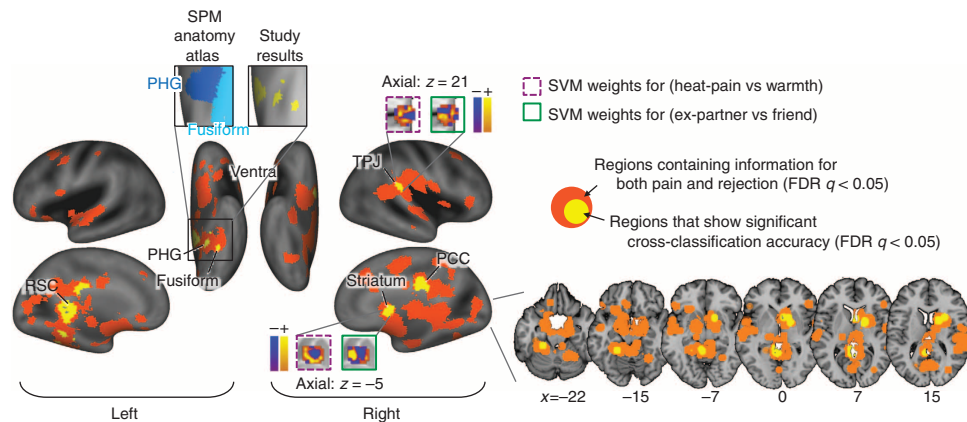
## Discussion

The idea that there is substantial overlap in the neural representations of pain and rejection within core pain-processing

regions such as dACC and aINS has been influential in various fields<sup>1,47</sup>. However, claims about shared representation in the previous studies have been based on findings of overlapping univariate fMRI activity between pain and rejection<sup>2,4,5</sup>, which is not anatomically specific enough to bear on the question of whether the underlying neural representations are similar. Here we used a more fine-grained analysis technique to demonstrate that the overlapping activity arises from *distinct neural representations*<sup>18</sup>. MVPA is more likely to reflect population codes across neurons<sup>43,45,48</sup>, as they have demonstrably greater sensitivity and specificity to particular types of mental events<sup>17,31,32</sup>. We demonstrate that multivariate patterns encoding the intensity of pain and rejection are separately modifiable<sup>30</sup> by showing that pain and rejection each influenced distinct, uncorrelated fMRI patterns at the whole-brain level and within 'pain-processing' regions.

The criterion of separate modifiability, which is a stronger version of the classic 'double dissociation' logic, strongly implies the existence of distinct, non-shared neural representations for pain and rejection. The criterion also rules out a single-process interpretation (for example, based on salience, arousal or general aversiveness) of the fMRI multivariate patterns that we observed within and across regions. The logic of this claim is laid out more completely in ref. 30, but it deserves a brief treatment here. We start by hypothesizing that the effects of both pain and rejection manipulations on dACC (or any other region discussed here) are related to a single process (for example, salience), and asking whether this hypothesis can be falsified with the separate modifiability criterion. Let us call the pain-related fMRI pattern in dACC 'Pattern A.' If rejection does not activate Pattern A, then there are three possibilities: (i) Pattern A does not reflect a common process engaged by both pain and rejection; (ii) Pattern A is not sensitive to the common process (for example, is too noisy) or (iii) rejection activates the common process too weakly. We can rule out (ii), because pain does activate Pattern A. Alternative (iii) is possible, although we note that subjective ratings of distress were comparable across pain and rejection (Fig. 1b). Now, let us consider that rejection activates Pattern B, but pain does not. If Pattern B also reflects activation of the same common process (for example, salience), then (iii) is ruled out, because rejection is demonstrated to be a more potent manipulation of Pattern B than pain. Thus, we conclude (i): 'Pattern A does not reflect a common process engaged by both pain and rejection'. The same logic applies to Pattern B, and we can infer that neither pattern encodes a mental process that is shared by both pain and rejection. This is important in relation to recent literature interpreting fMRI activity in the dACC, aINS and other regions in terms of general processes (for example, both regions are part of the 'salience network'<sup>24</sup>). Based on our findings and the separate modifiability logic, standard brain activation results may reflect non-specific, domain-general processes, but the multivariate patterns we identify here cannot.

The existence of at least some neural dissociation of pain and rejection is trivial, in part because pain and rejection-related stimuli are initially processed in different sensory systems (that is, somatosensory versus visual), and also because they differ in their cognitive associations and implications. What is at stake, however, is whether core affective processes (for example, 'pain affect') are shared<sup>2,47</sup>, and whether the type of emotional distress elicited during pain is the same as that elicited during social rejection. Germaine to this issue, pain- and rejection-related multivariate patterns were uncorrelated and separately modifiable even within the regions thought to encode 'pain affect,' including the dACC and aINS. In addition, although both pain and rejection activate somatosensory areas, such as S2 and dpINS, we



**Figure 5 | Cross-classification test results with local pattern classifiers.** We conducted whole-brain searches for candidate regions for shared neural processes between pain and rejection with 6-mm radius spherical searchlights around centre voxels. To determine the candidate regions, we conducted cross-classification tests (with leave-one-subject-out cross-validation) among the regions that accurately classify both pain (Heat-pain versus Warmth) and rejection (Ex-partner versus Friend). The cross-classification test consisted of the following steps: (i) A local classifier was trained and cross-validated separately for one condition, and (ii) the classifier was tested on the other condition in the same leave-one-subject-out cross-validation test sample. Here we thresholded all results at FDR  $q < 0.05$  ( $P < 0.0038$ ) and visualized the 6-mm sphere coverage. PCC, posterior cingulate cortex; PHG, parahippocampal gyrus; RSC, retrosplenial cortex; TPJ, temporal parietal junction. The anatomical atlas was from the SPM anatomy toolbox<sup>70</sup>.

demonstrate here that they activate different multivariate patterns within these areas. Thus, together, the findings indicate separable neural representations that are co-localized at the gross anatomical level.

It is important to consider how the multivariate patterns we identified might relate to those at more spatially precise levels of analysis, including optical imaging and single-unit recording. Ultimately, neural processes that underlie representations of mental constructs can be investigated at multiple levels of analysis. Among methods available in healthy humans, fMRI multivariate patterns may be particularly useful in inferring representational similarity<sup>43</sup>, as such patterns can be sensitive to population codes distributed across large numbers of individual neurons. fMRI multivariate patterns have obvious limitations in spatial resolution compared with single-unit recording<sup>49</sup>, in large part because blood-oxygen-level-dependent (BOLD) fMRI is sensitive to changes in microvascular beds that subserve many neurons, and the BOLD response thus acts as a spatial low-pass (blurring) filter sensitive to some, but not all, neuronal-level processes<sup>48,50</sup>. Because of the spatial blurring properties of fMRI, finding that two manipulations activate *similar* fMRI patterns does not necessarily imply similar neuron-level representations. However, finding *dissociable* fMRI patterns, as we did here, implies that the neuron-level population codes are different. An additional advantage of working at the fMRI pattern level is that fMRI can detect representations encoded across millions or billions of neurons, including interactions among large-scale brain networks, complementing approaches at finer spatial scales.

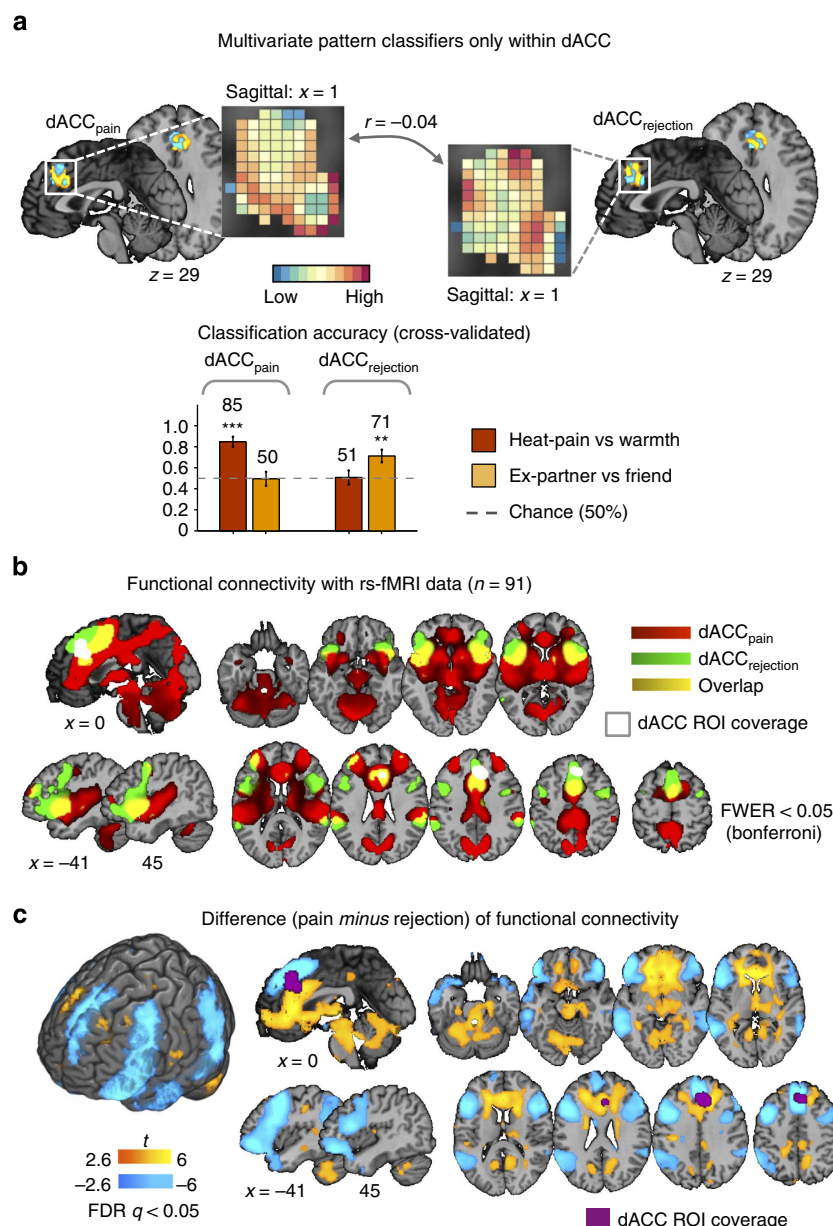
It is also important to note that our findings do not imply that there are *no* shared representations anywhere in the brain. To the contrary, we identified a number of regions showing similar fMRI patterns for pain and rejection. However, these regions were located outside of core pain-processing brain systems<sup>36,42</sup>. Thus, shared representations may exist (provisio considerations outlined above), but they are not likely associated with nociceptive pain. Rather, they are likely to be broadly associated with processes related to context, memory, motivated action, social inferences, and/or endogenous regulation (for example, opioid response in striatum<sup>13,51</sup>), among other processes. However, following the logic outlined above, we cannot claim that regions showing similar fMRI patterns are also similar at lower levels of analysis.

The idea that pain and rejection share neurophysiological mechanisms has also been supported by neurochemical and pharmacological findings<sup>11,13,52</sup>, which at first may seem contradictory to our conclusion. However, similar neurochemical responses to pain and rejection (for example, endogenous opioids and oxytocin<sup>12</sup>) may reflect common physiological *regulatory* responses to general aversive events (for example, stress). For example, the endogenous opioid system, implicated in both pain and rejection, also plays a key role in stress-induced analgesic effects<sup>53</sup>, which could explain the analgesic effects of acute experiences of social exclusion or separation<sup>54,55</sup>. In addition, the effects on rejection experience of 'pain medications' like acetaminophen<sup>11</sup> may be related to nonspecific neurochemical pathways for aversive affective states, such as the serotonin system<sup>56</sup>. Therefore, the involvement of common neuropeptides and regulatory mechanisms in both pain and rejection do not necessarily imply shared neural representations.

More broadly, the multivariate pattern-based approach we use here has potential for identifying and targeting many types of affective processes more specifically, which could enable significant progress in understanding the structure of emotion and its regulatory mechanisms<sup>9</sup> and providing more specific and robust links between brain representations and psychopathology<sup>57</sup>. Importantly, our results do not imply that there is no functional relationship between pain and rejection, or that one cannot influence the other. In fact, such interactions may underlie, for example, the increased rate of pain disorders observed following emotional trauma<sup>58</sup>. Developing and using fMRI pattern-based markers that are sensitive and specific to different psychological experiences will provide new leverage points for examining their underlying neural processes, the regulatory effects on the processes, and inter-relationships among different psychological processes. This approach could ultimately contribute to developing tailored interventions for specific psychological problems by identifying unique components of each condition, and it could also guide diagnosis and treatment of mental disorders by providing provisional fMRI-based biomarkers for them<sup>59</sup>.

This study also has implications for how networks are identified in fMRI studies. We showed that fMRI activity within localized anatomical regions (for example, the dACC region





**Figure 6 | Difference in functional connectivity patterns with dACC pattern classifiers for pain and rejection.** (a) The multivariate pattern classifiers for pain and rejection within the dorsal anterior cingulate cortex (dACC) region-of-interest (ROI) (a top). The dACC ROI was defined by the searchlight analysis results presented in Fig. 5, which showed that the dACC ROI contained information for both pain and rejection conditions, but the patterns were distinct and non-transferrable. Using the dACC ROI, we trained linear SVMs to discriminate pain and rejection from their respective control conditions, and tested on out-of-sample participants. The pattern weights were uncorrelated with each other,  $r = -0.04$ . The cross-validated (leave-one-subject-out) accuracy in two-choice classification tests demonstrated separate modifiability of the pattern classifiers (a bottom). The dotted line indicates the chance level (= 50%), and the error bars represent standard error of the mean across subjects.  $**P < 0.01$ ,  $***P < 0.001$ , binomial test. (b) Seed-based functional connectivity with dACC pattern classifiers for pain and rejection. Here seeds were pattern expression values (the dot-product of a vectorized activation map and SVM weights within dACC) for pain and rejection. The functional connectivity for each condition was calculated with independent resting-state fMRI data ( $n = 91$ ). These maps were thresholded at family-wise error rate (FWER) < 0.05 using Bonferroni correction. Here we used Bonferroni correction instead of false discovery rate (FDR) because the latter provided too liberal thresholds for these functional connectivity patterns (uncorrected  $P < 0.03$ ) and therefore yielded non-sensible maps. (c) Paired  $t$ -test results between two seed-based functional connectivity patterns. The results were thresholded at FDR < 0.05.

corresponding to Vogt's anterior mid-cingulate zone<sup>60</sup>) can be decomposed into two orthogonal multivariate patterns that, when used as 'seed' patterns in connectivity analyses, are associated with two distinct large-scale brain networks. This finding supports the idea that local regions such as dACC can contain multiple, distinct neural population codes<sup>25</sup> that encode distinct mental processes. Previous neuroimaging studies have generally

assessed functional connectivity by averaging activity within localized 'seed' regions into a single variable, which implicitly assumes that all voxels within the 'seed' region share a common pattern of connectivity. Therefore, the current approach is qualitatively different from the traditional approach to functional connectivity, and has broad implications for network analysis in future fMRI studies.



Previously, we published two papers using part of the current data set ( $N=40$ , compared to  $N=60$  currently)<sup>4,32</sup>. The first publication<sup>4</sup> reported overlapping fMRI activity across pain and rejection within many pain-processing regions, based on univariate GLM analyses. We interpreted these findings in terms of shared representations between pain and rejection. The co-localized representations may still point to important interactions between pain and rejection. However, the current findings at the multivariate pattern level suggest that the representations of pain and rejection are in fact distinct within and across regions. This conclusion, which differs from our previous interpretation of the data, does not result from the addition of 20 more participants (see Supplementary Fig. 7), but rather from the different level of analysis. The second previous publication<sup>32</sup> used part of this data set ( $N=40$ ) to test the cross-study sensitivity and specificity of an fMRI-based brain marker for experimental pain. The focus of the current study is qualitatively different: The previous paper<sup>32</sup> focused on the development of a new fMRI-based marker for pain, whereas the current study focuses on examining, and ultimately challenging, current notions of shared representation across social and physical pain. We believe the current study provides a more comprehensive picture of the relationship between pain and rejection than previous work, and provides a step towards a better understanding of the interaction between pain and rejection and developing tailored interventions for each of them.

More specifically, the following analyses have been conducted only in the current study, not in the previous studies<sup>4,32</sup>: (i) we identified a multivariate fMRI pattern sensitive and specific to rejection; (ii) we demonstrated that the rejection pattern is qualitatively different from the pain pattern, and the rejection pattern did not respond to pain; (iii) we showed that pain and rejection yielded distinct multivariate patterns even in the overlapping regions activated by both pain and rejection; (iv) we identified brain regions with similar neural patterns for pain and rejection using a searchlight approach; and finally, (v) we showed that the distinct multivariate patterns for pain and rejection within the dACC engaged distinct functional brain networks using a separate resting-state data set ( $N=91$ ).

The present study has limitations that should be addressed in future studies. Given that we used a five-point Likert scale for pain and emotion ratings and applied only two different intensity levels of stimulation for both pain and rejection, the multivariate patterns we identify here may not be sensitive to fine-grained differences in the subjective intensity of pain and rejection experience. Therefore, these multivariate patterns cannot be regarded as fully characterized representations of pain and rejection experience. We cannot also rule out the possibility that some of the activity in the multivariate patterns we identified could be related to response preparation, attention and other consequences, instead of mental representations central to the *experience* of pain and rejection. This is an important issue with all MVPA analyses and fMRI results in general<sup>61</sup>, and therefore future work will be required to identify fMRI representations of pain and rejection by characterizing their specificity and generalizability across different tasks, individuals and studies. Particularly, the stimuli used in the current social rejection task (for example, ex-partner photos) refer to past experience and only indirectly to present circumstance, whereas the somatic pain task elicits acute thermal pain. Therefore, other social rejection tasks eliciting acute feelings of exclusion (for example, the Cyberball game task<sup>62</sup>) might provide a good additional comparison condition, although they are likely to elicit less intense rejection-related emotions than the current task. We also focused only on pattern-based classification across individuals

because this approach is more stable and generalizable than the within-subject approach. However, identifying individual-specific multivariate patterns might also be informative about the relationship between pain and rejection.

In conclusion, our results provide the first neuroimaging evidence that pain and rejection do not share neural representations within core pain-processing brain regions. Instead, the present findings suggest that pain and rejection are represented in distinct mesoscopic neural systems whose locations are conserved across individuals. This new approach, and the evidence for separately modifiable multivariate patterns for each type of 'pain,' can move the field beyond the search for shared processes across pain and rejection and provide a step towards identifying their unique components, developing tailored interventions for each of them, and obtaining a better understanding of the interactions between these two distinct affective processes.

## Methods

**Participants.** Sixty healthy, right-handed participants (31 females,  $M_{\text{age}}=20.8$ ,  $SD_{\text{age}}=3.0$ ) completed the social rejection and somatic pain tasks while undergoing scanning with fMRI (Study 1). All participants experienced an unwanted romantic relationship break-up within the past 6 months, and indicated that they felt rejected when thinking about their break-up; all participants rated high on a 1 (not at all rejected) to 7 (very rejected) scale asking how rejected they felt when they thought about their rejection experience ( $M=5.52$ ,  $s.d.=1.08$ ). Participants were recruited via flyers posted around Manhattan and advertisements posted on Facebook and Craigslist. The sample was comprised of 63% Caucasian, 15% Asians, 10% African Americans, 3% Hispanic and 8% other. All participants were native English speakers and received \$175 for their participation. fMRI task data (for example, trial onset times) were lost for one participant. Thus, all imaging analyses were completed on 59 participants. The study was approved by the Columbia University's Institutional Review Board.

In addition, 91 healthy, right-handed participants (48 females,  $M_{\text{age}}=20.8$ ,  $SD_{\text{age}}=2.8$ ) completed the resting-state fMRI task (Study 2), separately from the social rejection and somatic pain tasks. The sample was comprised of 87% Caucasian, 4% Asians, 2% African Americans and 7% Hispanic. All participants were native English speakers and received \$50 for their participation. The study was approved by the University of Colorado Boulder's Institutional Review Board.

All participants provided written informed consent. Participants were screened to ensure that they did not suffer from any neurological or psychiatric illness, experience chronic pain, take psychoactive medications, antihistamine or steroids, have metal in their bodies, or have a history of substance use or abuse.

**Social rejection task stimuli.** The social rejection task was designed after the following two lines of research: (i) fMRI research where photos provided by participants were used to elicit powerful emotions (for example, maternal love, romantic love and rejection)<sup>51</sup> and (ii) behavioural research where cues were used to recall autobiographical experiences of social rejection and to effectively reactivate distress related to social rejection<sup>63</sup>. The social rejection stimuli consisted of: (i) a *headshot photograph* of each participant's ex-partner or a same-gendered friend with whom they shared a positive experience around the time of their break-up, and (ii) *cue phrases* displaying beneath each photograph that led participants to focus on a specific experience they shared with each person. Participants provided these cue phrases before the day of fMRI scanning using a procedure developed in prior research<sup>64</sup>: they first wrote a specific description of break-up experience with their ex-partner and positive experience with their friend. Subsequently, they created a cue-phrase to capture the gist of their experience. They were reminded of the cues they generated and the experiences they referred to on the day of scanning following established procedures. We cropped all images to ensure the total area of the photograph taken up by the face to be constant across Ex-partner and Friend images. Ratings of the picture quality and attractiveness by a group of ten individuals who were blind to the study goals and hypotheses indicated that the level of picture quality and attractiveness for each photograph did not differ between Ex-partner and Friend photographs.

**Somatic pain task stimuli.** For somatic pain stimuli, thermal stimulations were delivered to left volar forearm, consisting of two levels: non-painful (rated as level 2 on a 10-point scale by participants) versus very painful (rated as level 8 on a 10-point scale; near the limit of pain tolerance). With a calibration procedure<sup>65</sup>, we chose two temperatures on a participant-by-participant basis to be sure that the subjective pain intensity was constant across participants. We carried out the calibration task before the day of fMRI scanning.

**Pain calibration task.** Pain calibration consisted of 24 trials where participants rated the pain induced by thermal stimulation ( $10^{\circ}\text{C s}^{-1}$  ramp up, 7 s at target temperature,  $10^{\circ}\text{C s}^{-1}$  ramp down) applied using a TSA-II Neurosensory Analyzer (Medoc Ltd) with a 16-mm thermode end-plate. Participants provided verbal ratings with a 0 (no sensation) to 10 (unbearable pain) scale. Thermal stimulations were applied in a fixed order through eight different, nonadjacent candidate skin sites on the participant's left volar forearm. On each trial, we used an adaptive procedure to estimate temperatures corresponding to level 2, 5 and 8 (referred to as 'low', 'medium' and 'high'). For the estimation procedure, a linear regression model continuously fit the data collected up to that point during the calibration session with *temperature* as the independent variable and *pain ratings* as the dependent variable. Although the temperature–pain relationship is not perfectly linear, it is relatively robust to noise and can be fit with few observations. To make the calibration more robust to idiosyncratic behaviour on individual trials, we defined trials for which the absolute value of the residual based on the initial regression was greater than three times the median absolute deviation as outliers, and the model was re-fit without the outliers. The predicted low, medium and high temperature levels from the regression model were used to determine the temperature applied on the subsequent trial. Throughout the calibration, we used a fixed, counterbalanced order of heat-levels (high, medium or low). The order was chosen to equalize the transitional probabilities and to balance stimulation at high, medium and low temperatures across the eight skin locations.

**Pre-scan task training.** Before fMRI scanning, the experimenter described each step of the Social Rejection task (referred to as the 'Photograph' task to participants) and the Somatic Pain task (referred to as the 'Heat' task to participants). The experimenter explained that during the 'Photograph' task, participants would see photographs of their ex-partner and friend and cue-phrases beneath each photograph. Participants were asked to look directly at each photograph and reflect on the thoughts and emotions each photograph-cue pair elicited in them. During the Somatic Pain task, participants were asked to stare at the fixation cross during each trial and focus on the sensations elicited by the somatic stimulus. Participants were then instructed how to rate their affect following each type of trial and how to perform the visuospatial control task (see Fig. 1a for a detailed description of these tasks).

**Pain and affect ratings and post-scan questionnaire.** Following each trial, participants rated how they felt or the pain intensity they experienced using five-point Likert scales (1 = 'very bad' or 'very painful', 5 = 'very good' or 'not painful'). For display purposes, these scores were *reversed* such that higher scores indicate worse mood or more pain (Fig. 1b). After participants exited the scanner, they were asked to write freely about the thoughts that came to mind regarding (i) their recent romantic relationship break-up and (ii) their positive experience with friends during the experiment in the fMRI scanner.

**Resting-state fMRI task.** During the resting-state fMRI task (5 min 34 s), participants were asked to relax with their eyes closed while remaining awake.

**fMRI acquisition and preprocessing for study 1.** Whole-brain fMRI data were acquired on a GE 1.5 T scanner using a T2\*-weighted spiral in-out sequence developed by Dr Gary Glover: repetition time (TR) = 2,000 ms, echo time (TE) = 40 ms, flip angle = 84 and field of view = 22 cm, 24 axial slices ( $3.5 \times 3.5 \times 4.5 \text{ mm}^3$  voxels) parallel to the anterior commissure-posterior commissure line. Structural data were acquired with a T1-weighted spoiled gradient-recalled sequence:  $1 \times 1 \times 1 \text{ mm}^3$ , TR = 19 ms, TE = 5 ms, flip angle = 20. Stimulus presentation and behavioural data acquisition were controlled using E-Prime software (PST Inc.).

Structural T1-weighted images were co-registered to the mean functional image for each subject using the iterative mutual information-based algorithm implemented in SPM8 and manual adjustment of the starting point until the co-registration was satisfactory. Structural images were normalized to Montreal Neurological Institute space using SPM8. Before preprocessing of functional images, we removed the first four volumes to allow for image intensity stabilization. We also identified image-intensity outliers by computing the mean and standard deviation (across voxels) of intensity values for each image for all slices to remove intermittent gradient and severe motion-related artefacts present to some degree in all fMRI data. To identify outliers, Mahalanobis distances for the matrix of slice-wise mean and standard deviation values (concatenated)  $\times$  functional volumes (time) were computed, and any values with a significant  $\chi^2$  value (corrected for multiple comparisons based on the more stringent of either FDR or Bonferroni methods) were considered outliers (less than 1% of images were outliers). Each time-point identified as outliers was included as a separate nuisance covariate in the first-level models.

Then, functional images were corrected for differences in the acquisition timing of each slice and were motion (realignment) corrected using SPM8. The functional images were warped to SPM's normative atlas using warping parameters estimated from co-registered, high-resolution structural images, interpolated to  $2 \times 2 \times 2 \text{ mm}^3$  voxels, and smoothed with an 8-mm FWHM Gaussian kernel. This smoothing level has been shown to improve inter-subject functional alignment,

while retaining sensitivity to mesoscopic activity patterns that are consistent across individuals<sup>48,66</sup>.

**fMRI acquisition and preprocessing for Study 2.** Whole-brain fMRI data were acquired on a 3T Siemens Trio system (Siemens) with a 12-channel receiver head coil using a T2\*-weighted functional EPI sequence: TR = 2,000 ms, TE = 29 ms, matrix size =  $64 \times 64 \times 33$ , flip angle = 75, slice thickness = 3.5 mm, field of view = 24 cm. The first four volumes of the resting state data were first removed to allow for image intensity stabilization, followed by detection of frame-to-frame motion and intensity outliers using the Artifact Detection Toolbox. Skull stripping was then conducted using FSL's Brain Extraction Tool (v5, FMRIB, Oxford, UK), followed by slice timing and motion correction (realignment). Functional images were smoothed with an 8-mm FWHM Gaussian kernel, and warped to the Montreal Neurological Institute template using FMRIB's Linear Image Registration Tool. Using the Conn SPM toolbox<sup>67</sup>, data were temporally filtered using a band-pass filter to retain frequencies between 0.009 and 0.08 Hz. Non-neural sources of variance were removed by regressing out the six estimated head movement parameters (*x*, *y*, *z*, roll, pitch and yaw) and their first temporal derivatives, the top five principal components from the signal averaged in the white matter and ventricles, and the frames identified as motion or intensity outliers, and the residual was retained for analyses of functional connectivity.

**Behavioural analysis.** Affect rating data (Fig. 1b) were analysed using a multi-level GLM analysis<sup>68</sup> (<http://wagerlab.colorado.edu/tools>). Linguistic Inquiry and Word Count software (LIWC<sup>33</sup>)—a word frequency-based text analysis tool—was used for text analysis (Fig. 1c). The written descriptions of participants' stream of thoughts while watching their ex-partner's photos were first filtered with the negative emotion category in the LIWC dictionary. Then, we calculated the word frequency for each negative emotion category using the sub-categories within the negative emotion category.

**First-level fMRI analysis for Study 1.** SPM8 was used to conduct first-level GLM analyses by modelling task regressors corresponding to the 15-s photo/heat period, the 5-s affect/pain rating period and the 18-s visuospatial control task period. The fixation-cross epoch was designated the unmodelled baseline. For each task regressor, a Boxcar function was convolved with SPM8's canonical haemodynamic response function. A high-pass filter of 180 s, which is well suited for pain and emotion elicited for longer duration, was then applied. Other regressors of non-interest (that is, nuisance variables) included (i) 'dummy' coding regressors for each run (intercept for each run); (ii) linear drift across time within each run; (iii) the six estimated head movement parameters (*x*, *y*, *z*, roll, pitch and yaw; mean-centered) and their squares, their derivatives and squared derivative for each run (24 columns per run); (iv) indicator vectors for outlier time points identified based on their multivariate distance from the other images in the sample (see above); (v) indicator vectors for the first two images in each run. Voxel-wise statistical parametric maps for different trial types were calculated for each participant and then entered into MVPA (Figs 2–5) or random-effects group GLM analyses (Supplementary Fig. 3). Robust regression analysis was implemented for the group-level GLM analyses<sup>69</sup>. All results were thresholded at  $q < 0.05$ , FDR corrected, two-tailed.

**Multivariate voxel pattern analysis.** We used linear SVMs<sup>34</sup> to train multivariate pattern classifiers for pain and rejection. For global pattern classifiers (Figs 2 and 3), we first conducted feature selection to maximize predictability and generalizability of SVMs. A priori voxels associated with 'pain', 'emotion' and 'social' terms were selected based on the union of forward and reverse inference maps from the automated large-scale meta-analytic database of more than 5,800 published neuroimaging studies (<http://neurosynth.org><sup>22</sup>; see Supplementary Fig. 1). The SVMs were implemented using custom Matlab code based on the Spider toolbox (<http://people.kyb.tuebingen.mpg.de/spider>). The pattern classifiers were trained on first-level contrast images for four different conditions (that is, Heat-Pain, Warmth, Ex-partner and Friend) to separately discriminate 'Heat-Pain' and 'Ex-partner' from the respective three other conditions using the one-against-all approach<sup>34</sup> (that is, Heat-pain against other conditions, including Ex-partner, and Ex-partner against other conditions, including Heat-pain). Bootstrap tests were conducted to provide *P*-values for voxel weights in order to threshold the classifier weights for display and interpretation (Fig. 2b). We constructed 10,000 bootstrap sample sets (with replacement) and ran SVMs on each. Two-tailed, uncorrected *P*-values were calculated for each voxel based on the proportion of weights above or below zero. With a leave-one-subject-out cross-validation procedure, we assessed classification accuracy of the SVM classifiers using the forced-choice test, where pattern expression values (that is, the dot-product of a vectorized activation image with the classifier weights) were compared for two conditions tested within the same out-of-sample individual, and the higher was chosen as pain or rejection. We calculated the accuracy for Heat-pain versus Ex-partner or Ex-partner versus Heat-pain, Heat-pain versus Warmth and Ex-partner versus Friend. We also conducted a multi-voxel pattern similarity analysis and a cross-classification test (Figs 4 and 5), and functional connectivity analysis with resting-state fMRI data (Study 2). For more details of these analyses, see Results.

## References

- MacDonald, G. & Leary, M. R. Why does social exclusion hurt? The relationship between social and physical pain. *Psychol. Bull.* **131**, 202–223 (2005).
- Eisenberger, N. I., Lieberman, M. D. & Williams, K. D. Does rejection hurt? An fMRI study of social exclusion. *Science* **302**, 290–292 (2003).
- Eisenberger, N. I. The pain of social disconnection: examining the shared neural underpinnings of physical and social pain. *Nat. Rev. Neurosci.* **13**, 421–434 (2012).
- Kross, E., Berman, M. G., Mischel, W., Smith, E. E. & Wager, T. D. Social rejection shares somatosensory representations with physical pain. *Proc. Natl Acad. Sci. USA* **108**, 6270–6275 (2011).
- Novembre, G., Zanon, M. & Silani, G. Empathy for social exclusion involves the sensory-discriminative component of pain: a within-subject fMRI study. *Soc. Cogn. Affect. Neurosci.* doi: 10.1093/scan/nsu038 (2014).
- Holt-Lunstad, J., Smith, T. B. & Layton, J. B. Social relationships and mortality risk: a meta-analytic review. *PLoS Med.* **7**, e1000316 (2010).
- Keller, M. C., Neale, M. C. & Kendler, K. S. Association of different adverse life events with distinct patterns of depressive symptoms. *Am. J. Psychiatry.* **164**, 1521–1529 quiz 1622 (2007).
- Cacioppo, J. T. & Hawkey, L. C. Perceived social isolation and cognition. *Trends Cogn. Sci.* **13**, 447–454 (2009).
- Barrett, L. F. & Wager, T. D. The structure of emotion—Evidence from neuroimaging studies. *Curr. Dir. Psychol. Sci.* **15**, 79–83 (2006).
- Cohen, S. & Janicki-Deverts, D. Can we improve our physical health by altering our social networks? *Perspect. Psychol. Sci.* **4**, 375–378 (2009).
- DeWall, C. N. *et al.* Acetaminophen reduces social pain: behavioral and neural evidence. *Psychol. Sci.* **21**, 931–937 (2010).
- Panksepp, J., Nelson, E. & Bekkedal, M. Brain systems for the mediation of social separation-distress and social-reward. Evolutionary antecedents and neuropeptide intermediaries. *Ann. N.Y. Acad. Sci.* **807**, 78–100 (1997).
- Hsu, D. T. *et al.* Response of the mu-opioid system to social rejection and acceptance. *Mol. Psychiatry* **18**, 1211–1217 (2013).
- Craig, A. D. How do you feel—now? The anterior insula and human awareness. *Nat. Rev. Neurosci.* **10**, 59–70 (2009).
- Etkin, A. & Wager, T. D. Functional neuroimaging of anxiety: a meta-analysis of emotional processing in PTSD, social anxiety disorder, and specific phobia. *Am. J. Psychiatry* **164**, 1476–1488 (2007).
- Rainville, P., Duncan, G. H., Price, D. D., Carrier, B. & Bushnell, M. C. Pain affect encoded in human anterior cingulate but not somatosensory cortex. *Science* **277**, 968–971 (1997).
- Peelen, M. V., Wiggett, A. J. & Downing, P. E. Patterns of fMRI activity dissociate overlapping functional brain areas that respond to biological motion. *Neuron* **49**, 815–822 (2006).
- Peelen, M. V. & Downing, P. E. Using multi-voxel pattern analysis of fMRI data to interpret overlapping functional activations. *Trends Cogn. Sci.* **11**, 4–5 (2007).
- Iannetti, G. D., Salomons, T. V., Moayed, M., Mouraux, A. & Davis, K. D. Beyond metaphor: contrasting mechanisms of social and physical pain. *Trends Cogn. Sci.* **17**, 371–378 (2013).
- Wager, T. D. *et al.* in *Handbook of Emotions* Ch. 15 (eds Lewis, M., Haviland-Jones, J. M. & Barrett, L. F.) 249–271 (Guilford, 2008).
- Dosenbach, N. U. *et al.* A core system for the implementation of task sets. *Neuron* **50**, 799–812 (2006).
- Yarkoni, T., Poldrack, R. A., Nichols, T. E., Van Essen, D. C. & Wager, T. D. Large-scale automated synthesis of human functional neuroimaging data. *Nat. Methods* **8**, 665–670 (2011).
- Egner, T. Surprise! A unifying model of dorsal anterior cingulate function? *Nat. Neurosci.* **14**, 1219–1220 (2011).
- Seeley, W. W. *et al.* Dissociable intrinsic connectivity networks for salience processing and executive control. *J. Neurosci.* **27**, 2349–2356 (2007).
- Kvitsiani, D. *et al.* Distinct behavioural and network correlates of two interneuron types in prefrontal cortex. *Nature* **498**, 363–366 (2013).
- Sikes, R. W. & Vogt, B. A. Nociceptive neurons in area 24 of rabbit cingulate cortex. *J. Neurophysiol.* **68**, 1720–1732 (1992).
- Shidara, M. & Richmond, B. J. Anterior cingulate: single neuronal signals related to degree of reward expectancy. *Science* **296**, 1709–1711 (2002).
- Sheth, S. A. *et al.* Human dorsal anterior cingulate cortex neurons mediate ongoing behavioural adaptation. *Nature* **488**, 218–221 (2012).
- Haxby, J. V. *et al.* Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science* **293**, 2425–2430 (2001).
- Sternberg, S. Separate modifiability, mental modules, and the use of pure and composite measures to reveal them. *Acta Psychol.* **106**, 147–246 (2001).
- Kamitani, Y. & Tong, F. Decoding the visual and subjective contents of the human brain. *Nat. Neurosci.* **8**, 679–685 (2005).
- Wager, T. D. *et al.* An fMRI-based neurologic signature of physical pain. *New Engl J. Med.* **368**, 1388–1397 (2013).
- Pennebaker, J. W., Chung, C. K., Ireland, M., Gonzales, A. & Booth, R. J. The development and psychometric properties of LIWC2007. *LIWC. Net* (2007).
- Vapnik, V. *The Nature of Statistical Learning Theory* (Springer-Verlag, 1995).
- Hastie, T., Tibshirani, R. & Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction* 2nd edn (Springer, 2009).
- Apkarian, A. V., Bushnell, M. C., Treede, R. D. & Zubieta, J. K. Human brain mechanisms of pain perception and regulation in health and disease. *Eur. J. Pain* **9**, 463–484 (2005).
- Bushnell, M. C., Ceko, M. & Low, L. A. Cognitive and emotional control of pain and its disruption in chronic pain. *Nat. Rev. Neurosci.* **14**, 502–511 (2013).
- Denny, B. T., Kober, H., Wager, T. D. & Ochsner, K. N. A meta-analysis of functional neuroimaging studies of self- and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. *J. Cogn. Neurosci.* **24**, 1742–1752 (2012).
- Amodio, D. M. & Frith, C. D. Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* **7**, 268–277 (2006).
- Wager, T. D., Davidson, M. L., Hughes, B. L., Lindquist, M. A. & Ochsner, K. N. Prefrontal-subcortical pathways mediating successful emotion regulation. *Neuron* **59**, 1037–1050 (2008).
- Urry, H. L., van Reekum, C. M., Johnstone, T. & Davidson, R. J. Individual differences in some (but not all) medial prefrontal regions reflect cognitive demand while regulating unpleasant emotion. *Neuroimage* **47**, 852–863 (2009).
- Willis, W. D. & Westlund, K. N. Neuroanatomy of the pain system and of the pathways that modulate pain. *J. Clin. Neurophysiol.* **14**, 2–31 (1997).
- Kriegeskorte, N., Mur, M. & Bandettini, P. Representational similarity analysis—connecting the branches of systems neuroscience. *Front. Sys. Neurosci.* **2**, 4 (2008).
- Corradi-Dell'Acqua, C., Hofstetter, C. & Vuilleumier, P. Felt and seen pain evoke the same local patterns of cortical activity in insular and cingulate cortex. *J. Neurosci.* **31**, 17996–18006 (2011).
- Kriegeskorte, N., Goebel, R. & Bandettini, P. Information-based functional brain mapping. *Proc. Natl Acad. Sci. USA* **103**, 3863–3868 (2006).
- Craddock, R. C., James, G. A., Holtzheimer, 3rd P. E., Hu, X. P. & Mayberg, H. S. A whole brain fMRI atlas generated via spatially constrained spectral clustering. *Hum. Brain Mapp.* **33**, 1914–1928 (2012).
- Eisenberger, N. I. & Cole, S. W. Social neuroscience and health: neurophysiological mechanisms linking social ties with physical health. *Nat. Neurosci.* **15**, 669–674 (2012).
- Shmuel, A., Chaimow, D., Raddatz, G., Ugurbil, K. & Yacoub, E. Mechanisms underlying decoding at 7 T: Ocular dominance columns, broad structures, and macroscopic blood vessels in V1 convey information on the stimulated eye. *Neuroimage* **49**, 1957–1964 (2010).
- Logothetis, N. K. What we can do and what we cannot do with fMRI. *Nature* **433**, 869–878 (2008).
- Chaimow, D., Yacoub, E., Ugurbil, K. & Shmuel, A. Modeling and analysis of mechanisms underlying fMRI-based decoding of information conveyed in cortical columns. *Neuroimage* **56**, 627–642 (2011).
- Fisher, H. E., Brown, L. L., Aron, A., Strong, G. & Mashek, D. Reward, addiction, and emotion regulation systems associated with rejection in love. *J. Neurophysiol.* **104**, 51–60 (2010).
- Way, B. M., Taylor, S. E. & Eisenberger, N. I. Variation in the mu-opioid receptor gene (OPRM1) is associated with dispositional and neural sensitivity to social rejection. *Proc. Natl Acad. Sci. USA* **106**, 15079–15084 (2009).
- Butler, R. K. & Finn, D. P. Stress-induced analgesia. *Prog. Neurobiol.* **88**, 184–202 (2009).
- Kehoe, P. & Blass, E. M. Opioid-mediation of separation distress in 10-day-old rats: reversal of stress with maternal stimuli. *Dev. Psychobiol.* **19**, 385–398 (1986).
- DeWall, C. N. & Baumeister, R. F. Alone but feeling no pain: Effects of social exclusion on physical pain tolerance and pain threshold, affective forecasting, and interpersonal empathy. *J. Pers. Soc. Psychol.* **91**, 1–15 (2006).
- Anderson, B. J. Paracetamol (Acetaminophen): mechanisms of action. *Paediatr. Anaesth.* **18**, 915–921 (2008).
- Kapur, S., Phillips, A. G. & Insel, T. R. Why has it taken so long for biological psychiatry to develop clinical tests and what to do about it? *Mol. Psychiatr.* **17**, 1174–1179 (2012).
- Landa, A., Peterson, B. S. & Fallon, B. A. Somatoform pain: a developmental theory and translational research review. *Psychosom. Med.* **74**, 717–727 (2012).
- Singh, I. & Rose, N. Biomarkers in psychiatry. *Nature* **460**, 202–207 (2009).
- Vogt, B. A. Pain and emotion interactions in subregions of the cingulate gyrus. *Nat. Rev. Neurosci.* **6**, 533–544 (2005).
- Davis, T. & Poldrack, R. A. Measuring neural representations with fMRI: practices and pitfalls. *Ann. N.Y. Acad. Sci.* **1296**, 108–134 (2013).
- Williams, K. D., Cheung, C. K. T. & Choi, W. Cyberostracism: effects of being ignored over the Internet. *J. Pers. Soc. Psychol.* **79**, 748–762 (2000).
- Ayduk, O., Mischel, W. & Downey, G. Attentional mechanisms linking rejection to hostile reactivity: the role of 'hot' versus 'cool' focus. *Psychol. Sci.* **13**, 443–448 (2002).



64. Kross, E., Davidson, M., Weber, J. & Ochsner, K. Coping with emotions past: the neural bases of regulating affect associated with negative autobiographical memories. *Biol. Psychiat.* **65**, 361–366 (2009).
65. Wager, T. D. *et al.* Placebo-induced changes in fMRI in the anticipation and experience of pain. *Science* **303**, 1162–1167 (2004).
66. Op de Beeck, H. P. Against hyperacuity in brain reading: spatial smoothing does not hurt multivariate fMRI analyses? *Neuroimage* **49**, 1943–1948 (2010).
67. Whitfield-Gabrieli, S. & Nieto-Castanon, A. Conn: a functional connectivity toolbox for correlated and anticorrelated brain networks. *Brain Connect.* **2**, 125–141 (2012).
68. Lindquist, K. A., Wager, T. D., Kober, H., Bliss-Moreau, E. & Barrett, L. F. The brain basis of emotion: a meta-analytic review. *Behav. Brain Sci.* **35**, 121–143 (2012).
69. Wager, T. D., Keller, M. C., Lacey, S. C. & Jonides, J. Increased sensitivity in neuroimaging analyses using robust regression. *Neuroimage* **26**, 99–113 (2005).
70. Amunts, K. *et al.* Cytoarchitectonic mapping of the human amygdala, hippocampal region and entorhinal cortex: intersubject variability and probability maps. *Anat. Embryol. (Berl.)* **210**, 343–352 (2005).

## Acknowledgements

Thanks to Gary Glover for generously providing the spiral in-out sequence (NIH P41 RR009784). This work was funded by NIMH R01MH076136 and NIDA R01DA035484-01 (T.D.W., PI) and NIBIB R01EB016061 (M.A.L., PI). Resting state

fMRI data were funded by NIMH P50 MH079485 (M.T.B., PI). C.-W.W. was supported by a Fulbright Graduate Study Fellowship, and L.K. was supported by a postdoctoral fellowship from the Swiss National Science Foundation (PBGE1\_142252). Matlab code implementing all analyses is available at <http://wagerlab.colorado.edu/tools>.

## Author contributions

C.-W.W. and T.D.W. analysed the data, interpreted the results and wrote the paper. E.K. and T.D.W. designed the experiment, collected the data and revised the paper. L.K. and L.R. analysed the data and revised the paper. M.A.L. analysed the data and revised the paper. M.T.B. and J.R.A.-H. provided the resting-state fMRI data and revised the paper.

## Additional information

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Woo, C.-W. *et al.* Separate neural representations for physical pain and social rejection. *Nat. Commun.* 5:5380 doi: 10.1038/ncomms6380 (2014).