PNAS RESEARCH ARTICLE | NEUROSCIENCE
PSYCHOLOGICAL AND COGNITIVE SCIENCES

The neural signature of the decision value of future pain

Michel-Pierre Coll[a,b], Hocine Slimani[c], Choong-Wan Woo[d,e], Tor D. Wager[f], Pierre Rainville[g,h], Étienne Vachon-Presseau[i,j,k], and Mathieu Roy[c,k,1]

Edited by Kent Berridge, University of Michigan, Ann Arbor, MI; received November 1, 2021; accepted April 4, 2022, by Editorial Board Member Michael S. Gazzaniga

Pain is a primary driver of action. We often must voluntarily accept pain to gain rewards. Conversely, we may sometimes forego potential rewards to avoid associated pain. In this study, we investigated how the brain represents the decision value of future pain. Participants ($n = 57$) performed an economic decision task, choosing to accept or reject offers combining various amounts of pain and money presented visually. Functional MRI (fMRI) was used to measure brain activity throughout the decision-making process. Using multivariate pattern analyses, we identified a distributed neural representation predicting the intensity of the potential future pain in each decision and participants' decisions to accept or avoid pain. This neural representation of the decision value of future pain included negative weights located in areas related to the valuation of rewards and positive weights in regions associated with saliency, negative affect, executive control, and goal-directed action. We further compared this representation to future monetary rewards, physical pain, and aversive pictures and found that the representation of future pain overlaps with that of aversive pictures but is distinct from experienced pain. Altogether, the findings of this study provide insights on the valuation processes of future pain and have broad potential implications for our understanding of disorders characterized by difficulties in balancing potential threats and rewards.

pain | value | MVPA | decision-making | reward

Significance

We often willingly experience pain to reach a goal. However, potential pain can also prevent reckless action. How do we consider future pain when deciding on the best course of action? To date, the precise neural mechanisms underlying the valuation of future pain remain unknown. Using functional MRI, we derive a whole-brain signature of the value of future pain capable of predicting participants' choices to accept pain in exchange for a reward. We show that this signature is characterized by a distributed pattern of activity with clear contributions from structures encoding reward and salience, notably the ventral and dorsal striatum. These findings highlight how the brain assigns value to future pain when choosing the best course of action.

The value of pain for survival resonates far beyond immediate behavior; painful experiences allow us to learn to predict and avoid future pain by ascribing an aversive value to potentially painful actions (1). However, avoiding future pain often comes at the cost of missed opportunities and foregone rewards, and excessive pain avoidance can have disastrous effects on health and well-being (2). It is therefore crucial for organisms to judiciously assess the right price to pay to avoid pain. To fulfill this function, the brain must accurately represent the intensity of potential pain and assess its value for comparison with competing goals (3). Thus, understanding how future pain is represented in the brain constitutes an essential missing piece of information linking pain predictions (1, 4) to decisions about pain.

One possibility could be that the value of future pain is represented in the form of a shared "common currency" exchangeable with other goods, such as money. Following that hypothesis, we would predict that regions known to positively encode the value of expected rewards, such as the ventral striatum [VS (5–8)] and orbitofrontal cortex [OFC (9, 10)], should negatively encode the value of future pain. In other terms, future pain would be predicted to have a signed value representation. It should be noted that this hypothesis is compatible with reports of aversive–appetitive gradients in the VS and OFC (8, 11–13), in which case opposite patterns of activation/deactivation in these regions would be expected to track the value of future pain and competing rewards.

By contrast, future pain and rewards could also share an unsigned value representation, as suggested by the fact that cues signaling pain and rewards are both motivationally salient. Indeed, the prospect of high levels of future pain or rewards may mark the need for increased cognitive control over the decision process to prevent making potentially important mistakes. Accumulating evidence now indicates that the anterior insula (aINS) tracks stimulus saliency across a wide range of domains and shows a U-shaped response to values, with an increased response to both relatively low and high values (5, 14). Moreover, recent models of the dorsal anterior cingulate cortex (dACC) function propose that it plays an important role in compiling the expected value of control, that is, how much control should be allocated to a given decision as a function of the potential costs or benefits (15, 16). Finally, other regions densely connected with the dACC, such as the dorsal striatum (DS) (7), could also serve in the implementation of pain-related decisions. Indeed, the DS has been shown to contribute to goal-directed action selection and initiation in both appetitive and aversive decision-making contexts

Author contributions: H.S., C.-W.W., T.D.W., P.R., É.V.-P., and M.R. designed research; M.-P.C., H.S., and M.R. performed research; M.-P.C., H.S., É.V.-P., and M.R. analyzed data; and M.-P.C., C.-W.W., T.D.W., P.R., É.V.-P., and M.R. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission. K.B. is a guest editor invited by the Editorial Board.

[1]To whom correspondence may be addressed. Email: mathieu.roy3@mcgill.ca.

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2119931119/-/DCSupplemental.

Published June 3, 2022.

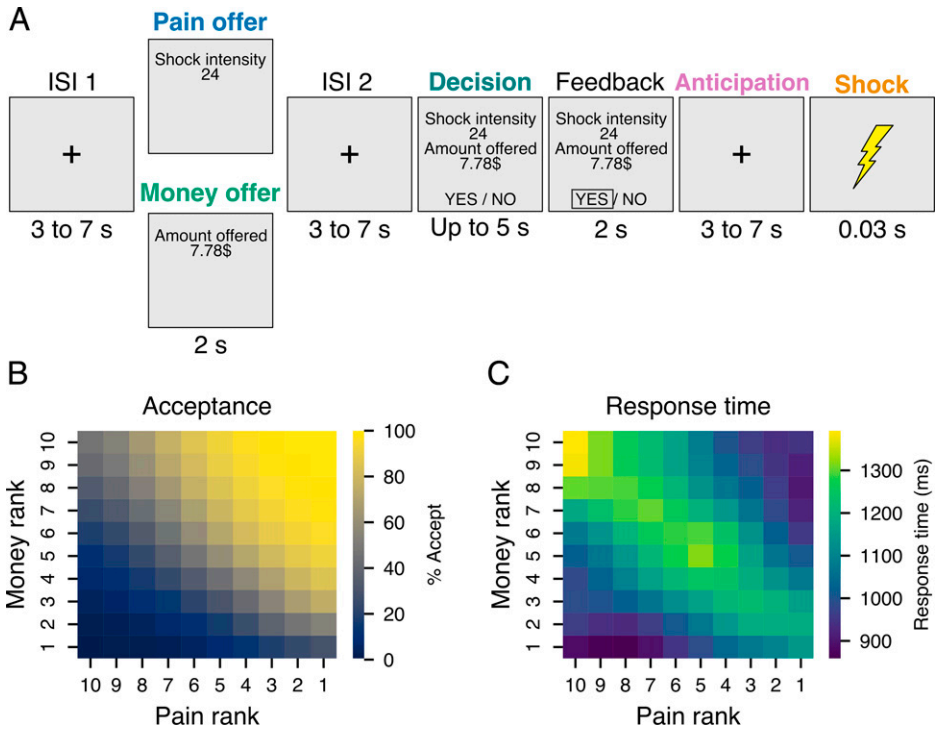

Fig. 1. Experimental task and behavioral results. (*A*) Schematic representation of an exemplar experimental trial. For each trial, participants first saw an offer screen indicating the amount of money or pain involved in the next decision, followed by an interstimulus interval (ISI) and a decision screen indicating a complementary amount of pain (or money). Participants could accept or reject the offer by pressing a key. If accepted, participants received both the pain and the opportunity to gain the money at the end of the experiment. If rejected, participants received no pain and lost the chance to earn the money. (*B*) The average proportion of offers accepted and (*C*) average response time as a function of pain and money levels offered. Matrices in *B* and *C* were smoothed with a Gaussian kernel for display only.

(17–19), suggesting that it could also harbor an unsigned value representation shared between pain and rewards.

A third possibility could be that the cerebral representation of future pain is shared with that of immediately experienced pain. Under that hypothesis, the main difference between immediate and future pain would be the degree of urgency, and we would therefore predict that future pain should be simply represented as reduced activity in pain-processing regions (e.g., ACC, insula, and somatosensory cortices), as was observed for anticipated, imagined, or recalled pain (20–23).

Finally, the decision to escape or confront future anticipated pain could also recruit the brain's defensive circuitry, including limbic structures such as the amygdala and periaqueductal gray [PAG (24, 25)]. From that perspective, the cerebral representation of future pain would be predicted to share important similarities with that of negative valence (26).

In order to test these four hypotheses (hypothesis 1, signed representation shared with rewards; hypothesis 2, unsigned representation shared with rewards; hypothesis 3, shared representation with experienced pain; and hypothesis 4, shared representation with negative valence), we asked participants to accept or reject offers combining various levels of physical pain and sums of money (Fig. 1*A*). Importantly, we presented the amount of pain and money involved in each decision in two distinct stages, allowing us to separately consider the representation of the decision value of future pain and money before their integration. This particular methodological consideration represents an important innovation of the current study compared with previous studies presenting pain and rewards simultaneously (27–30). The main conclusion from these studies is that pain exerts a discounting effect on the representation of rewards in the OFC and VS. By contrast, the present study specifically assessed the cerebral representation of pain value before it is compared to that of money when making decisions to accept or reject future pain.

More specifically, to identify the cerebral representation of future pain, we employed a multivariate approach to generate whole-brain predictive patterns for future pain and monetary outcomes and the immediate painful sensation produced by electric shocks. We also leveraged open data from a previously developed whole-brain predictive pattern of negative affect (26). Using these multivariate brain patterns, we asked whether the representation of future pain was similar to the representation of future monetary rewards, immediately experienced pain, and aversive pictures. We found that a distributed pattern of activity encodes the value of future pain—the pain value pattern (PVP)—comprising both unsigned saliency and signed valuation signals. This representation is partly related to the representation of monetary reward and negative valence but distinct from the representation of physical pain. Finally, we trained a machine-learning algorithm to predict the decision to accept or reject pain based in part on PVP pattern expression, confirming that the PVP is functionally related to decisions about pain.

Results

Behavioral Results. Fifty-seven participants (26 females, mean age = 24.91 y, SD = 5.56; see *Materials and Methods* for recruitment and exclusion details) took part in a functional MRI (fMRI) task in which they were offered, at each trial, 1 of 10 amounts of money (0 to 10 $CAD) and 1 of 10 individually calibrated levels of painful electrical shock (from pain threshold to pain tolerance level). At each trial of the task (Fig. 1*A*), participants first saw a screen presenting either the pain level or the

monetary amount involved in this trial (pain or money offer). Then, after an interstimulus interval, participants received the second part of the offer and were asked to decide within 5 s using two adjacent keys on a response box (decision phase). If they accepted the offer, they received the shock after a jittered anticipation phase and a chance to obtain the monetary reward. On the other hand, if they refused the offer, they did not receive the shock and lost the opportunity to gain the money.

On average, participants accepted 52% of the offers (SD = 17%, range: 16 to 89%). Using a binomial generalized linear mixed-model, we confirmed that their decision to accept or reject each offer was influenced by both the amount of pain (B = −0.57, SE = 0.04, $P < 0.001$) and money (B = 1.38, SE = 0.06, $P < 0.001$; Fig. 1*B*). An interaction (B = −0.07, SE = 0.01, $P < 0.001$) between pain and money level indicated that pain level had a lower influence on decisions at higher levels of money and vice versa. The presentation order of pain and money levels did not influence choice behavior (B = 0.07, SE = 0.09, $P = 0.437$). Response times to the offers were on average 1,117 ms (SD = 281 ms, range: 635 to 1,712 ms), and the effect of pain and money levels interacted with choice, indicating that participants were slower to accept higher levels of pain (B = 137.33, SE = 14.28, $P < 0.001$) and to reject higher levels of money (B = −145.58, SE = 15.50, $P < 0.001$). Participants were also faster to accept offers when the monetary amount was presented first (presentation order × choice interaction: B = −561.46, SE= 139.76, $P < 0.001$). To assess the influence of choice difficulty on behavior, we computed an index of choice difficulty using the following equation: 10 − |money level − pain level|. A value of 10, therefore, represents the most difficult choices in which levels of pain and money were of equal rank, and a value of 1 indicates easy choices in which the level of pain or money maximally exceeded the other level. Response times were positively related to choice difficulty (B = 59.11, SE = 2.94, $P < 0.001$) with slower responses at the point of pain–money equivalence, especially when both pain and money offers were at their maximum levels (Fig. 1*C*). Detailed regression tables for all models and additional figures for behavioral results as a function of presentation order are shown in *SI Appendix, Supplementary Results*, S1. Altogether, these results are in line with previous investigations using a similar task (31) and indicate that the aversive value of the pain stimulation led participants to refuse opportunities to gain significant amounts of money and that the similarity between pain and money offers and their high values increased decision times.

Multivariate Pattern Predicting the Decision Value of Future Pain. We used brain activity during the pain offer phase of the task (Fig. 1*A*), during which the level of potential future pain was processed in isolation, to develop a multivariate model capable of predicting the pain offer level (PVP). We trained a least absolute shrinkage and selection operator principal component regression (LASSO-PCR) algorithm to predict the intensity of the pain in the offers using the whole-brain parametric maps for each level and each participant as input features. The resulting pattern of voxel weights was considered as the PVP. We assessed the performance of each pattern in predicting the level of hypothetical future pain or money with 10-fold cross-validation in which the same LASSO-PCR was iteratively trained on the images from ∼90% of participants and tested on the images of the remaining 10% of participants (*Materials and Methods*).

The analysis indicated that across all participants and images, the PVP was able to predict the pain offer level (cross-validated $r = 0.44$, root mean squared error [RMSE] = 2.60 on a 10 point scale, $R^2 = 0.18$) with a prediction accuracy significantly above chance accuracy (permutation test: $P < 0.001$). The slope between the actual and predicted values for each participant and the distribution of the related correlations are shown in Fig. 2*A*. Visual inspection of the average pattern similarity for each level of future pain (Fig. 2*B*) indicated that the PVP did not discriminate between very low levels of pain offers but could discriminate between higher levels. We confirmed the capacity of the pattern to discriminate between low, medium, and high levels of pain using single interval classification tests between levels 1 and 5, 1 and 10, and 5 and 10 (Fig. 2*C*), which indicated significantly above chance discrimination for all binary choices (binomial tests, all $P < 0.05$ Bonferonni corrected for the three tests). We further tested the PVP predictive performance in an independent open dataset [$n = 28$; openneuro.org/datasets/ds001814 (32)], which used a similar design, with the difference that the future pain involved in each decision was cued with a gauge indicating one of three levels of risk of receiving a shock. We calculated the cosine similarity between the PVP weights and the parametric maps of each participant, trial, and risk level. As shown in Fig. 2*C*, the PVP similarity was significantly related to the risk of pain involved in the decisions (B = 0.07, SE= 0.03, $P = 0.02$), suggesting that the PVP could track the decision value of future pain levels in both our dataset and in independent data. It could be the case, however, that the PVP mainly tracks the level of the offer without having a direct impact on decision. For instance, someone who does not value pain at all (always chooses money) could still represent pain levels. To address this issue, we additionally show that the PVP similarity is related to individual differences in pain valuation and to participant-specific estimates of pain value based on computational modeling of choice behavior (*SI Appendix, Supplementary Results*, S2 and S3). By directly relating PVP similarity to subject-specific estimates of the influence of pain value on choice behavior, these results provide additional evidence for the idea that the PVP tracks the subjective value of pain and not simply input to the decision process.

To identify which voxels made the most stable contribution to the PVP, we used a bootstrap resampling procedure to construct 10,000 bootstrap samples of the same size as the original dataset. We trained the LASSO-PCR algorithms in each of these samples, and we used the resulting bootstrap distribution for each voxel weight to threshold the pattern maps using the false discovery rate procedure [FDR (33)] with a threshold of $q < 0.05$, two-tailed. Note that the patterns were thresholded for display and interpretation only and that the full pattern was used in all analyses.

As shown in Fig. 2*E*, the PVP was characterized by a distributed pattern of weights with positive weights mainly located in the left primary somatosensory and motor cortices, the dACC, the PAG, the DS and the precuneus and negative weights in the bilateral VS, the bilateral mid/lateral OFC, right paracentral lobule, and the bilateral superior temporal sulcus. To further characterize the nature of the PVP, we assessed the distribution of the unthresholded pattern weights across large-scale cortical (34) and striatal (35) resting-state networks by calculating the Pearson correlation between each network relative to the other networks and the PVP weights. As shown in Fig. 2*F*, at the cortical level, the PVP was positively associated with the ventral attention and default mode networks at the cortical level and showed a negative association with the limbic and somatomotor networks. Given the importance of the striatum weights in the PVP, we performed the same analysis performed within the striatal networks and observed an apparent dissociation between the ventral and dorsomedial striatum, with a positive relationship between the

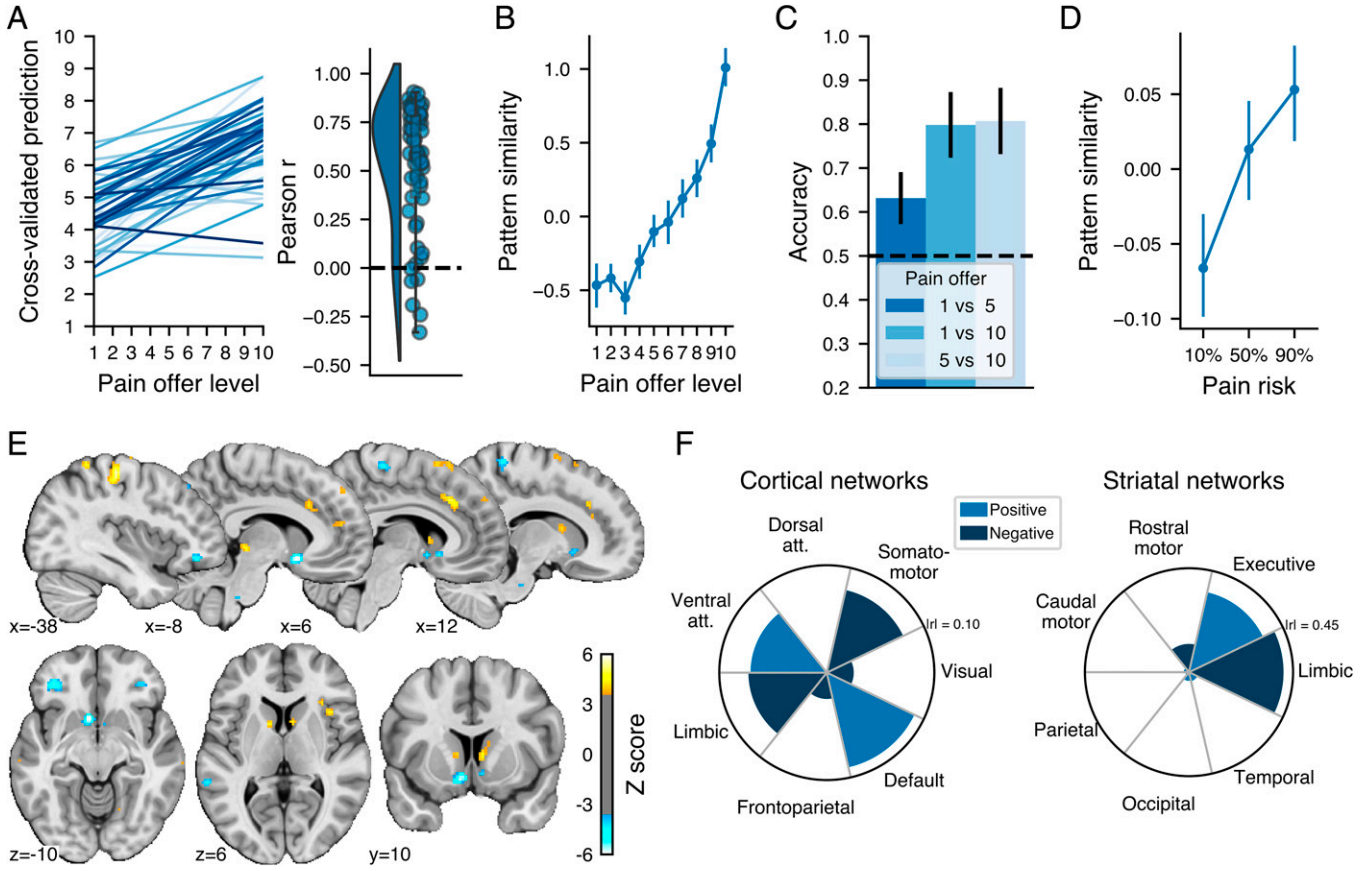

Fig. 2. Validation and performance of the multivariate pattern predicting pain offer levels and its spatial distribution. (*A*) Relationship between the actual and predicted pain offer level for each participant (shown in different shades of blue) and the corresponding distribution of the Pearson correlation between actual and predicted levels for each participant. (*B*) Average z-scored pattern similarity between the multivariate pattern and the parametric maps corresponding to each pain offer level. (*C*) Binary classification accuracy for different combinations of pain offer levels using the pain pattern similarity. The black dashed horizontal line shows the chance accuracy level. (*D*) Average z-scored pattern similarity between the multivariate pattern and the parametric maps corresponding to the level of pain risk in an independent dataset. (*E*) Multivariate pattern predicting the pain offer level thresholded at FDR q < 0.05 using a bootstrap distribution built from 10,000 samples drawn with replacement. The color bar shows the regression weights z-scored using the bootstrap distribution. Note that the multivariate patterns were thresholded for display and interpretation only. All error bars show the SEM. The full unthresholded weight map is available at https://neurovault.org/collections/10410. (*F*) Correlation between the pattern weights and resting-state (*Left*) cortical and (*Right*) striatal networks.

pattern and striatal regions primarily connected to the executive network and a negative relationship between the pattern and striatal regions related principally to the limbic network. In *SI Appendix, Supplementary Results*, S4, we show that the whole-brain PVP was better than any particular subregion in predicting pain level, therefore confirming the distributed nature of the neural representation of the decision value of future pain.

Multivariate Pattern Predicting Monetary Value and Shock Intensity. To test the hypothesis that the cerebral representation of future pain value shares characteristics with the pattern of brain activity associated with monetary rewards and immediately experienced pain, we used the same approach to developed two additional patterns to predict the level of monetary offers in the money offer phase of the task (money value pattern [MVP]) and the intensity of the delivered electrical shocks in the shock phase (shock intensity pattern [SIP]).

For the money offers, the MVP showed a good performance in predicting the amount of money involved (Fig. 3*A*; cross-validated $r = 0.56$, RMSE = 2.38, $R^2 = 0.31$, $P < 0.001$ permutation test). There was a linear relationship between the MVP similarity score and each level of money offer (Fig. 3*B*), and these pattern similarity values could be used to significantly distinguish between different levels of money offers using single interval classification tests (Fig. 3*C*; binomial tests $P < 0.05$, Bonferonni corrected for the three tests). The MVP was mainly characterized by a large cluster of positive weights covering the striatal region (Fig. 3*D*). Additional small positive clusters were located in the bilateral thalamus, right precuneus, right superior frontal gyrus, and bilateral primary somatosensory cortex, and small negative clusters were located in the right midinsula, the lingual gyrus, and the right supplementary motor area. However, these additional clusters did not meaningfully contribute to the prediction of the money level over the striatal clusters (*SI Appendix, Supplementary Results*, S4).

For the shock intensity, the SIP was highly accurate in predicting the intensity of the shocks received (Fig. 3 *E* and *F*; $r = 0.70$, $R^2 = 0.48$, RMSE= 1.96, $P < 0.001$ permutation test) and could accurately discern between shocks of various intensities (Fig. 3*G*). The SIP response to shock intensities was similar to other well-established patterns reflecting pain experience: the neurological pain signature [NPS (36)] and the stimulus-intensity independent pain signature [SIIPS (37)] (*SI Appendix, Supplementary Results*, S5). The weight map thresholded at q < 0.05 FDR corrected (Fig. 3*H*) indicated that the SIP was highly distributed across the brain but mainly characterized by large positive clusters in the bilateral aINS, the cerebellum, the PAG, and sensorimotor areas and negative clusters in the bilateral caudate nucleus, precuneus, and the dorsolateral prefrontal cortex.

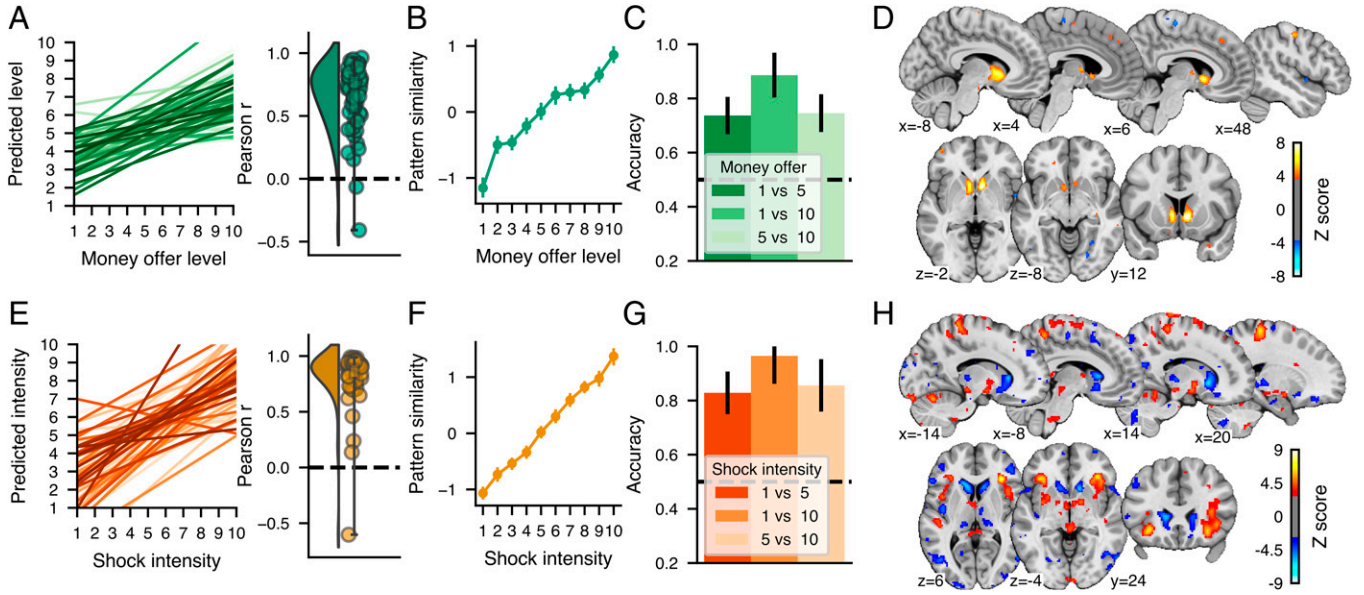

Fig. 3. Validation and performance of the multivariate pattern predicting (*A*–*D*) money offer levels (MVP) and (*E*–*H*) shock intensity (SIP) and their spatial distribution. *A* and *E* show the relationship between the actual and predicted money offer level or shock intensity for each participant (shown in different shades of green or orange) and the corresponding distribution of the Pearson correlation between actual and predicted levels for each participant. *B* and *F* show the average pattern similarity between the multivariate pattern and the parametric maps corresponding to each money offer level or shock intensity. *C* and *G* show binary classification accuracy for different pairs of money offer levels using the MVP similarity and the combinations of pairs of shock intensity using the SIP similarity. The black dashed horizontal line shows the chance accuracy level. *D* and *H* show multivariate patterns predicting the money offer level and the shock intensity thresholded at FDR q < 0.05 using a bootstrap distribution built from 10,000 samples drawn with replacement. The color bars show the regression weights z-scored using the bootstrap distribution. Note that the multivariate patterns were thresholded for display and interpretation only. The full unthresholded weight maps are available at https://neurovault.org/collections/10410.

Cross-Prediction of Pain and Monetary Values. We first assessed the separate modifiability of the PVP and MVP by testing their capacity to predict the other kind of offer (MVP predicting pain offer levels and PVP predicting money offer levels). As shown in Fig. 4*A*, the MVP similarity did not discriminate between the different levels of pain offered, suggesting that MVP is different from pattern associated with future pain. However, the PVP similarity with the different levels of money offers showed a U-shaped relationship indicating that the PVP negatively tracked monetary offers in the lower range and positively tracked monetary offers in the higher range (Fig. 4*B*). Accordingly, single interval classification indicated that the PVP similarity could discriminate between low and average money levels ($P < 0.001$, binomial test, Bonferonni corrected for the three tests) and high and average money levels ($P < 0.001$) but not between the two extreme levels ($P = 0.9$).

Since whole-brain analyses revealed a partial overlap between the representation of pain and money values, we sought to identify which brain regions harbored signed and unsigned shared value representations. To this end, we used the searchlight technique (38) to perform cross-validated predictions of the pain and money levels in 6-mm spheres centered around each voxel. We report the results corrected at $P < 0.05$, corrected for the family-wise error rate (FWE). As shown in Fig. 4*C*, significant positive cross-prediction values were found in the DS and in a large frontal cluster covering the ACC, dACC, dorsomedial prefrontal cortex (DMPFC) and extending to the supplementary motor area, indicating that these regions harbor a shared unsigned value representation between future pain and money. By contrast, negative cross-prediction was observed only in a few voxels in the vicinity of left VS, indicating that this region encodes a shared value representation between pain and money. Further visual inspection of the unthresholded weight distribution of the two patterns in the striatum (Fig. 4 *D* and *E*) indicates that the PVP and MVP show a similar pattern of positive weights in the DS but that the signed weights in the PVP were in the most ventral section of the VS and showed little overlap with the MVP weights in the more dorsal part of the VS.

Cross-Prediction of Pain Value and Physical Pain. To test the hypothesis that the cerebral representation of future pain value shares characteristics with the pattern of brain activity associated with immediately experienced pain, we examined PVP pattern expression in response to the intensity of painful electric shocks. We found no clear relationship between the PVP and experienced pain (Fig. 5*A*) or anticipated pain (*SI Appendix*, *Supplementary Results*, S6). Similarly, the signature trained on physical pain intensity showed no clear association with changes in pain offers (Fig. 5*B*). Moreover, the same was found for the NPS and the SIIPS (*SI Appendix*, *Supplementary Results*, S5), confirming that hypothetical and actual pain are separately modifiable. However, the analysis of local patterns using a searchlight analysis indicated that cross-prediction between pain values and physical pain intensity was possible in an OFC region bordering the aINS, while the bilateral DS showed negative cross-prediction (Fig. 5*C*), suggesting a common role for these regions in processing hypothetical and actual pain.

Cross-Prediction of Pain Value and Negative Affect. To test the possibility that the PVP reflects a nonspecific response to aversive visual stimuli, we additionally used an independent open dataset to assess the similarity of the pattern with brain responses to different levels of aversive pictures rated by the participants on a scale of 1 to 5, with 1 being neutral and 5 highly aversive [$n = 182$ (26)]. As shown in Fig. 6*A*, we found that the PVP could discriminate between neutral and aversive pictures (1 vs. 5, $P < 0.001$; 1 vs. 3, $P < 0.001$) but not between different levels of unpleasant pictures (3 vs. 5, $P = 0.9$). Additionally, a multivariate pattern trained on these data, the picture induced negative emotion signature (PINES) (26), responded to high levels of pain

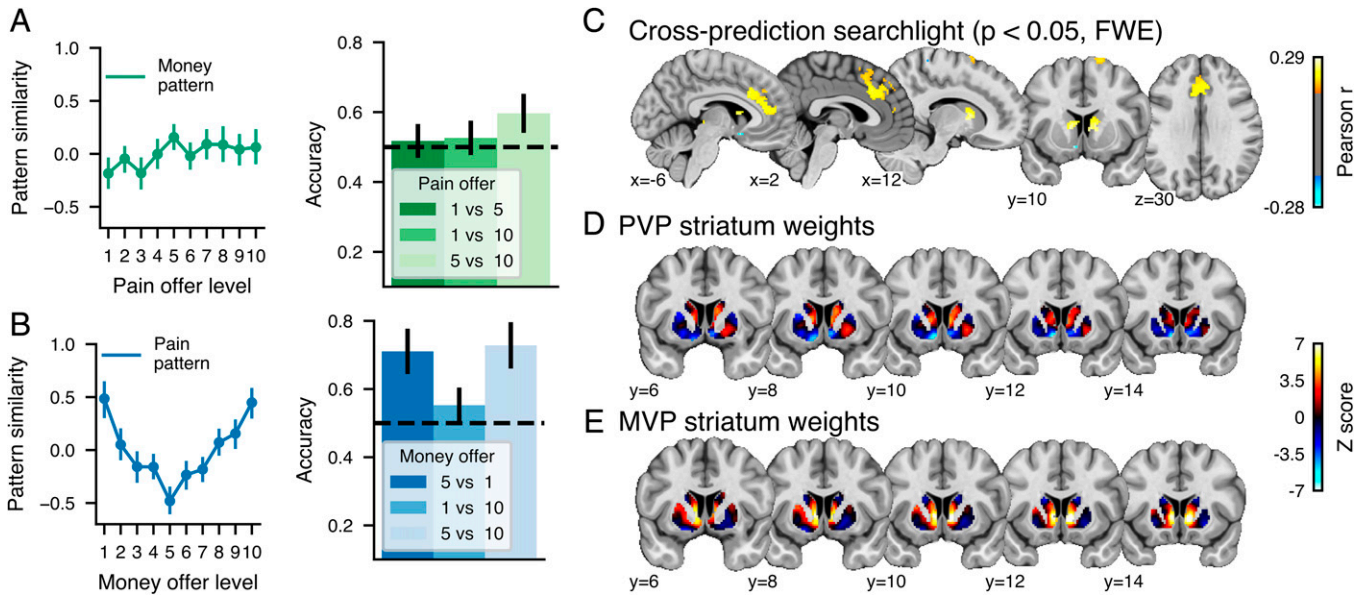

Fig. 4. Cross-prediction of pain and monetary offers. Pattern similarity and discrimination accuracy between different levels for the (*A*) MVP applied to each level of the pain offers and (*B*) the PVP applied to each level of the money offers. (*C*) Results for the whole-brain cross-prediction searchlight analysis in which a principal component regression was trained on pain or money offers and tested on the other modality ($P < 0.05$, FWE corrected using a bootstrap distribution built from 5,000 samples drawn with replacement). Unthresholded striatum z-scored weights for the (*D*) PVP and the (*E*) MVP.

offers (Fig. 6*B*) and could accordingly discriminate between high and low pain offers (1 vs. 10, $P = 0.009$) and high and medium pain offers (5 vs. 10, $P = 0.03$) but not between low and medium offers (1 vs. 5, $P = 1$). A cross-prediction searchlight between the pain offers and the emotional ratings indicated that shared patterns between the two experiences were present in the bilateral aINS, the dACC/DMPFC, and the PAG. Altogether, these results show that the PVP is sensitive to emotional arousal and can distinguish between neutral and aversive pictures but does not share the same fine-grained patterns as the PINES allowing the discrimination between different aversive pictures of different intensities.

Predicting Participants' Choices. Finally, to demonstrate that the PVP was functionally related to the decision to accept or reject the offers, we predicted participants' choices based on PVP and MVP expression during the decision period. To this end, we estimated the pattern expression of the PVP and MVP for each single-trial parametric map of the decision phase for each participant. As shown in Fig. 7 *A*–*C* the average PVP and MVP expressions during the offer phase tracked the pain and money level, respectively, and their differences closely resemble the decision heat map presented in Fig. 1*B*. We then used these two pattern expression values as features in a linear support vector machine classifier to predict participants' choices within a 10-fold cross-validation procedure. To avoid using the same data to train the predictive pattern and predict choices, we performed our analyses within the same cross-validation loop used to train the patterns, meaning that in each cross-validation fold, the binary classifier was tested in a subset of participants not used to train the pattern or the classifier. The classifier trained on the PVP and MVP similarity was able to predict participants' choices with a cross-validated balanced accuracy of 0.67 ($P < 0.001$, binomial test; Fig. 7*D*). We also confirmed that the presentation order of the pain and money values did not influence the classification of the decision as similar accuracies were obtained independently of the order of presentation (*SI Appendix, Supplementary Results*, S7). As expected, the MVP was weighted positively and predicted acceptance of the offer, while the PVP similarity was weighted negatively in the classifier and predicted rejection. To test whether the signature trained at the group level could also make accurate predictions within participants, we applied the same procedure within participants within a 10-fold cross-validation performed loop across trials. The classifier was similarly accurate when trained and tested within participants Fig. 7*E*), with an average accuracy of 0.68 [SD = 0.10, range: 0.47 to 0.88, one-sample t test against chance accuracy: $t(56) = 11.06$, $P < 001$]. Finally, we examined the relative importance of the PVP and MVP to the decision process by assessing their unique and combined predictive accuracies. Results revealed that the MVP had a higher predictive accuracy [0.61, SD = 0.10, range: 0.47 to 0.87, $t(56) = 7.78$, $P < 0.001$] than pain [0.56, SD = 0.08, range: 0.47 to 0.8, $t(56) = 5.48$, $P < 0.001$] but that their combination outperformed the use of the MVP alone [MVP vs. both patterns, $t(56) = 4.82$, $P < 0.001$], suggesting that the PVP contributes unique information not found in the MVP. The same analyses performed using the SIP pattern similarity indicated that the pattern predicting shock intensity could not predict participant's choices (balanced accuracy = 0.50, $P = 0.46$). We also performed the same analyses using the univariate maps of the effect sizes for the parametric effect of pain and money level, which indicated that the univariate map of the effect of pain level could not discriminate between choices across participants (balanced accuracy = 0.49, $P = 0.06$) but the univariate map of the effect of money level could predict choices but with a lower accuracy than the multivariate money pattern (balanced accuracy = 0.60, $P < 0.001$).

Discussion

The capacity to make judicious decisions about future potential pain is crucial for our survival and well-being. In this study, participants were asked to accept or reject painful electric shocks in exchange for monetary rewards. Importantly, pain and money offers were presented sequentially so as to allow assessing the cerebral representation of pain separately from that of rewards. More specifically, we developed a multivariate fMRI signature, the PVP, that tracked future anticipated pain levels and predicted the decision to accept or reject pain. The PVP was characterized by

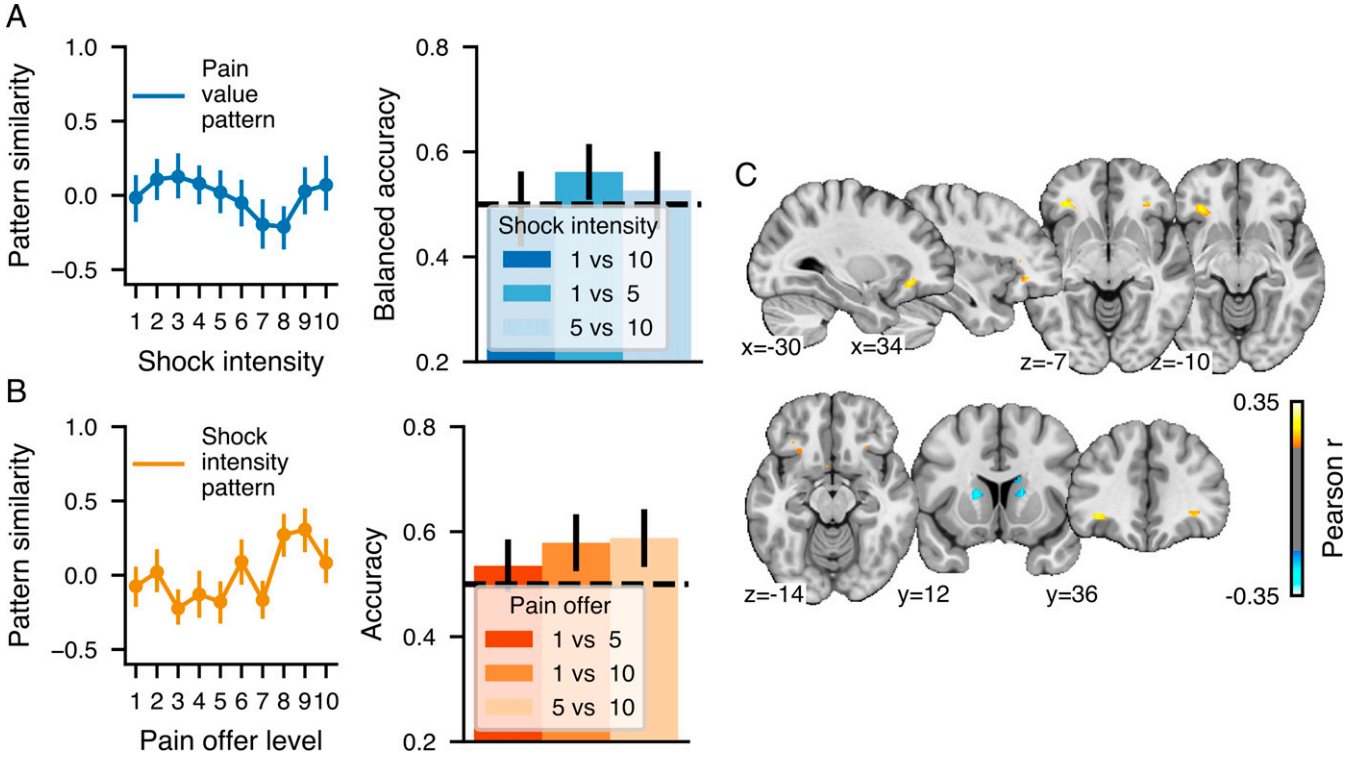

Fig. 5. Cross-prediction of pain offers and physical pain. Pattern similarity and discrimination accuracy between different levels for the (*A*) PVP applied to each level of the shock intensities and (*B*) the SIP applied to each level of the pain offers. (*C*) The whole-brain cross-prediction searchlight analysis in which a principal component regression was trained on pain offers or shock intensities and tested on the other modality ($P < 0.05$, FWE corrected using a bootstrap distribution built from 5,000 samples drawn with replacement).

a distributed pattern of voxel weights across several different regions. It included negative weights located in areas related to the valuation of rewards [VS and OFC (3, 5–10)] and positive weights in regions associated with saliency [aINS (5, 14)], executive control [dACC (15, 16)], and goal-directed action [DS and PAG (17–19, 39)]. The diversity of regions associated with the PVP may be reflective of the multidimensional nature of pain and is in line with previous brain imaging studies showing that perceived pain is associated with brain activity distributed over several brain regions (40). By contrast, the multivariate signature for monetary outcomes, the MVP, was predominantly composed of positive voxel weights in the VS and DS (*SI Appendix*, *Supplementary Results*, S4).

When trying to assess the separate modifiability (41) of the PVP and MVP, we found an asymmetry between the cross-predictive properties of the two predictive patterns. While the MVP could be dissociated from future pain, the PVP presented a U-shaped relationship with the value of monetary outcomes, with an increased pattern similarity to both extremely high and low monetary offers. This intriguing feature of the PVP can be explained by the fact that the PVP comprises a mixture of positive and negative voxel weights, while the brain's response to monetary rewards appears to be mainly characterized by a graded striatal response flipping from negative to positive close to the average level of monetary offers (level 5; Fig. 3*B*). Thus, the higher PVP response to low (e.g., level 1) and high (e.g., level 10) monetary values could be driven by a positive match between the graded negative-to-positive striatal response to monetary outcomes and the PVP's negative and positive weights, respectively. The fact that positive and negative voxel weights do not systematically cancel each other out to produce a straight line suggests that the negative and positive PVP weights tracking money level may lose their sensitivity above and below the average monetary offer. This would cause the PVP to positively track the absolute deviation of monetary offers from the average offer, with negative PVP weights driving higher similarity scores for below-average monetary offers and positive weights driving higher similarity scores for above-average monetary offers.

A searchlight examination of the local similarities between pain and money patterns revealed the presence of unsigned value signals in the DS and dACC (Figs. 2*E*, 3*D*, and 4*C*). Thus, the observation of increased activity as a function of higher pain or money levels in the DS and dACC could reflect the role of these regions in action selection and action monitoring (15–17). In addition, we found some evidence for a signed value representation in the left VS that was positively associated with money and negatively associated with pain. The reversed sign was clearly observable when comparing the PVP and MVP patterns within the striatum (Fig. 4), although the region negatively associated with pain was slightly more ventral to the one positively associated with money. This relatively small spatial shift may be consistent with the presence of an aversive–appetitive gradient in the VS (11, 13) and with the remaining uncertainty in the literature regarding the role of the VS in the representation of the negative value of aversive stimuli (3, 6). Still, the searchlight procedure confirmed the presence of a small, but statistically significant, cluster of voxels showing negative cross-prediction values between pain and money. This confirms the presence of a signed value representation in the VS, which could serve to compare pain with money for the type of decision participants were asked to make.

Therefore, there appears to be a substantial portion of the cerebral representation of future pain that is shared with that of monetary rewards. Moreover, part of this shared representation seems to be signed (hypothesis 1 above), as observed in the VS, and part of this shared representation appears to be unsigned

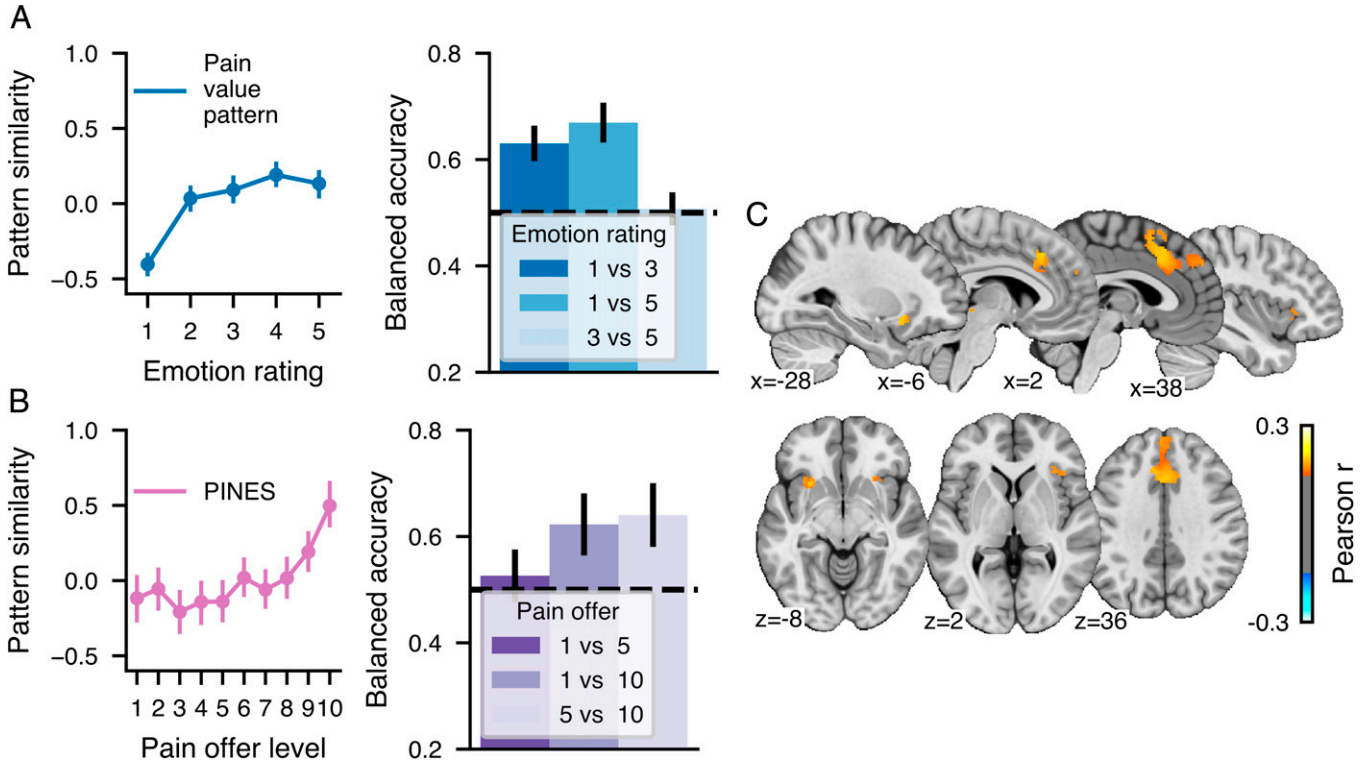

Fig. 6. Cross-prediction of pain offers and intensity of aversive pictures. Pattern similarity and discrimination accuracy between different levels for the (*A*) PVP applied to each level of affective ratings and (*B*) the PINES applied to each level of the pain offers. (*C*) The whole-brain cross-prediction searchlight analysis in which a principal component regression was trained on pain offers or shock intensities and tested on the other modality ($P < 0.05$, FWE corrected using a bootstrap distribution built from 5,000 samples drawn with replacement).

(hypothesis 2 above), such as in the DS and dACC. In addition, these findings may also have important consequences for our understanding of the cerebral representation of rewards, in that they confirm that VS and DS activations to monetary rewards may represent signed and unsigned value representations, respectively. This had been previously suggested by prior studies comparing monetary gains and losses (3), but the interpretation of these findings was complicated by uncertainty regarding how monetary gains and losses are framed in an experimental context. Therefore, comparing monetary rewards with pain provides even more convincing evidence that the VS and DS harbor signed and unsigned value representations, respectively, to monetary rewards.

Interestingly, whole-brain predictive patterns for future and immediate pain were found to be separately modifiable from one another, indicating that future pain may not be represented in the same fashion as immediate pain. Indeed, the PVP did not respond to variations in the intensity of nociceptive stimulations, and the SIP developed to predict pain intensity in the present dataset, as well as two additional well-established pain signatures [NPS (36) and SIIPS (37)], did not respond to variations in the value of future pain. While these results contrast with prior studies showing a considerable overlap between the anticipation and immediate experience of pain in the ACC and insula (21, 23), they are consistent with previous findings showing that the multivariate representation of physical pain was insensitive to the anticipation of pain or other aversive experiences (26, 36). One explanation for the lack of correspondence between the whole-brain patterns of activity associated with future and immediate pain is that they may drastically differ along some of their most important components. For instance, the PVP is almost completely devoid of activity in the somatosensory cortices. Finally, another reason for the lack of correspondence could be that even in regions where there seems to be an overlap, like in the ACC, for instance, there may still be important differences in the fine-grained pattern of activity associated with immediate and future pain.

The searchlight analysis conducted on immediate vs. future pain revealed that the only regions showing significant cross-prediction were the DS and bilateral aINS extending into the OFC. The negative cross-prediction in the DS was characterized by positive weights for future pain and negative weights for immediate pain. Given the known role of the DS in action selection (17), this inverse relationship may reflect the fact that participants had to make decisions about future pain but then passively experienced painful electric shocks as a function of their decision. This would suggest that the significant cross-prediction observed in the DS may reflect task design rather than the immediately experienced vs. hypothetical nature of the pain stimulus. By contrast, the positive cross-prediction in the bilateral aINS and OFC was associated with positive weights for both immediate and future pain and therefore truly seems to represent a region of shared representation between immediate and future pain. Indeed, brain imaging studies have frequently shown that this region is activated by pain (42). Moreover, the same region was associated with relative pain level in a previous study examining the decision value of pain (30). Based on the more general role of this structure in the processing of different types of punishers (43), it can be inferred that activity in this region reflects the general aversiveness of immediate and future pain. Thus, the shared representation between immediate and future pain in this region seems to reflect the shared unpleasantness between the two experiences, rather than a shared representation of pain's sensory dimension (hypotheses 3 and 4 above).

This conclusion seems to be confirmed by the presence of positive cross-predictions between future pain and unpleasant

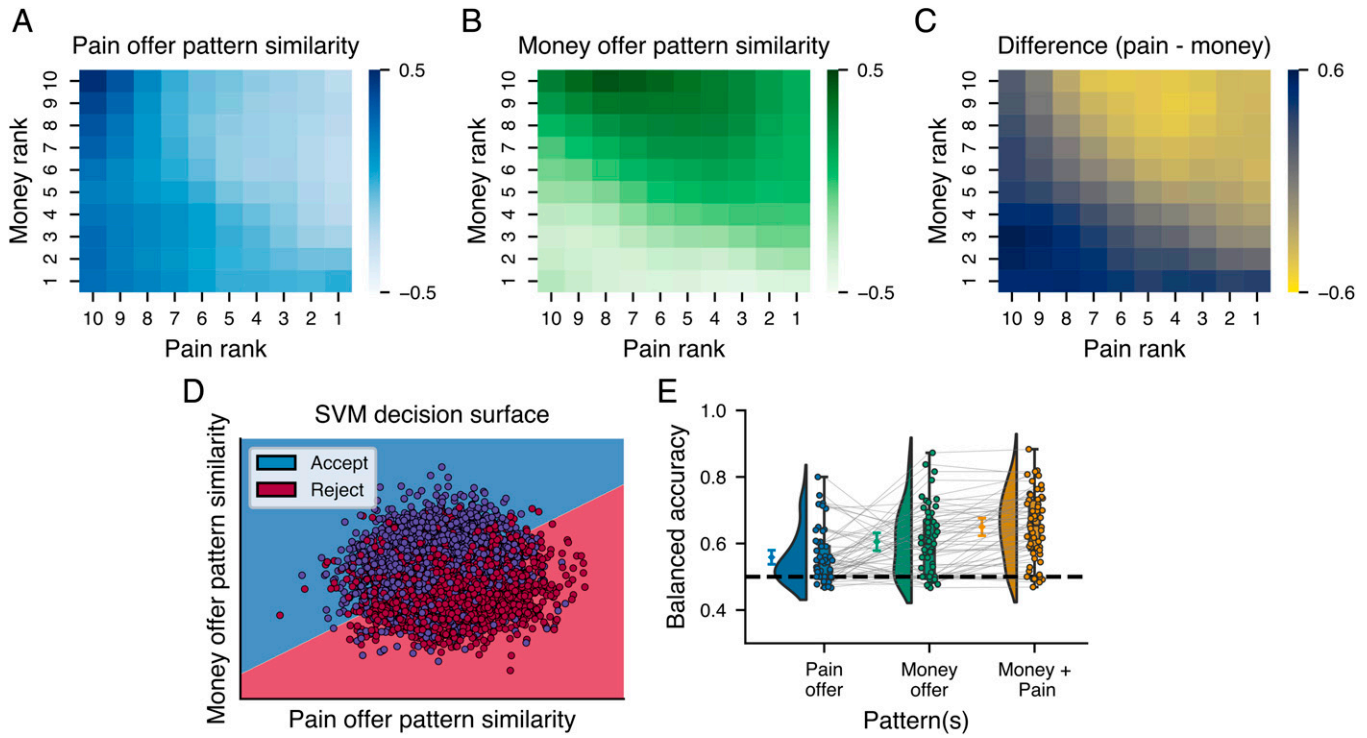

Fig. 7. Prediction of participants' choices using pain and money offer patterns. (*A*–*C*) Pattern expression during the decision phase of the pain (*A*) and money patterns (*B*) developed in the first part of the offers as well as their difference (*C*). Matrices were smoothed with a Gaussian kernel for display only. (*D*) Decision surface of the SVM classifier trained on the pattern expression of the pain and money patterns during the decision. (*E*) Classification accuracies of SVM classifiers predicting participants' decisions based on the expression of the pain and money patterns or the combination of both. Boxes show the quartiles of the dataset, while the whiskers extend to show the range of the distribution. The diamonds show the mean accuracy across all trials, and the raincloud plots show the distribution of accuracies across participants. Error bars show the 95% confidence interval.

pictures in a similar bilateral aINS cluster located slightly more caudally from the one exhibiting a shared representation between future and immediate pain (Fig. 6*C*). Altogether, this suggests that the aINS may harbor a representation of the negative affect shared between aversive pictures and immediate and future pain. The fact that there were no significant cross-predictions with monetary rewards in the aINS may help rule out the possibility that this aINS region would encode general arousal, although we should note that it is difficult to conclude from the absence of an effect with money. Still, this interpretation is consistent with previous studies on economic decision-making showing that the aINS is particularly sensitive to the prospect of monetary losses (3).

Moreover, the ACC and PAG also presented significant cross-predictions between future pain and aversive pictures (Fig. 6*C*). The PAG plays an important role in coordinating defensive behavior (25) and learning about threats (4) and could therefore be associated with activation of avoidance circuits in response to aversive pictures or threat of pain. By contrast, the ACC seems to have a more general role to play in processing saliency and the need for cognitive control (15, 16), and it was shown here to harbor shared representations between future pain and rewards, immediate pain, and aversive pictures (44). At the whole-brain level, this prevented dissociating the PVP from the previously developed PINES (26), suggesting that the cerebral representation of future pain shares important similarities with that of negative emotions.

This finding is consistent with a recent model of value-based decision-making proposing that the decision process begins with a representation of the affective state associated with the different features of the decision at hand, such as basic approach or withdrawal tendencies, and that there are at least two affective systems which activity precede and predict choices (3, 45). The observed association between the PVP and the PINES seems to indicate that an important portion of the cerebral representation of future pain may be shared with that of negative affective valence (hypothesis 4 above). The fact that many of the regions showing shared representation between future pain and negative emotions, such as the PAG or aINS, do not track monetary rewards seems to support the notion that approach and avoidance tendencies may be distinctly represented when making decisions about pain and rewards.

Finally, one last question pertains to the use of electric shocks to induce pain. While the cerebral representation of pain induced by electric shocks appears to be broadly similar to that induced by thermal stimuli (46), it seems that electrical pain is perceived as more unpleasant than thermal pain for the same levels of subjective pain intensity (47). Given that the choice to reject pain may mostly depend on its degree of unpleasantness, it is possible that anticipating future electric shocks may have been associated with a stronger negative affect, and more rejections, than if we had employed thermal stimuli. Still, the calibration procedure did not ask participants to distinguish pain sensation vs. affect, and therefore, participants' ratings represented a combination of both. Consequently, it is very likely that the same calibration procedure with thermal pain would be associated with only small differences in the ratio between pain sensation and affect.

In summary, the present study identified a pattern of brain activity, the PVP, that is predictive of the level of pain one would have to pay in order to obtain rewards and of the decision to accept or reject pain. We assessed the cerebral representation of pain-predictive cues in a decision-making context (but see ref. 30) and therefore provide an important missing piece of information linking pain predictions to decisions about pain.

This predictive pattern was found to be partly dissociable from that of monetary rewards. However, a signed shared value representation between pain and rewards was found in the VS, and unsigned shared value representations were found in the DS and ACC. Moreover, we found that the value of future pain did not share representation with immediately experienced pain but shared resemblance with the representation of negative emotions. Altogether, the present findings significantly contribute to further our understanding of the cerebral mechanisms responsible for making decisions about future potential pain, which could have potentially important implications for our understanding of disorders characterized by excessive or insufficient harm avoidance.

Materials and Methods

Participants. Sixty-six healthy adults took part in this experiment. Due to technical issues or excessive movement in the scanner, the data from six participants could not be used. Three further participants were excluded from further analyses since they either accepted or rejected almost all offers (>90%). The final sample included 57 participants (31 identified as men and 26 as women, mean age = 24.91 y, SD = 5.56). All procedures were approved by the ethics committee of the Centre de Recherche de l'Institut Universitaire de Gériatrie de Montréal (project no. CER VN 16-17-12), where the study was performed. Participants provided written informed consent at the start of the experimental session and received a fixed honorarium for their participation in addition to their gains during the decision task (*Decision Task*).

Electrical Stimulation. Electrical stimulations were administered using a Grass S48 square pulse generator (Grass Instruments) and Digitimer DS7A constant current stimulator (Digitimer) via two Ag/Cl radiotranslucent electrodes (1 cm^2 with a distance of 2 cm between electrodes), positioned at the level of the sural nerve on the participants' left ankle. Prior to placing the electrodes, the participants' skin was cleaned using alcohol swabs, and any hair was removed using a single-use razor. Electrical shocks lasted ∼30 ms and consisted in a series of ten 1-ms pulses.

Participants' pain tolerance levels were selected using a classical staircase method in which the intensity started at 1 mA and was gradually increased in steps of 2 mA. After each stimulation, participants rated their pain using a visual analog scale (VAS) with the labels "no pain" and "extremely painful." The stimulation intensity was increased until the participants reported not being able to tolerate the next stimulus, and the last intensity delivered was set as the pain tolerance level. The intensity corresponding to each participant's pain tolerance level was divided by 10 to determine 10 evenly distributed shock intensities between the pain detection and tolerance thresholds. The participants then received 32 stimulations presented in a randomized order and rated the intensity of the pain perceived on the VAS. With these ratings, we used a cross-validation procedure to determine the function that best describes the participant's pain sensitivity. The selected model allowed the extrapolation of 10 shock intensities corresponding to 10 levels of pain ranging from 10 to 100% of the pain tolerance level. On average, participants' tolerance level (maximum stimulation) corresponded to a stimulation of 25.6 mA (SD = 12.11, range: 9 to 57 mA).

Decision Task. Participants were told that they would take part in a task in which they would need to accept or reject offers combining various levels of pain and money (verbatim instructions are reported in *SI Appendix, Supplementary Methods*). Participants were made aware that rejecting the offer would allow them to avoid the pain and prevent them from gaining the money, while accepting the offer would result in the immediate delivery of the pain and the chance to earn the money. They performed 100 trials of the decision task spread across 5 runs of 20 trials. These 100 trials fully sampled the 10 × 10 matrix of possible offers (e.g., Fig. 1*B*) and were presented in one of four pseudorandomized orders. Specifically, the 100 trials were split in 5 blocks of 20 trials. In each block, the lower half of the monetary rewards was randomly paired to a pain level, amounting to 10 trials. For the 10 remaining trials, the upper half of the monetary rewards was paired to the pain levels such as each pain level had an overall average compensation of 5$/block and appeared first and second in each block. Pain levels corresponded to 10 values ranging from 10 to 100% of each participant's tolerance threshold evenly spaced on the estimated sensitivity curve, while money levels were fixed at 0, 1.11, 2.22, 3.33, 4.44, 5.56, 6.67, 7.78, 8.89, and 10 $CAD. At each trial of the task (Fig. 1*A*), participants first saw a fixation cross indicating the start of the trial for 3, 4, 5, 6, or 7 s followed by a screen presenting either the pain level or the monetary amount involved in this trial for 2 s. After another fixation cross, participants received the second part of the offer and were asked to make a decision within 5 s using two adjacent keys on a response box. After selecting to accept or reject the offer, participants received a confirmation of their decision for 2 s, followed by either a waiting period until the next trial if they refused the offer or an anticipation period of 3 to 7s followed by the electrical shock and a wait period if they accepted the offer. After the task, the program randomly selected 10 trials, and the money earned in these trials was added to the participants' honorarium. The task was presented using the E-prime 2.0 software (Psychology Software Tools Inc.).

fMRI Data Acquisition. MR data were acquired using a 3-Tesla Siemens Prisma Fit MRI scanner.

T1-weighted segmented k-space, spoiled, and magnetization prepared gradient recalled and inversion recovery structural MRI data were collected using a multiecho MPRAGE sequence (176 slices; repetition time [TR] = 2,200 ms; echo time [TE] = 1.87, 4.11, 6.35, and 8.59 ms; flip angle [FA] = 8°; field of view, field of view [FOV] = 256 × 256 mm; matrix size = 256 × 256; voxel size = 1 × 1 × 1 mm). The root mean square average of the four echo times was used in all analyses.

A spoiled gradient recalled field map (phase encoding: anterior to posterior; 51 slices; TR = 540 ms; TE = 4.92 and 7.38 ms; FA = 60°; FOV = 192 × 192 mm; matrix size = 64 × 64; voxel size = 3 × 3 × 3 mm) was acquired and used to correct for field inhomogeneity during the fMRI acquisition.

Five runs of segmented k-space echo planar single-echo fMRI data were collected (51 slices in interleaved ascending order; TR = 832 ms; TE = 20 ms; FA = 58°; FOV = 192 × 192 mm; matrix size = 64 × 64; voxel size = 3 × 3 × 3 mm; MB factor = 3; in-plane acceleration factor = 2). Each run was 7:05 min in length, during which 510 functional volumes were acquired. Dicoms were converted to NIfTI-1 format. This section was (in part) generated automatically using PyBIDS [0.10.2 (48)].

Analyses.

Behavioral analyses. Participants' choices were scored as 1 if they accepted the offer and 0 if they rejected it. A binomial generalized linear mixed model was used to predict decisions with the level of pain (1 to 10) and the level of money (1 to 10) as fixed effects and participants as a random effect allowing for random intercepts. Response times were calculated as the time elapsed between the onset of the second part of the offer. Trials for which the response time was below 200 ms or for which no response was recorded during the response screen (maximum 5,000 ms) were removed from the analyses (0.3% of trials). The effects of pain and money level on response times were tested using a linear mixed model with pain and money levels as fixed effects and participants as the random effect allowing for random intercepts. The effect of decision on response time was tested in a model with decision, pain level, and money level as fixed effects and participant as the random effect allowing for random intercepts. Choice difficulty was computed as 10 − |money level − pain level|; a value of 10 therefore represents the most difficult choices in which levels of pain and money were equal, and a value of 1 indicates easy choices in which the level of pain or money maximally exceeded the other level. Choice difficulty was entered as the fixed effect in a linear mixed model predicting response times at each trial with participants as the random effect allowing for random intercepts. The significance of all fixed effects was tested using likelihood ratio tests comparing models with and without the effect of interest.

fMRI preprocessing. We preprocessed raw fMRI and anatomical MRI data using fMRIPrep 20.1.1 standardized pipeline (49). We normalized all data to the 2 × 2 × 2 mm International Consortium for Brain Mapping 152 Nonlinear Asymmetrical template version 2009c, RRID: SCR_008796; TemplateFlow ID: MNI152NLin2009cAsym), and additional spatial smoothing was applied to the blood-oxygen-level-dependent (BOLD) data using a 6-mm full width at half maximum Gaussian kernel. See *SI Appendix, Supplementary Methods*, for the full description of the preprocessing operations performed by fMRIPrep.

First-level general linear models for BOLD data. We first created general linear models (GLMs) at the participant level using the nilearn GLM module to estimate the BOLD response to various task events. For all models, the signal was scaled to the mean value in the time axis, and the event regressors were generated using a boxcar function equal to the event duration (for visual stimuli) or a delta function (for shocks) convolved with a canonical hemodynamic response function. The design matrix was filtered using a high-pass filter with a cutoff of 128 s, and autocorrelation was modeled using an autoregressive AR(1) model. The five BOLD runs were concatenated, and a constant regressor was added to model each run's mean. All first-level models presented in this manuscript included nuisance regressors modeling nonneural sources: 24 movement-related regressors comprising the six estimated head movement parameters (x, y, z, roll, pitch, and yaw), their first temporal derivatives, their squares, their squared derivatives, and the first five components of the anatomical CompCor extracted using the CompCor method (50) implemented in fMRIprep. The interstimulus periods were used as the unmodeled baseline in all models (ISIs in Fig. 1*A*).

In the main first-level model of interest, we focused on the effect of the pain and money offers. To this end, we entered regressors for each level of pain or money offered separately for offers in which the pain level was presented first and offers in which the money level was presented first (20 regressors per participant). We also added regressors for the second part of the offers, the key press/feedback, the shock anticipation screen, and the shock delivery. In all multivariate patterns analyses, we used the beta values maps of the regressors of interest scaled by variance to downweight noisy voxels [t values (51)]. We used two additional models with the same structure to model the different levels of anticipated and the intensity of delivered shocks.

To model brain activity during the decision phase, we created two separate GLMs with regressors for the onset of each pain offer level and/or money offer level in the decisions. To predict decisions, we modeled each trial by entering the onset of each decision event as an independent regressor (100 regressors per participant) with a duration corresponding to the duration of the event. Additional regressors of no interest in this model included the onsets for the pain/money offer, the key press/feedback, the shock anticipation screen, and the shock delivery.

Univariate analyses. For the sake of completeness, we considered the univariate effects of the pain and money levels offered, the shock intensity, and the anticipation phase. For each effect (pain level, money level, anticipated intensity, and shock intensity) and participant, we performed a mass univariate regression between the beta maps from the first-level GLM and the predicted levels. The resulting beta coefficients were entered into a group-level two-tailed one-sample *t* test against 0 at the second level. The resulting t-value maps were thresholded using a threshold of $P < 0.05$, FWE corrected. We report these results in addition to a comparison between the univariate maps and the multivariate patterns in *SI Appendix, Supplementary Results*, S8.

Multivariate pattern analysis. To develop multivariate patterns predicting pain and money levels as well as shock intensity, we used the whole-brain parametric maps calculated for each participant and each of the 10 levels of pain and 10 levels of money. Using scikit-learn, we trained separate LASSO-PCR algorithms to predict the level of pain (1 to 10) and money (1 to 10) and assessed their performance using a 10-fold cross-validation with participants as the grouping factor. The LASSO alpha hyperparameter was fixed at 1.0, and we retained all PCA components (scikit-learn default values; 52). In the cross-validation procedure, the participants were split in 10 sets, and the algorithm was trained on ∼90% of the participants and tested on the remaining participants; the procedure was repeated 10 times so that each set was used as the test set. The training data were standardized within each cross-validation loop, and the same scaling was applied on the test data. We report the Pearson correlation coefficient (r) between the actual and predicted values, the coefficient of determination (R^2 as calculated in scikit-learn), and the root mean squared error between actual and predicted levels as measures of performance for the predictions. We used a similar procedure to perform predictions in 6-mm spherical searchlights or in each of the 486 parcels of the Cognitive and Affective Neuroscience lab (CANlab) combined atlas (*SI Appendix, Supplementary Results*, S4; github.com/canlab/Neuroimaging_Pattern_Masks) with the exception that we did not use a penalized regression when predicting within each searchlight/parcel to avoid the impossibility of interpreting results when all coefficients are shrunk to zero in noninformative searchlights/parcels. Instead, we used a principal component regression while retaining components explaining 80% of the variance of the signal when performing predictions in searchlights/parcels.

To calculate the pattern similarity between the pain/money patterns and the parametric maps of each participant and phases of the task, we calculated the cosine similarity between the weights of the full patterns and the parametric values of each map. In this procedure, we always respected the cross-validation splits used to develop the patterns, so that pattern similarity was always calculated on images that were not used in the training of the pattern. Using this pattern similarity, we assessed the capacity of each pattern to discriminate between various levels of pain and money (1 vs. 10, 1 vs. 5, and 5 vs. 10) with receiver operating characteristic curves and single-interval classification and tested the significance of the accuracy using binomial tests. When classes were unbalanced, we report the balanced accuracy measure as calculated in scikit-learn. We used the same approach to apply the patterns to the open datasets, but in these cases we used the weights averaged across all cross-validation folds (52).

To assess which voxels made stable contributions to each pattern, we used a bootstrap resampling procedure in which 10,000 bootstrap samples of the same size as the original dataset were created by randomly sampling observations from the original dataset with replacement. We then applied the whole-brain LASSO-PCR in each of these samples and used the distribution of value at each voxel to calculate a z score and a *P* value. Using these *P* values, we thresholded the maps at $q < 0.05$, FDR corrected, two-tailed for display and interpretation only. All analyses were performed using the full weight maps. For searchlight analyses, we used the same procedure but used 5,000 bootstrap samples to limit the computational demands, and we report the FWE corrected results at $P < 0.05$ for inference.

To confirm that the predictions were above chance and that the cross-validation procedure was unbiased, we performed the cross-validated predictions in 5,000 samples with permuted labels. Specifically, for each cross-validation fold, we permuted the labels within participants, trained the algorithm on the permuted labels, and tested it on the original labels. We used the formula (C + 1)/(N permutations + 1) to calculate the permutation *P* values using the Pearson's correlation coefficient between actual and predicted outcomes, where C is the number of permutations whose correlation coefficient was higher than or equal to the true correlation coefficient.

To predict participants' choices, in the decision phase we applied the multivariate patterns obtained in each cross-validation fold of the LASSO-PCR and applied it to the single trial data of the decision phase of the left out participants. We removed trials with variance inflation factors higher than 2 in the single-trial GLM and for which the response time was lower than 200 ms or for which no response was recorded (0.33% of trials). We used each trial pattern similarity values as input features in a linear SVM classifier (C parameter = 1) to predict the binary outcome of each trial (accept or reject) using the same 10-fold cross-validation split used to develop the patterns, and we report the balanced accuracy as the measure of classification performance. We also performed the same analysis within each participant by performing the same 10-fold cross-validated classification within each participant's 100 trials.

Data Availability. The raw data cannot be openly shared due to local legal restrictions, but all raw, Brain Imaging Data Structure (BIDS)-formatted data necessary to reproduce all analyses will be shared on request. Thresholded and unthresholded patterns, custom code, and a containerized version of the open source software have been deposited in Neurovault, GitHub, and Dockerhub, respectively (https://neurovault.org/collections/10410/, https://github.com/mpcoll/coll_painvalue_2021, and https://hub.docker.com/r/mpcoll2/coll_painvalue_2021).

ACKNOWLEDGMENTS. M.-P.C. was supported by a fellowship from the Canadian Institutes of Health Research and a career award from the Fonds de Recherche Québec–Santé during this work. C.-W.W. was supported by Grants IBS-R015-D1 (Institute for Basic Science) and 2019R1C1C1004512 (National Research Foundation of Korea). T.D.W.'s effort was supported by NIH Grants R01MH076136 and R01DA046064. This work was supported by a discovery grant from the Natural Sciences and Engineering Research (Grant RGPIN-2016-06682) to M.R.

Author affiliations: [a]École de Psychologie, Université Laval, Québec, QC, G1V 0A6, Canada; [b]Center for Interdisciplinary Research in Rehabilitation and Social Integration, Québec, QC, G1M 2S8, Canada; [c]Department of Psychology, McGill University,

Montréal, QC, H3A 1G1, Canada; [d]Center for Neuroscience Imaging Research, Institute for Basic Science, Suwon, 16419, South Korea; [e]Department of Biomedical Engineering, Sungkyunkwan University, Suwon, 16419, South Korea; [f]Department of Psychological and Brain Sciences, Dartmouth College, Hanover, NH 03755; [g]Research Center of the Institut Universitaire de Gériatrie de Montréal, Université de Montréal, Montréal, QC, H3W 1W5, Canada; [h]Department of Stomatology, Université de Montréal, Montréal, QC, H3T 1J4, Canada; [i]Faculty of Dentistry, McGill University, Montréal, QC, H3V 1H9, Canada; [j]Department of Anesthesia, McGill University, Montréal, QC, H4A 3J1, Canada; and [k]Alan Edwards Centre for Research on Pain, McGill University, Montreal, QC, H3A 0G1, Canada

1. B. Seymour, Pain: A precision signal for reinforcement learning and control. *Neuron* **101**, 1029–1041 (2019).
2. G. Crombez, C. Eccleston, S. Van Damme, J. W. S. Vlaeyen, P. Karoly, Fear-avoidance model of chronic pain: The next generation. *Clin. J. Pain* **28**, 475–483 (2012).
3. B. Knutson, S. A. Huettel, The risk matrix. *Curr. Opin. Behav. Sci.* **5**, 141–146 (2015).
4. M. Roy *et al.*, Representation of aversive prediction errors in the human periaqueductal gray. *Nat. Neurosci.* **17**, 1607–1612 (2014).
5. O. Bartra, J. T. McGuire, J. W. Kable, The valuation system: A coordinate-based meta-analysis of BOLD fMRI experiments examining neural correlates of subjective value. *Neuroimage* **76**, 412–427 (2013).
6. B. Knutson, S. M. Greer, Anticipatory affect: Neural correlates and consequences for choice. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **363**, 3771–3786 (2008).
7. W. M. Pauli, R. C. O'Reilly, T. Yarkoni, T. D. Wager, Regional specialization within the human striatum for diverse psychological functions. *Proc. Natl. Acad. Sci. U.S.A.* **113**, 1907–1912 (2016).
8. B. Seymour, N. Daw, P. Dayan, T. Singer, R. Dolan, Differential encoding of losses and gains in the human striatum. *J. Neurosci.* **27**, 4826–4831 (2007).
9. S. Ballesta, W. Shi, K. E. Conen, C. Padoa-Schioppa, Values encoded in orbitofrontal cortex are causally related to economic choices. *Nature* **588**, 450–453 (2020).
10. C. Padoa-Schioppa, J. A. Assad, Neurons in the orbitofrontal cortex encode economic value. *Nature* **441**, 223–226 (2006).
11. A. M. Brooks, G. S. Berns, Aversive stimuli and loss in the mesocorticolimbic dopamine system. *Trends Cogn. Sci.* **17**, 281–286 (2013).
12. J. O'Doherty *et al.*, Dissociable roles of ventral and dorsal striatum in instrumental conditioning. *Science* **304**, 452–454 (2004).
13. S. M. Reynolds, K. C. Berridge, Emotional environments retune the valence of appetitive versus fearful functions in nucleus accumbens. *Nat. Neurosci.* **11**, 423–425 (2008).
14. Z. Zhang *et al.*, Distributed neural representation of saliency controlled value and category during anticipation of rewards and punishments. *Nat. Commun.* **8**, 1907 (2017).
15. A. Shenhav, J. D. Cohen, M. M. Botvinick, Dorsal anterior cingulate cortex and the value of control. *Nat. Neurosci.* **19**, 1286–1291 (2016).
16. A. Shenhav, M. M. Botvinick, J. D. Cohen, The expected value of control: An integrative theory of anterior cingulate cortex function. *Neuron* **79**, 217–240 (2013).
17. B. W. Balleine, M. R. Delgado, O. Hikosaka, The role of the dorsal striatum in reward and decision-making. *J. Neurosci.* **27**, 8161–8165 (2007).
18. M. R. Delgado, J. Li, D. Schiller, E. A. Phelps, The role of the striatum in aversive learning and aversive prediction errors. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **363**, 3787–3800 (2008).
19. E. M. Tricomi, M. R. Delgado, J. A. Fiez, Modulation of caudate activity by action contingency. *Neuron* **41**, 281–292 (2004).
20. M.-C. Albanese, E. G. Duerden, P. Rainville, G. H. Duncan, Memory traces of pain in human cortex. *J. Neurosci.* **27**, 4612–4620 (2007).
21. G. S. Berns *et al.*, Neurobiological substrates of dread. *Science* **312**, 754–758 (2006).
22. M. Khoshnejad *et al.*, Brain processing of the temporal dimension of acute pain in short-term memory. *Pain* **158**, 2001–2011 (2017).
23. S. Palermo, F. Benedetti, T. Costa, M. Amanzio, Pain anticipation: An activation likelihood estimation meta-analysis of brain imaging studies. *Hum. Brain Mapp.* **36**, 1648–1661 (2015).
24. C. Herry, J. P. Johansen, Encoding of fear learning and memory in distributed neuronal circuits. *Nat. Neurosci.* **17**, 1644–1654 (2014).
25. D. Mobbs, D. B. Headley, W. Ding, P. Dayan, Space, time, and fear: Survival computations along defensive circuits. *Trends Cogn. Sci.* **24**, 228–241 (2020).
26. L. J. Chang, P. J. Gianaros, S. B. Manuck, A. Krishnan, T. D. Wager, A sensitive and specific neural signature for picture-induced negative affect. *PLoS Biol.* **13**, e1002180 (2015).
27. S. Becker, W. Gandhi, F. Pomares, T. D. Wager, P. Schweinhardt, Orbitofrontal cortex mediates pain inhibition by monetary reward. *Soc. Cogn. Affect. Neurosci.* **12**, 651–661 (2017).
28. S. Q. Park, T. Kahnt, J. Rieskamp, H. R. Heekeren, Neurobiology of value integration: When value impacts valuation. *J. Neurosci.* **31**, 9307–9314 (2011).
29. D. Talmi, P. Dayan, S. J. Kiebel, C. D. Frith, R. J. Dolan, How humans integrate the prospects of pain and reward during choice. *J. Neurosci.* **29**, 14617–14626 (2009).
30. J. S. Winston, I. Vlaev, B. Seymour, N. Chater, R. J. Dolan, Relative valuation of pain in human orbitofrontal cortex. *J. Neurosci.* **34**, 14526–14535 (2014).
31. H. Slimani, P. Rainville, M. Roy, The aversive value of pain in human decision-making. *Eur. J. Pain* **26**, 668–679 (2022).
32. S. Zorowitz *et al.*, The neural basis of approach-avoidance conflict: A model based analysis. *eNeuro* **6**, ENEURO.0115-19 (2019).
33. Y. Benjamini, Y. Hochberg, Controlling the false discovery rate: A practical and powerful approach to multiple testing. *J. R. Stat. Soc. B* **57**, 289–300 (1995).
34. B. T. T. Yeo *et al.*, The organization of the human cerebral cortex estimated by intrinsic functional connectivity. *J. Neurophysiol.* **106**, 1125–1165 (2011).
35. E. Y. Choi, B. T. T. Yeo, R. L. Buckner, The organization of the human striatum estimated by intrinsic functional connectivity. *J. Neurophysiol.* **108**, 2242–2263 (2012).
36. T. D. Wager *et al.*, An fMRI-based neurologic signature of physical pain. *N. Engl. J. Med.* **368**, 1388–1397 (2013).
37. C.-W. Woo *et al.*, Quantifying cerebral contributions to pain beyond nociception. *Nat. Commun.* **8**, 14211 (2017).
38. N. Kriegeskorte, R. Goebel, P. Bandettini, Information-based functional brain mapping. *Proc. Natl. Acad. Sci. U.S.A.* **103**, 3863–3868 (2006).
39. C. Silva, N. McNaughton, Are periaqueductal gray and dorsal raphe the foundation of appetitive and aversive control? A comprehensive review. *Prog. Neurobiol.* **177**, 33–72 (2019).
40. B. Petre *et al.*, A multistudy analysis reveals that evoked pain intensity representation is distributed acress brain systems. PLoS Biol. **20**, e3001620 (2022).
41. S. Sternberg, Separate modifiability, mental modules, and the use of pure and composite measures to reveal them. *Acta Psychol. (Amst.)* **106**, 147–246 (2001).
42. E. G. Duerden, M.-C. Albanese, Localization of pain-related brain activation: A meta-analysis of neuroimaging data. *Hum. Brain Mapp.* **34**, 109–149 (2013).
43. M. L. Kringelbach, The human orbitofrontal cortex: Linking reward to hedonic experience. *Nat. Rev. Neurosci.* **6**, 691–702 (2005).
44. A. J. Shackman *et al.*, The integration of negative affect, pain and cognitive control in the cingulate cortex. *Nat. Rev. Neurosci.* **12**, 154–167 (2011).
45. G. R. Samanez-Larkin, B. Knutson, Decision making in the ageing brain: Changes in affective and motivational circuits. *Nat. Rev. Neurosci.* **16**, 278–289 (2015).
46. M. Zunhammer, T. Spisák, T. D. Wager, U. Bingel; Placebo Imaging Consortium, Meta-analysis of neural systems underlying placebo analgesia from individual participant fMRI data. *Nat. Commun.* **12**, 1391 (2021).
47. P. Rainville, J. S. Feine, M. C. Bushnell, G. H. Duncan, A psychophysical comparison of sensory and affective responses to four modalities of experimental pain. *Somatosens. Mot. Res.* **9**, 265–277 (1992).
48. T. Yarkoni *et al.*, PyBIDS: Python tools for BIDS datasets. *J. Open Source Softw.* **4**, 1294 (2019).
49. O. Esteban *et al.*, fMRIPrep: A robust preprocessing pipeline for functional MRI. *Nat. Methods* **16**, 111–116 (2019).
50. Y. Behzadi, K. Restom, J. Liau, T. T. Liu, A component based noise correction method (CompCor) for BOLD and perfusion based fMRI. *Neuroimage* **37**, 90–101 (2007).
51. M. Misaki, Y. Kim, P. A. Bandettini, N. Kriegeskorte, Comparison of multivariate classifiers and response normalizations for pattern-information fMRI. *Neuroimage* **53**, 103–118 (2010).
52. G. Varoquaux *et al.*, Assessing and tuning brain decoders: Cross-validation, caveats, and guidelines. *Neuroimage* **145** (Pt B), 166–179 (2017).