

# Opponent appetitive-aversive neural processes underlie predictive learning of pain relief

Ben Seymour<sup>1</sup>, John P O'Doherty<sup>1,2</sup>, Martin Koltzenburg<sup>3</sup>, Katja Wiech<sup>1</sup>, Richard Frackowiak<sup>1,4</sup>, Karl Friston<sup>1</sup> & Raymond Dolan<sup>1</sup>

Termination of a painful or unpleasant event can be rewarding. However, whether the brain treats relief in a similar way as it treats natural reward is unclear, and the neural processes that underlie its representation as a motivational goal remain poorly understood. We used fMRI (functional magnetic resonance imaging) to investigate how humans learn to generate expectations of pain relief. Using a pavlovian conditioning procedure, we show that subjects experiencing prolonged experimentally induced pain can be conditioned to predict pain relief. This proceeds in a manner consistent with contemporary reward-learning theory (average reward/loss reinforcement learning), reflected by neural activity in the amygdala and midbrain. Furthermore, these reward-like learning signals are mirrored by opposite aversion-like signals in lateral orbitofrontal cortex and anterior cingulate cortex. This dual coding has parallels to 'opponent process' theories in psychology and promotes a formal account of prediction and expectation during pain.

Self-preservation and evolution ordain that animals act optimally or near-optimally to minimize harm. One of the principal mechanisms for detecting harm is the pain system, and early prediction is essential to direct appropriate pre-emptive behavior. However, any simple correspondence between predicted sensory input and behavioral output is challenged by considering the nature of relief: for example, mild pain will be rewarding if it directly follows severe pain. This illustrates a critical issue in our understanding of pain relief as an affective and motivational state<sup>1–3</sup> and poses a broader question in emotion research: how do the neural processes that underlie motivation adapt to the context provided by the ongoing affective state?

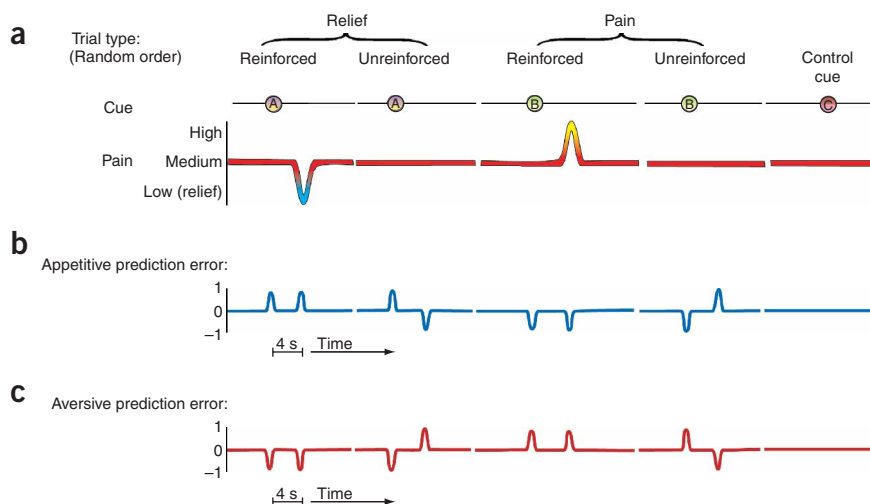
According to psychological theories<sup>4–7</sup>, tonic aversive states recruit reward processes to help direct behavior toward homeostatic equilibrium (which becomes the motivational goal). This may offer insight into why relief is often pleasurable: for example, the experience of cooling oneself in a swimming pool on a hot day. Indeed, the euphoria of relief has been used to help explain a number of seemingly paradoxical behaviors, from sky diving to sauna bathing<sup>8</sup>, in which relief is thought to become the dominant motivational drive. Despite supportive psychological evidence<sup>9–12</sup>, direct observations of neural activity consistent with such appetitive processes are lacking.

Conceptually related issues arise in diverse areas such as engineering, economics and computer science and offer potential insight into the underlying neural processes involved in relief in animals. Notably, computational reinforcement learning models have proved particularly useful in formalizing how the brain learns to predict rewards and punishments<sup>13–19</sup>. These models learn to make

predictions by assessing previous contingencies between environmental cues and motivationally salient outcomes. In theory, these models can be extended to deal with tonic reinforcement and relief, by computing predictions relative to an average rate of reinforcement, rather than according to absolute values<sup>20,21</sup>. However, the extent to which average reward/loss reinforcement learning strategies are implemented in the brain is still unclear. With respect to pain, this may have added importance, as motivational predictions (of pain or relief) are thought to exert substantial influence on the subsequent perception of pain<sup>22,23</sup>. Understanding the neural mechanisms by which predictions are learned is therefore key to our understanding of how the brain intrinsically modulates pain in physiological and clinical situations.

We used fMRI to investigate the pattern of brain responses in nineteen healthy subjects as they learned to predict the occurrence of phasic relief from or exacerbation of tonic pain (see Methods). We employed a first-order pavlovian conditioning procedure with a partial (50%) reinforcement schedule (**Fig. 1a**). Tonic pain was induced using the capsaicin-heat model. Capsaicin is the pain-inducing component of chili pepper; it induces sensitization to heat by activation of temperature-dependent TRPV1 ion channels expressed on peripheral nociceptive neurons. This temperature sensitivity allowed us to deliver constant but easily modifiable levels of pain for long durations, adapted for each individual subject, at temperatures which do not cause skin damage. This provides a unique experimental tool to study pain, as it specifically permits investigation of the neural processes underlying the offset of pain: that is, relief. The model has the further advantage that it

<sup>1</sup>Wellcome Department of Imaging Neuroscience, 12 Queen Square, London WC1N 3BG, UK. <sup>2</sup>Division of the Humanities and Social Sciences 228-77, California Institute of Technology, Pasadena, California 91125, USA. <sup>3</sup>Institute of Child Health, University College London, 30 Guildford Street, London WC1N 1EH, UK. <sup>4</sup>Neuroimaging Laboratory, Fondazione Santa Lucia, Rome 00179, Italy. Correspondence should be addressed to B.S. (bseymour@fil.ion.ucl.ac.uk).



**Figure 1** Experimental design and computational model. **(a)** Experimental design. There were five trial types: cue A was followed by a temperature/pain decrease on 50% of occasions (reinforced and unreinforced relief cue), cue B was followed by a temperature/pain increase on 50% of occasions (reinforced and unreinforced pain cue) and cue C was followed by no change in temperature/pain (control cue). **(b)** Appetitive computational model: predicted neuronal response. Schematic showing the mean representation of the temporal difference prediction error according to the different cue types, where relief is represented as reward. **(c)** Aversive computational model: predicted neuronal response. Schematic showing the aversive temporal difference prediction error, which treats pain exacerbation as punishment. **b** and **c** represent the average predicted neuronal response; the corresponding predicted BOLD response is shown in **Figs. 3c** and **4c**, respectively, following convolution with a canonical hemodynamic response function.

induces the characteristic molecular and cellular changes that mimic physiological injury, and so presents a biologically realistic model of relief in natural and clinical environments.

We applied capsaicin topically to an area (12.5 cm<sup>2</sup>) of skin on the left leg, which caused a localized area of burning pain (which feels similar to sunburn), and manipulated the intensity of this pain with an overlying temperature thermode that matched the capsaicin-treated area. Temperature was adjusted for individual subjects to aim for evoking an average baseline magnitude of pain rated as 6 on a 0–10 categorical scale. Phasic decreases in the baseline temperature to 20 °C caused complete relief of pain, and temperature increases caused exacerbation. We used visual cues (which were abstract colored images) as pavlovian conditioned predictors of these changes. Thus, in the fMRI scanner, subjects learned that certain images tended to predict imminent relief or exacerbation of pain.

We used a computational reinforcement learning (temporal difference) model to identify neural activity consistent with reward-like processing. The characteristic teaching signal of these models is the

prediction error, which is used to direct acquisition and refinement of predictions relating to individual cues. The prediction error records any change in expected affective outcome, and it thus occurs whenever predictions are generated, updated or refuted. By treating relief of pain as reward, and exacerbation as negative reward, we sought to identify activity that correlated with this prediction error signal. We calculated the value of the prediction error for each subject according to the sequence of stimuli they received in order to provide a statistical predictor of fMRI data (as has been done previously<sup>17,18,24</sup>). The use of a partial (probabilistic) reinforcement strategy, in which the cues are only 50% predictive of their outcomes, ensures constant learning and updating of predictions and generates both positive and negative prediction errors throughout the course of the experiment (**Fig. 1b,c**). Thus, inference is based on identification of this dynamic and highly characteristic signal.

In support of the model, our data show that brain activity (that is, blood oxygen level-dependent, or BOLD, activity) in the amygdala and midbrain correlates with the reward prediction error signal predicted by average reward temporal difference learning. In addition,

we show an opponent, aversive representation of the prediction error in lateral orbitofrontal and genual anterior cingulate cortex. Furthermore, these two signals appear to be coexpressed in the ventral striatum.

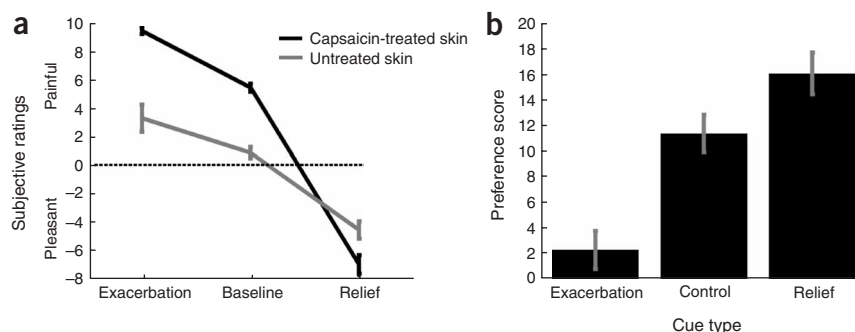
## RESULTS

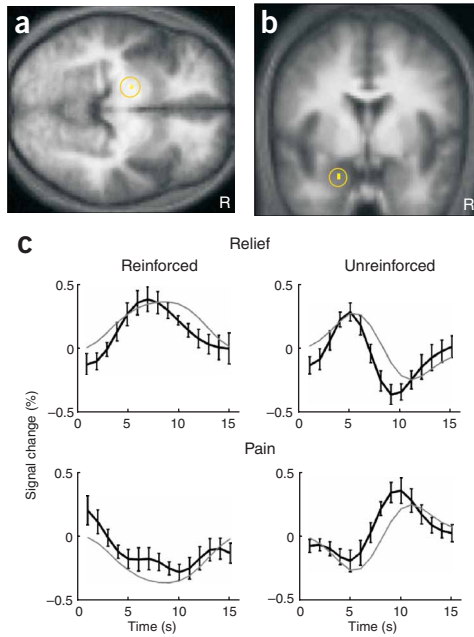
### Behavioral and autonomic results

Subjects rated the baseline thermal stimulation as painful and the decreases and increases in temperature as pleasant or more painful, respectively (**Fig. 2a**). In addition, pleasantness and pain ratings were significantly greater than equivalent temperature changes on adjacent skin not treated with capsaicin ( $P < 0.05$ , all pair-wise comparisons; see Methods).

In a behavioral version of the task outside of the fMRI scanner, we demonstrated conditioning to the relief and exacerbations of pain by engaging the subjects in a supplementary cue-preference task, after the learning task. In this, subjects ( $n = 14$ ) made a forced choice preference judgement of pairs of cues, presented side by side. This demonstrated a

**Figure 2** Behavioral measures. **(a)** Pain ratings. Pain and pleasantness ratings for the baseline level of thermal stimulation, and the phasic increases and decreases in temperature. Scores are on a 0–10 magnitude rating, with error bars representing the s.e.m. The graph shows results for the capsaicin-treated skin and an adjacent area of unaffected skin. **(b)** Preference scores. After the learning experiment, subjects made forced choices between randomized pairs of cues. The scores are out of a maximum of 20 pairings for each cue (with higher scores indicating more preferred).





**Figure 3** Appetitive temporal difference prediction error. **(a,b)** Statistical parametric maps ( $P < 0.001$ ) showing **(a)** left substantia nigra (axial plane) and **(b)** left amygdala (coronal plane). **(c)** Time course of inferred mean neuronal activity for the four principal trial types in left amygdala. Black line represents the data (error bars represent 1 s.e.m.), and thin gray line is the model appetitive temporal difference prediction error (from **Fig. 1b**) after convolution with a canonical hemodynamic response function.

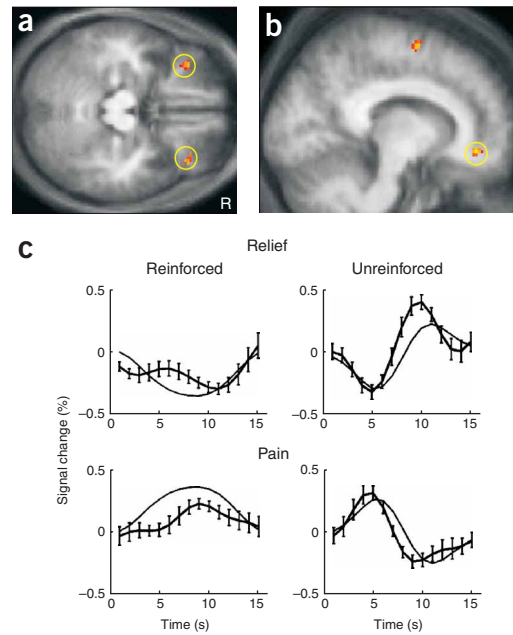
significant preference ordering, with the relief cue preferred to the neutral cue ( $P < 0.05$ , Wilcoxon sign rank test), which was, in turn, preferred to the exacerbation cue ( $P < 0.01$ , Wilcoxon sign rank test; **Fig. 2b**). On post-experimental debriefing (see Methods), only four out of the 14 subjects could report any contingent relationship between the cues and the outcomes.

During the fMRI version of the task, we used physiological measures to assess the acquisition of cue expectations. Heart rate changes induced by the cues correlated with the magnitude of expectations (that is, cue-specific temporal difference values) both of pain relief ( $P < 0.01$ ) and pain exacerbation ( $P < 0.01$ ), calculated from the model (see Methods). This supports the hypothesis that cue expectations are acquired in a manner consistent with the (temporal difference) learning model, albeit in a valence-insensitive manner. That is, we observed increased heart rate with higher valued cues, whether positive or negative, consistent with a learned arousal-like response associated with the expectations.

### fMRI results

We used the model to identify a representation of the appetitive prediction error in the brain (**Fig. 1b**, appetitive model). Activity in left amygdala and left midbrain (in a region consistent with the substantia nigra) correlated with this signal (**Fig. 3a,b**). Time-course analysis illustrates the average pattern of response associated with the different trial types in the amygdala, illustrating a strong correspondence with the predictions of the model (**Fig. 3c**). These data support the hypothesis that relief learning involves a reward-like learning signal.

Recent evidence indicates that temporal difference models also provide an accurate description of aversive learning, suggesting the existence of a separate reinforcement learning mechanism encoding aversive events<sup>18</sup>. We therefore sought to identify whether an aversive



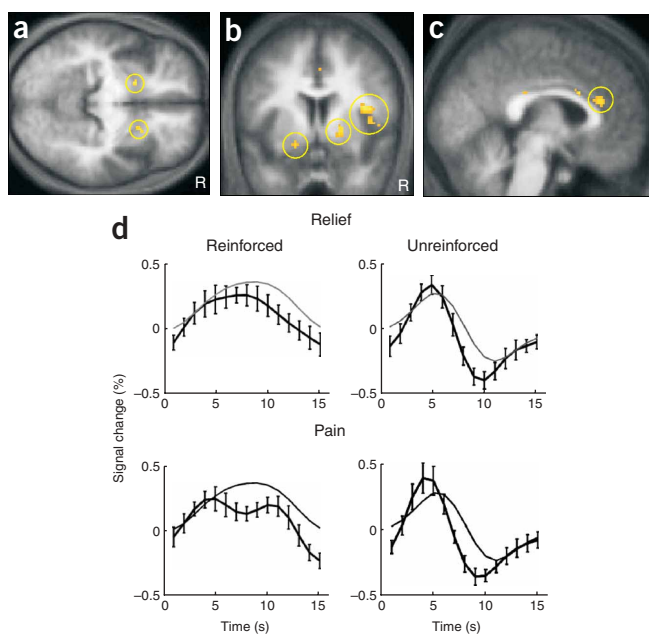
**Figure 4** Aversive temporal difference prediction error. Statistical parametric maps ( $P < 0.001$ ) showing **(a)** lateral orbitofrontal cortex (axial plane) and **(b)** genu anterior cingulate cortex, highlighted (sagittal plane). **(c)** Time course of inferred mean neuronal activity for the four principal trial types in left orbitofrontal cortex. Black line shows data (error bars represent 1 s.e.m.), and thin black line is the model aversive temporal difference prediction error (**Fig. 1c**) after convolution with a canonical hemodynamic response function.

representation of the prediction error was expressed, in which exacerbation of pain was treated as positive punishment, and relief as negative punishment (**Fig. 1c**, aversive model). Activity in bilateral lateral orbitofrontal cortex and genu anterior cingulate cortex correlated with this signal (**Fig. 4a,b**). The time-course of this activity (**Fig. 4c**) illustrates the opposite pattern of response to the appetitive prediction error. These data indicate the existence of an aversive reinforcement signal, distinct from the reward-like signal.

Psychological studies of appetitive-aversive interactions predict that opposing, learning-related activities should converge in some areas<sup>10</sup>. This might occur in areas such as the ventral striatum (and insula cortex), where predictive activity has been observed in both reward and pain learning tasks, albeit in separate studies<sup>17,18,25–28</sup>. This raises a question about how coexpressed aversive and appetitive prediction errors would be represented by the BOLD signal, particularly if they interact. We therefore created a new statistical model that included two regressors, modelling prediction error for relief and exacerbation separately. This model revealed coexpression in the ventral putamen, anterior insula and rostral anterior cingulate cortex (**Fig. 5a–c**). The responses in these regions showed an appetitive prediction error for the relief-related cue, and an aversive prediction error for the exacerbation-related cue (**Fig. 5d**). This pattern of activity is notable, as it cannot result simply from the linear superposition of appetitive and aversive signals, but implies either an interaction between prediction error and cue-valence, or the expression of a single valence-independent prediction error.

### DISCUSSION

Drawing on theoretical considerations provided by computational reinforcement learning<sup>11</sup>, our data provide evidence in support of an opponent motivational model of tonic pain. We observed two distinct



**Figure 5** Appetitive relief-related plus aversive exacerbation-related prediction error. Statistical parametric maps showing activity that correlates with the appetitive prediction error for the relief cue ( $P < 0.001$ ), masked with the aversive prediction error for the exacerbation cue ( $P < 0.001$ ). (a) Bilateral ventral putamen. (b) Bilateral ventral putamen and right anterior insula. (c) Rostral anterior cingulate cortex. (d) Time course of inferred mean neuronal activity for the four principal trial types in left ventral putamen. Thick black line shows the data (error bars represent 1 s.e.m.), and the thin gray and black lines are the model appetitive and aversive temporal difference prediction error, respectively (from Fig. 1b,c) after convolution with a canonical hemodynamic response function.

patterns of neural activity, distinguishable by their expression in separate brain areas, that correlated with the prediction error signals of an opponent temporal difference model. This extends our understanding of human predictive learning beyond the occurrence of phasic events arising from a neutral baseline. Thus, during tonic pain, aversive and appetitive systems seem to be simultaneously involved to encode appropriate goal-directed predictions across the spectrum of positive and negative outcomes. Our observations suggest a formal framework for understanding the homeostatic and motivational processes engaged by pain and may offer a paradigmatic account of motivation during tonic affective states.

The use of the temporal difference algorithm to represent positive and negative deviations of pain intensity from a tonic background level approximates the class of reinforcement learning model termed average-reward models<sup>20,21,29</sup>. Accordingly, predictions are judged relative to the average level of pain, rather than according to an absolute measure. This comparative treatment of motivationally salient predictions is consistent with both neurobiological and economic accounts of homeostatic motivation, which rely critically on change in affective state<sup>2,30,31</sup>.

Implicit in any such model is a representation of the average rate of reinforcement, although the short time window of fMRI precludes investigation of this directly. From an implementational perspective, one argument for opponency relates to consideration of how a long-run average affective state might be represented. Given our demonstration that positive and negative prediction errors are both encoded by one system and are fully mirrored by opposite signals in an opponent

system, the requirement for one system to fully represent both the tonic levels of reinforcement (that is, by sustained elevated activity) with positive and negative phasic predictions simply superimposed, would seem to be obviated. If this is the case, the tonic level of pain would be free to have a distinct representation, a signal that has been suggested to be conveyed by tonic dopamine release<sup>11</sup>.

Mirror opponency has many similarities to the appetitive-aversive reciprocity characteristic of early psychological ‘opponent process’ theories<sup>4–7</sup>. In their various forms, these theories grew out of a requirement both to explain the adaptive changes that occur during and after tonic reinforcement, and to understand the interactions between appetitive and aversive processes that arise in certain specific learning procedures such as conditioned inhibition and trans-reinforcer blocking. Notably, recent electrophysiological recordings of neuronal activity in mice directly indicate the involvement of opponent processes in (context-related) conditioned inhibition, specifically implicating the ventral striatum and amygdala<sup>32</sup>. Thus it seems possible (and fully consistent with a computational account) that, at least in the ventral striatum, a ‘safety signal’ that predicts the absence of future pain might share the same neural substrate as the relief-prediction error seen here. However, we show an appetitive representation in the amygdala, rather than an opponent aversive representation (which we observe in lateral orbitofrontal and genu anterior cingulate cortex). This points to the expression of multiple learning-related neural signals in the amygdala, consistent with the complex, integrative role of this structure (and the various nuclei within) in associative learning and pain<sup>33,34</sup>.

The finding that lateral orbitofrontal cortex demonstrates an aversive prediction error signal is consistent with previous reports of a role for this region in aversive learning<sup>35</sup>. In particular, this area has been shown to be involved in evaluation of aversive stimuli in the context of different motivational states<sup>36</sup> as well as in short-time-scale pain prediction relative to a changing (learned) baseline rate of phasic pain<sup>37</sup>. Taken with the present results, this suggests that learning of aversive value predictions in this region may be mediated by an aversion-specific prediction error signal, particularly in circumstances that require adaptive representations following changing motivational state or context. However, it should also be noted that lateral orbitofrontal cortex may not be exclusively involved in aversive processing, as reward-related responses have also been reported in this region in some circumstances.

In relation to pain, other cortical areas, specifically insula and anterior cingulate cortex, have clear motivational roles and have previously been implicated in the processing of relief-related information<sup>3</sup>. For example, recent neuroimaging studies investigating the expectation and receipt of placebo analgesia implicate these areas in endogenously mediated analgesia<sup>38,39</sup>. Our findings provide further support that these areas have a key role in homeostatic functions relating to pain<sup>2</sup>.

The BOLD signal is thought to correspond to changes (increases or decreases) in synaptic activity, and thus the activity we describe may reflect specific afferent neuromodulatory influences that originate elsewhere<sup>40,41</sup>. Substantial evidence indicates that mesolimbic dopamine neurons both encode reward-related prediction error<sup>16,19</sup> and have a key role in analgesia<sup>42</sup>, suggesting that dopamine could convey an appetitive relief-related prediction error. This draws attention to activity in the ventral striatum, a region that receives strong mesolimbic dopaminergic projections. Comparison with previous data in this area highlights the observation that cues signaling lower-than-predicted pain cause deactivation in the context of a neutral baseline, as opposed to activation in the context of a tonic pain baseline<sup>18,26</sup>. This implicates adaptive changes occurring during tonic pain, influencing ventral striatal activity and consistent with the representation of an appetitive signal for relief-related cues. However, taken alone, it is possible that



this ventral striatal activity is modulated by a single prediction-error signal for both relief and exacerbation cues<sup>43,44</sup>, although recent electrophysiological evidence demonstrating suppression of midbrain dopaminergic neurons to aversive stimuli would seem to require a distinct aversive opponent<sup>45</sup>. Either way, this signal must interact with valence-specific information by some additional mechanism, possibly through the involvement of different intrinsic sub-populations of appetitive and aversive neurons within the ventral striatum<sup>46</sup>.

That pain relief and reward might share a common neural substrate is also suggested by the fact that many drugs that have rewarding effects have analgesic properties. Aside from dopamine, there are many neurotransmitters with clear combined roles in appetitive and aversive motivation, for example opioid peptides, serotonin, substance P and glutamate<sup>3,47,48</sup>. Of particular interest are serotonin-releasing neurons projecting from the dorsal raphe nucleus to the ventral striatum, which have emerged as a plausible candidate to mediate an aversive prediction error<sup>11</sup>.

In addition to a role in pavlovian motivation, it is also clear that pain and relief-related expectations exert a strong influence on the actual subsequent experience of pain, in that perception (of intensity) is weighted by the prior expectancies acquired through conditioning. How predictive motivational values influence perceptual inferences such as pain intensity is not yet clear, although probabilistic perceptual models that incorporate economic cost functions, such as decision theory, may offer insight at a theoretical level<sup>49</sup>. From an implementational perspective, one putative mechanism exploits an influence of 'higher' brain areas on ascending pain pathways via descending modulatory control centers. A possible target is the 'on-' and 'off-' cells of the periaqueductal grey and rostral ventromedial medulla, which show opponent anticipatory pain-related activity under apparent higher control<sup>3</sup>. Whatever the mechanisms, these influences are thought to be clinically important both in endogenous pain modulation (including placebo analgesia) and in the pathogenesis of some chronic pain syndromes<sup>3,23,38,39</sup>, and we suggest that integrated psychological, neurophysiological and computational approaches offer some promise in furthering their understanding.

## METHODS

**Subjects.** Thirty-three healthy right handed subjects (14 in a behavioral version of the task, and 19 in the fMRI version of the task), free of pain or medication, gave informed consent and participated in the study, approved by the Joint National Hospital for Neurology and Neurosurgery (University College London, National Health Service Trust) and Institute of Neurology (University College London) Ethics Committee. Subjects were remunerated for their inconvenience (40 GBP).

**Stimuli: capsaicin model.** We applied topical 1% capsaicin (8-methyl-*N*-vanillyl-6-nonenamide, 98%, Sigma, diluted in 5% ethanol-KY jelly) to the lateral aspect of the left leg over an area of 2.5 × 5 cm, under an occlusive dressing, and left it for 40 min, after which all subjects reported feeling persistent (though bearable) pain, at which time the capsaicin and dressing was removed and the skin cleaned. A thermode matching the size of the capsaicin application area was applied with a loose tourniquet (easily removable in case of unbearable pain) to the treated skin. Temperature was then manipulated using an fMRI-compatible Peltier thermode (MSA thermostet, Somedic). Phasic variations in temperature were made at a rate of 5 °C/s to the predetermined upper and lower levels and were controlled by in-house software.

**Stimuli and pre-experimental set-up.** Before the experiment, required temperature levels for each individual subject were set by slowly increasing the cutaneous temperature overlying the capsaicin treatment site from 20 °C in steps of 0.5 °C, with continual monitoring of pain ratings (on a 0–10 rating scale) to achieve a baseline level of 6/10. Subsequently, subjects received progressively higher phasic increases to determine a satisfactory temperature

for the pain exacerbations, to at least 8/10 ('just tolerable'). Pain relief was induced by phasic cooling to 20 °C, which abolished pain in all subjects.

We obtained subjective ratings of pain for the increase, baseline and decreases in pain. We asked the subjects, "Can you give a score, on a scale of 0 to 10, as to how painful the pain is, where 0 is no pain at all, and 10 is the worst imaginable pain?" We also took subjective ratings of pleasantness for the phasic relief. We first asked the subjects, "Did you find the change in temperature unpleasant or pleasant?" to check that no subjects found the cooling as unpleasant, and then, "Can you give a score, on a scale of 0 to 10, as to how pleasant you found it, where 0 is not at all, and 10 is highest imaginable pleasure?" Phasic changes were repeated with pain and pleasantness ratings on capsaicin-treated skin and on a distant area of non-capsaicin treated skin on the same limb well beyond the area of secondary hyperalgesia, and repeated at the end of the experiment. We achieved mean ratings (s.e.m. in parentheses) for the baseline tonic pain of 5.5/10 (1.1) on capsaicin treated skin and 0.9/10 (1.5) on untreated skin. Phasic increases were rated at 9.3/10 (0.9) for capsaicin-treated skin and 3.3/10 (3.6) on untreated skin. Phasic decreases (relief; measured on the pleasantness scale) were rated at 7.0/10 (2.4) and 4.6/10 (2.3) on untreated skin. All comparisons (treated versus untreated) were significant at  $P < 0.01$  with corresponding  $t$ -tests. After transfer into the scanner or behavioral testing room (with the thermode attached) subjects were in pain for approximately 40 min to 1 h by the time the experiment started. The visual cues were abstract colored pictures.

**Task.** The task was a classical pavlovian delay-conditioning procedure of temperature increases (exacerbations of pain) or decreases (relief of pain). Visual cues were presented for 4 s, at the end of which the phasic pain perturbation was applied for 5 s. The precise timing was determined in psychophysical pilot testing (to accommodate thermode and C-fiber latencies). There were three different visual cues, each presented 30 times. Cue A (relief-related cue) was followed by decreased temperature on 15/30 occasions (50%), cue B (pain exacerbation related cue) was followed by increased temperature on 15/30 occasions (50%), and cue C was followed by no change in temperature on 30/30 occasions. The control condition provides additional control in our parametric design, although it was initially included to permit a more conventional analysis (data not shown). The five different trial types were presented in random order.

**Behavioral measures.** Subjects performed a reaction-time task which consisted of judging whether the visual cue appeared to the left or right of center on the display monitor, as quickly as possible. The resulting reaction times were taken as a behavioral index of conditioning. Performance on this task was not contingent on the stimuli presented, and subjects were told before imaging that their success or failure at quickly judging the position would not affect the amount of pain or relief received. The task was performed with a two-button key press using the right hand. Heart rate was recorded using a pulse oximeter in conjunction with Spike 2 software (CED).

A behavioral version of the task was performed that was identical to that performed in the fMRI scanner, except that it was performed in a testing room with the subject seated in front of a computer monitor. After this task, we performed a supplementary cue-preference task designed to investigate whether the subjects had acquired appetitive and aversive preferences for the cues as a result of the conditioning procedure. In this task, we presented two cues side-by-side and asked the subject to judge which cue they preferred, indicated by a left or right key-press. Each cue-pairing was repeated ten times and was randomized as to which side the cue appeared on. We calculated the preference scores by summing the total number of preference choices made for each cue (as in an all-play-all games table, with a maximum score of 20). Mean scores for each cue were compared across subjects using Wilcoxon sign rank tests.

We did not attempt to formally address the issue of conscious versus non-conscious acquisition of conditioned expectancies. However, to gain some insight into the level of explicit expectancy learning, we asked the question, "Did you recognize any relationship between the pictures and subsequent change in pain level?" at the end of the experiment (for the behavioral version of the task only). Subjects were not told the experiment was a learning and conditioning study beforehand but rather were simply told that it was a study

**Table 1 MNI coordinates and statistical z-scores for the appetitive, aversive and joint coexpressed appetitive-aversive temporal difference prediction error**

Region	Laterality	x	y	z	z-score
<b>Appetitive prediction error</b>					
Midbrain (substantia nigra)	L	-18	-12	-8	3.99
Amygdala	L	-20	2	-26	3.33
<b>Aversive prediction error</b>					
Lateral orbitofrontal cortex	R	40	34	-20	3.72
	L	-34	34	-20	3.71
Genual anterior cingulate cortex	R	10	42	-6	4.24
Motor cortex	R	14	0	60	5.35 <sup>a</sup>
<b>Combined appetitive-aversive prediction error</b>					
Ventral putamen	R	18	8	0	4.08
		22	10	-10	3.32
	L	-18	8	-12	3.62
Anterior insula	R	30	22	6	3.87
		36	2	16	4.78
	L	-34	12	12	4.55
Rostral anterior cingulate cortex	R	2	34	20	3.61

<sup>a</sup>Significant after whole brain correction.

of pain and temperature processing. Ten of fourteen subjects were unable to report any association between cues and outcomes.

**Computational model.** We used a temporal difference model to generate a parametric regressor corresponding to the appetitive prediction error, which was applied to the imaging data, as previously described<sup>17,18</sup>. Here, we used a two-time point temporal difference model with a learning rate ( $\alpha = 0.3$ ) determined from behavioral results (see below). In this model, the value  $v$  of a particular cue (referred to as a state  $s$ ) is updated according to the learning rule:  $v(s) \leftarrow v(s) + \alpha \delta$ , where  $\delta$  is the prediction error. This is defined as  $\delta = r - a + v(s)_{t+1} - v(s)_t$ , where  $r$  is the return (that is, the amount of pain) and  $a$  is the average amount of reinforcement (tonic pain) that was assumed to be constant. We assigned relief and exacerbations of pain as returns of 1 and -1, respectively (that is, a linear scale of pain from relief to exacerbation). This is an arbitrary specification, given that it is difficult to precisely scale the relative oppositely valenced utilities of relief and exacerbations of pain. Thus, the model treats predictions relating to relief of pain on equal par with unexpected omission of exacerbation of pain, and, similarly, it treats exacerbation-related predictions equivalently to unexpected omissions of relief.

**Data acquisition and analysis: behavioral and autonomic measures.** These were taken as measures of cue reinforcement and correlated with the temporal difference value (that is, the cue expectancy). Reaction time data were individually (that is, on a subject-by-subject basis) fit to a gamma cumulative distribution function (using a maximum likelihood function), to allow analysis across subjects, and correlated with the temporal difference value. This yielded a best fit with a learning rate of 0.3, and a significant correlation for both the relief-related and exacerbation-related trials, independently, and in the same direction. That is, reaction times were shorter for both high reward values and high aversive values. To remove any possible confounding effects of early trials, during which reaction time data habituate substantially, we repeated this procedure after removing the first ten trials. This yielded a correlation which just failed to reach significance ( $P = 0.056$ ), across both cue types. We also looked at sensitivity to the initial temporal difference value by setting this to the average value of 0.5, which yielded a non-significant correlation.

The heart rate was found to be approximately normally distributed and was normalized to permit analysis across subjects. We found significant heart rate correlations with both relief and pain cue types (independently, as for the reaction time). For both exacerbation and relief trial types, this yielded a best fit

with a learning rate of 0.3. Across both cue types, this remained significant ( $P < 0.05$ ,  $r = 0.19$ ) after removal of the first ten trials and with use of different initial temporal difference values. This is a robust correlation and is reported in the main text. Consequently, we used a learning rate of 0.3 for the temporal difference model used in the fMRI analysis.

**fMRI.** Functional brain images were acquired on a 3-T Allegra Siemens scanner. Subjects lay in the scanner with foam head restraint pads to minimize any movement associated with the painful stimulation. Images were realigned with the first volume, normalized to a standard EPI template and smoothed using a 6-mm FWHM Gaussian kernel. Realignment parameters were inspected visually to identify any potential subjects with excessive head movement; none was found. Images were analyzed in an event-related manner using the general linear model, with the onsets of each stimulus represented as a delta function to provide a stimulus function. We used a parametric design, in which the temporal difference prediction errors modulated the stimulus functions on a stimulus-by-stimulus basis. The statistical basis of this approach has been described previously<sup>50</sup>. Regressors were then generated by convolving the stimulus function with a hemodynamic response function (HRF). Effects of no interest included the onsets of visual cues, the pain relief and exacerbations themselves and realignment parameters from the image preprocessing to provide additional correction for residual subject motion. Linear contrasts of appetitive prediction errors were taken to a group level (random effects) analysis by way of a one-sample  $t$ -test, and the aversive prediction error was taken as the inverse. MNI coordinates and statistical z-scores are found in **Table 1**. This analysis determines areas which correlate to univalent appetitive or aversive prediction error and does not identify areas in which these signals overlap. To explore the possible representation of distinct prediction error signals for the pain relief and exacerbation trials, we generated two independent regressors for the prediction error occurring at each. Then, we took the appetitive relief and aversive exacerbation components of the prediction error to a second level analysis of variance and exclusively masked the two individual contrasts (that is, we looked for areas of overlap of the independent appetitive-relief and aversive-exacerbation prediction errors, both at  $P < 0.001$ ; **Fig. 5a-c**).

Group level activations were localized according to the group-averaged structural scan. Activations were checked on a subject-by-subject basis using their individual normalized structural scans to ensure correct localization, as some of the reported activations are in small nuclei (for example, substantia nigra). We report activity in areas in which we had prior hypotheses on the basis of previous data, though without specification of laterality. These regions have established roles in both aversive and appetitive predictive learning, and included ventral putamen, head of caudate, midbrain (substantia nigra), anterior insula cortex, cerebellum, anterior cingulate cortex, amygdala, lateral orbitofrontal cortex, medial orbitofrontal cortex, dorsal raphe and ventral tegmental area. We report activations at a threshold of  $P < 0.001$ , with a minimum size of five contiguous voxels. We also report brain activations outside our areas of interest that survive whole-brain correction for multiple comparisons (**Table 1**) using family-wise error correction at  $P < 0.05$ .

We performed a supplementary fixed-effects analysis on a trial basis to determine impulse responses, as previously described<sup>18</sup>. Note that this analysis refers to the average impulse response across each trial throughout the experiment and does not embody the time-dependent nature of learning incorporated within the main parametric analysis.

#### ACKNOWLEDGMENTS

We wish to thank P. Dayan and N. Daw for many helpful discussions and O. Josephs, B. Johanssen and C. Rickard for technical assistance. This research was funded by The Wellcome Trust.

#### COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.

Received 15 June; accepted 2 August 2005

Published online at <http://www.nature.com/natureneuroscience/>

1. Cabanac, M. Physiological role of pleasure. *Science* **173**, 1103–1107 (1971).
2. Craig, A.D. A new view of pain as a homeostatic emotion. *Trends Neurosci.* **26**, 303–307 (2003).

3. Fields, H. State-dependent opioid control of pain. *Nat. Rev. Neurosci.* **5**, 565–575 (2004).
4. Solomon, R.L. & Corbit, J.D. An opponent-process theory of motivation. I. Temporal dynamics of affect. *Psychol. Rev.* **81**, 119–145 (1974).
5. Konorski, J. *Integrative Activity of the Brain: an Interdisciplinary Approach* (Chicago, University of Chicago Press, 1967).
6. Schull, J. A conditioned opponent theory of Pavlovian conditioning and habituation. in *The Psychology of Learning and Motivation* (ed. Bower, G.) 57–90 (Academic, New York, 1979).
7. Grossberg, S. Some normal and abnormal behavioral syndromes due to transmitter gating of opponent processes. *Biol. Psychiatry* **19**, 1075–1118 (1984).
8. Solomon, R.L. The opponent-process theory of acquired motivation: the costs of pleasure and the benefits of pain. *Am. Psychol.* **35**, 691–712 (1980).
9. Solomon, R.L. Recent experiments testing an opponent-process theory of acquired motivation. *Acta Neurobiol. Exp. (Wars.)* **40**, 271–289 (1980).
10. Dickenson & Dearing, M.F. Appetitive-aversive interactions and inhibitory processes. in *Mechanisms of Learning and Motivation*. (eds. Dickinson, A. & Boakes, R.A.) 203–231 (Erlbaum, Hillsdale, New Jersey, 1979).
11. Daw, N.D., Kakade, S. & Dayan, P. Opponent interactions between serotonin and dopamine. *Neural Netw.* **15**, 603–616 (2002).
12. Tanimoto, H., Heisenberg, M. & Gerber, B. Experimental psychology: event timing turns punishment to reward. *Nature* **430**, 983 (2004).
13. Barto, A.G. Adaptive critics and the basal ganglia. in *Models of Information Processing in the Basal Ganglia* (eds. Houk, J.C., Davis, J.L. & Beiser, D.G.) 215–232 (MIT Press, Cambridge, Massachusetts, 1995).
14. Sutton, R.S. & Barto, A.G. *Reinforcement Learning: an Introduction*. (MIT Press, Cambridge, Massachusetts, 1998).
15. Montague, P.R., Dayan, P. & Sejnowski, T.J. A framework for mesencephalic dopamine systems based on predictive Hebbian learning. *J. Neurosci.* **16**, 1936–1947 (1996).
16. Schultz, W., Dayan, P. & Montague, P.R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
17. O'Doherty, J.P., Dayan, P., Friston, K., Critchley, H. & Dolan, R.J. Temporal difference models and reward-related learning in the human brain. *Neuron* **38**, 329–337 (2003).
18. Seymour, B. *et al.* Temporal difference models describe higher-order learning in humans. *Nature* **429**, 664–667 (2004).
19. Dayan, P. & Balleine, B.W. Reward, motivation, and reinforcement learning. *Neuron* **36**, 285–298 (2002).
20. Schwartz, A. A reinforcement learning method for maximizing undiscounted rewards. in *Proceedings of the Tenth International Conference on Machine Learning*. 298–305 (Morgan Kaufmann, San Mateo, California, 1993).
21. Mahadevan, S. Average reward reinforcement learning: Foundations, algorithms and empirical results. *Mach. Learn.* **22**, 1–38 (1996).
22. Fields, H.L. Pain modulation: expectation, opioid analgesia and virtual pain. *Prog. Brain Res.* **122**, 245–253 (2000).
23. Price, D.D. *Psychological Mechanisms of Pain and Analgesia* (IASP, Seattle, 1999).
24. Tanaka, S.C. *et al.* Prediction of immediate and future rewards differentially recruits cortico-basal ganglia loops. *Nat. Neurosci.* **7**, 887–893 (2004).
25. Ploghaus, A. *et al.* Dissociating pain from its anticipation in the human brain. *Science* **284**, 1979–1981 (1999).
26. Jensen, J. *et al.* Direct activation of the ventral striatum in anticipation of aversive stimuli. *Neuron* **40**, 1251–1257 (2003).
27. McClure, S.M., Berns, G.S. & Montague, P.R. Temporal prediction errors in a passive learning task activate human striatum. *Neuron* **38**, 339–346 (2003).
28. Setlow, B., Schoenbaum, G. & Gallagher, M. Neural encoding in ventral striatum during olfactory discrimination learning. *Neuron* **38**, 625–636 (2003).
29. Daw, N.D. & Touretzky, D.S. Long-term reward prediction in TD models of the dopamine system. *Neural Comput.* **14**, 2567–2583 (2002).
30. Markowitz, H. The utility of wealth. *J. Polit. Econ.* **60**, 151–158 (1952).
31. Camerer, C., Loewenstein, G. & Prelec, D. Neuroeconomics: how neuroscience can inform economics. *J. Econ. Lit.* (in the press).
32. Rogan, M.T., Leon, K.S., Perez, D.L. & Kandel, E.R. Distinct neural signatures for safety and danger in the amygdala and striatum of the mouse. *Neuron* **46**, 309–320 (2005).
33. Watkins, L.R. *et al.* Neurocircuitry of conditioned inhibition of analgesia: effects of amygdala, dorsal raphe, ventral medullary, and spinal cord lesions on antianalgesia in the rat. *Behav. Neurosci.* **112**, 360–378 (1998).
34. Holland, P.C. & Gallagher, M. Amygdala-frontal interactions and reward expectancy. *Curr. Opin. Neurobiol.* **14**, 148–155 (2004).
35. O'Doherty, J., Kringelbach, M.L., Rolls, E.T., Hornak, J. & Andrews, C. Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat. Neurosci.* **4**, 95–102 (2001).
36. Small, D.M., Zatorre, R.J., Dagher, A., Evans, A.C. & Jones-Gotman, M. Changes in brain activity related to eating chocolate: from pleasure to aversion. *Brain* **124**, 1720–1733 (2001).
37. Glascher, J. & Buchel, C. Formal learning theory dissociates brain regions with different temporal integration. *Neuron* **47**, 295–306 (2005).
38. Petrovic, P., Kalso, E., Petersson, K.M. & Ingvar, M. Placebo and opioid analgesia—imaging a shared neuronal network. *Science* **295**, 1737–1740 (2002).
39. Wager, T.D. *et al.* Placebo-induced changes in fMRI in the anticipation and experience of pain. *Science* **303**, 1162–1167 (2004).
40. Logothetis, N.K., Pauls, J., Augath, M., Trinath, T. & Oeltermann, A. Neurophysiological investigation of the basis of the fMRI signal. *Nature* **412**, 150–157 (2001).
41. Stefanovic, B., Warnking, J.M. & Pike, G.B. Hemodynamic and metabolic responses to neuronal inhibition. *Neuroimage* **22**, 771–778 (2004).
42. Altier, N. & Stewart, J. The role of dopamine in the nucleus accumbens in analgesia. *Life Sci.* **65**, 2269–2287 (1999).
43. Horvitz, J.C. Mesolimbocortical and nigrostriatal dopamine responses to salient non-reward events. *Neuroscience* **96**, 651–656 (2000).
44. Smith, A.J., Becker, S. & Kapur, S. A computational model of the functional role of the ventral-striatal D2 receptor in the expression of previously acquired behaviors. *Neural Comput.* **17**, 361–395 (2005).
45. Ungless, M.A., Magill, P.J. & Bolam, J.P. Uniform inhibition of dopamine neurons in the ventral tegmental area by aversive stimuli. *Science* **303**, 2040–2042 (2004).
46. Roitman, M.F., Wheeler, R.A. & Carelli, R.M. Nucleus accumbens neurons are innately tuned for rewarding and aversive taste stimuli, encode their predictors, and are linked to motor output. *Neuron* **45**, 587–597 (2005).
47. Johansen, J.P. & Fields, H.L. Glutamatergic activation of anterior cingulate cortex produces an aversive teaching signal. *Nat. Neurosci.* **7**, 398–403 (2004).
48. Gadd, C.A., Murtra, P., De Felipe, C. & Hunt, S.P. Neurokinin-1 receptor-expressing neurons in the amygdala modulate morphine reward and anxiety behaviors in the mouse. *J. Neurosci.* **23**, 8271–8280 (2003).
49. Dayan, P. & Abbott, L.F. *Theoretical Neuroscience: Computational and Mathematical Modeling of Neural Systems* (MIT Press, Cambridge, Massachusetts, 2001).
50. Buchel, C., Holmes, A.P., Rees, G. & Friston, K.J. Characterizing stimulus-response functions using nonlinear regressors in parametric fMRI experiments. *Neuroimage* **8**, 140–148 (1998).