Learning about pain: The neural substrate of the prediction error for aversive events

Alexander Ploghaus*†‡, Irene Tracey*, Stuart Clare*, Joseph S. Gati§, J. Nicholas P. Rawlins†, and Paul M. Matthews*

*Centre for Functional MRI of the Brain, Department of Clinical Neurology, University of Oxford, John Radcliffe Hospital, Oxford OX3 9DU, United Kingdom; †Department of Experimental Psychology, University of Oxford, South Parks Road, Oxford OX1 3UD, United Kingdom; and §Laboratory for Functional Magnetic Resonance Research, John P. Robarts Research Institute, 100 Perth Drive, London, ON N6A 5K8, Canada

Communicated by Lawrence Weiskrantz, University of Oxford, Oxford, United Kingdom, June 9, 2000 (received for review February 10, 2000)

Associative learning is thought to depend on detecting mismatches between actual and expected experiences. With functional magnetic resonance imaging (FMRI), we studied brain activity during different types of mismatch in a paradigm where contrasting-colored lights signaled the delivery of painful heat, nonpainful warmth, or no stimulation. When painful heat stimulation was unexpected, there was increased FMRI signal intensity in areas of the hippocampus, superior frontal gyrus, cerebellum, and superior parietal gyrus that was not found with mismatch between expectation and delivery of nonpainful warmth stimulation. When painful heat stimulation was unexpectedly omitted, the FMRI signal intensity decreased in the left superior parietal gyrus and increased in the other regions. These contrasting activation patterns correspond to two different mismatch concepts in theories of associative learning (Rescorla-Wagner, temporal difference vs. Pearce-Hall, Mackintosh). Searching for interventions to specifically modulate activation of these brain regions therefore offers an approach to identifying new treatments for chronic pain, which often has a substantial associative learning component.

Learning cues of impending pain allows future painful events to be anticipated and avoided (1, 2). Thus, learning associations between pain and predictive cues has fundamental adaptive value. However, such learning also can have adverse effects. It can exacerbate the unpleasantness of pain (3, 4) and can contribute to chronic pain states (5). We previously have identified brain regions activated by cues associated with pain (6). In this paper, we isolate brain regions that play a critical role in learning cue–pain associations.

Learning of a cue–outcome association only takes place when there is a mismatch between outcome and the expectations based on perceived cues (7, 8). The first goal of the present study, therefore, was to identify brain regions whose activation pattern is consistent with detecting mismatches between the expectation and the delivery of painful stimulation.

The related second goal was to study the extent to which this mismatch-related brain activity conforms to predictions derived from theories of Pavlovian conditioning (9–12), which have formalized the relationship between mismatch detection and associative learning. In these models, mismatch is represented as the difference $\lambda - V$. Applied to associative learning about pain, λ represents the intensity of the actual pain and V represents the accumulated strength of the cue–pain association (i.e., the expectation of pain based on the cue). Associative learning corresponds to changes in V. When the intensity of a painful stimulus exceeds expectation, V is increased proportionally to the magnitude of the mismatch. The rate of learning decelerates over successive learning trials as V approaches λ.

What happens, however, when the expectation exceeds the actual pain? In this situation, different models of associative learning have distinct predictions. In one theory (9), mismatch is formulated as $(\lambda - V)$. Thus, when V exceeds λ, the term takes on a negative value. The term $(\lambda - V)$ will be called *signed mismatch*. An alternative formulation (11) calculates mismatch as an absolute value, $|\lambda - V|$. Hence, mismatch assumes a positive value irrespective of whether the actual pain exceeds the expectation or vice versa. The term $|\lambda - V|$ will therefore be called *absolute mismatch*. Under conditions where the expectation of pain exceeds the actual pain, a brain region signaling absolute mismatch should activate, whereas a region signaling signed mismatch should deactivate in comparison to a low-mismatch control condition.

The current experiment had three phases. In the *acquisition* phase, we presented subjects with three intensities of thermal stimulation (painful hot, nonpainful warm, and no stimulation). Subjects learned to anticipate the type of stimulation to be delivered, because each type was signaled in advance by a specific color of light. When the color-temperature associations had been firmly acquired so that occurrences of the painful stimulus were anticipated, subjects received a painful stimulus during the colored light signaling the absence of stimulation. This stimulus marked the phase of *counterexpected pain*. In the *extinction* phase, the colored lights were presented, but no thermal stimulation was delivered.

Areas signaling mismatch were determined by comparing the functional magnetic resonance imaging (FMRI) signal during events representing high mismatch (e.g., first pain, counterexpected pain) with the FMRI signal during temporally adjacent events representing low mismatch (second pain and last pain of acquisition, respectively). The signed and absolute mismatch models were compared by contrasting brain responses to high- and low mismatch under conditions where the expectation exceeded the actual pain, that is, in extinction (first and second presentation of the light formerly associated with pain, respectively). Similar comparisons involving warm stimulation tested the specificity of the identified activation regions for mismatch involving pain.

Methods

Subjects and Neuroimaging. Twelve healthy, right-handed subjects (seven male and five female) with ages ranging from 23 to 30 years participated in the study. All subjects gave informed consent, and the study was approved both by the Oxfordshire Committee for Research Ethics and The University of Western Ontario Ethics Review Board. Data were acquired on a Varian/Siemens 4 T whole-body scanner with a quadrature birdcage head coil. Head movements were restrained with foam pads. In each of 15 contiguous planes, 280 blood oxygenation level-dependent (BOLD) images were acquired by using multishot echo-planar imaging (EPI) with TE = 15 ms, TR = 2.5 s, flip angle = 45°, in-plane resolution = 3.5 mm, slice thickness = 8 mm, and no slice gap. Slices were prescribed parallel to the anterior commissure–posterior commissure (AC-PC) line and

Abbreviations: FMRI, functional magnetic resonance imaging; LED, light-emitting diode; CP, counterexpected pain; GFs, superior frontal gyrus; BA, Brodmann's areas; GPs, superior parietal gyrus; SEF, supplementary eye field; PPC, posterior parietal cortex.

‡To whom reprint requests should be addressed. E-mail: alex@fmrib.ox.ac.uk.

The publication costs of this article were defrayed in part by page charge payment. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. §1734 solely to indicate this fact.

Article published online before print: *Proc. Natl. Acad. Sci. USA*, 10.1073/pnas.160266497. Article and publication date are at www.pnas.org/cgi/doi/10.1073/pnas.160266497

covered the entire brain volume. Structural images were obtained with a standard T1-weighted pulse sequence.

Psychological Task. Noxious and warm thermal stimuli were applied to the dorsum of the left hand with a 3 × 3 cm Peltier thermode, designed and built in-house. Subjects were individually thresholded three times for their levels of pain and warm stimulation after having been positioned in the scanner. Two stimuli that were consistently described by the subject as "painfully hot" and "clearly warm, but not hot" were chosen. Three color light-emitting diodes (LEDs; red, green, and blue) were mounted at the subjects' feet level and could be viewed through a mirror in the head coil.

The experiment had three phases. In *acquisition*, subjects received five painful (P1 to P5) and five warm (W1 to W5) stimulations in pseudorandom (PR) order. Each type of stimulation was consistently signaled by a certain LED for each subject (randomized across subjects), which preceded the onset of thermal stimulation by a PR interval (mean ± SD, 7.5 ± 5 s) and remained on during the 11 s of thermal stimulation. Between conditioning trials, a third LED was presented signaling a rest period of PR duration (26.5 ± 9 s) in which no stimulation occurred. In *counterexpected pain* (CP), subjects received a painful stimulus during the signaled rest period immediately after the last acquisition trial. In *extinction*, the LEDs previously associated with painful and warm stimulation each were presented twice followed by no thermal stimulation (EP1 and -2, and EW1 and -2, respectively). Presentation parameters were as in acquisition. Subjects were instructed to determine the contingencies between light color and intensity of thermal stimulation presented to them during the experiment. Immediately after the study, subject awareness of these contingencies was assessed in a postexperimental interview. Subjects also were asked to rate the two thermal stimuli on two 11-point visual analogue scales of pain intensity and unpleasantness.

Data Analysis. Image processing and statistical analysis were carried out with MEDX 3.0 (Sensor Systems, Sterling, VA). All volumes were realigned to the first volume and a mean image was created. The data were smoothed by using a 3.5 × 3.5 × 5 mm (full width at half maximum) Gaussian kernel, and the average of every volume was normalized to the same mean value. Because the question under study implies a confound between experimental condition and time, particular care was taken to minimize the effect of FMRI signal fluctuations unrelated to the experimental paradigm. Statistical comparisons were performed on instances of experimental conditions closest in time (e.g., by comparing P1 with P2 rather than with P5). High- and low-frequency noise was removed from the time series by using linear detrending and wavelet temporal filtering. In addition, local baseline changes in FMRI signal were estimated by using the rest period before and after an experimental condition, and volumes collected during the condition corrected for the baseline change (13).

We carried out three independent statistical comparisons to reveal brain activation corresponding to the signed mismatch model, $(\lambda - V)$, and the absolute mismatch model, $|\lambda - V|$. First, we compared brain activation during P1 and P2. During P1, signed and absolute mismatch have a high positive value because the delivered pain ($\lambda > 0$) is not expected ($V = 0$). Both types of mismatch have a smaller positive value during P2 because the delivered pain is expected ($V > 0$). Second, we compared brain activation during CP and P5. During CP, signed and absolute mismatch have a maximum positive value because pain is experienced ($\lambda > 0$) but the absence of pain has been predicted ($V < 0$). Both types of mismatch are approximately 0 during P5 because the experienced pain has been predicted ($V \cong \lambda$). Finally, we compared brain activation after offset of EP1 and EP2 (i.e., at the time pain would have been applied in acquisition). At EP1, signed mismatch has a high negative value and absolute mismatch a high positive value, because pain has been predicted ($V > 0$) but is not experienced ($\lambda = 0$). At EP2, signed mismatch has a smaller negative value and absolute mismatch a smaller positive value, because subjects have now already experienced the light followed by no pain ($V \cong \lambda$). This comparison does not involve painful stimulation and therefore also serves to exclude alternative accounts of our observations based on unreported changes in pain perception or differential motor activity.

Any brain voxel that signals absolute mismatch should be more active in P1 than in P2, more active in CP than in P5, and more active in EP1 than in EP2. Any brain voxel that signals signed mismatch should also be more active in P1 than in P2, and more active in CP than in P5; however, signed mismatch regions should contrast with absolute mismatch regions, by being more active in EP2 than in EP1 (as error negativity decreases). We conducted two conjunction analyses, the first assessing absolute mismatch and the second assessing signed mismatch.

The correction for hemodynamic delay was 5 s, and analyses were conducted with 3 vol related to each of the trials P1 to P5, W1 to W5, and CP, and 4 vol related to the termination of each of the trials EP1, EP2, EW1, and EW2. Activation maps were calculated for each of the comparisons for each subject by using Student's parametric unpaired t test. Cluster detection was performed on all voxels above $z = 2$ to determine significant activations ($P < 0.05$), and the cluster maps were combined for the conjunction analyses. The resulting activations were rendered on the subject's individual anatomical scan, and their anatomical location was determined in coronal, sagittal and axial views by using surface features, three-dimensional sectional anatomy, and MRI images (14, 15). The group-averaged time courses are discontinuous and show mean ± SEM of FMRI percentage signal change for two voxels within the conjunction activation region. Time courses during P1 to P5 were tested for significant trends by polynomial contrasts analysis of variance (familywise $P < 0.05$; Bonferroni correction).

Activation areas were tested for their specificity for pain by carrying out analogous comparisons involving the nonpainful, warm stimulus. For acquisition, W1 was compared with W2, and for extinction, EW1 was compared with EW2.

Results

Behavioral Results. Postexperimental interview established that all subjects who participated in the FMRI study became aware of the stimulus contingencies during the experiment. Subjects correctly reported the relationship between light color and intensity of thermal stimulation for acquisition, as well as the subsequent changes in this relationship. Subjects rated painful heat significantly higher than nonpainful warmth on two 11-point visual analogue scales measuring intensity [7.3 ± 1.3 ("moderate-strong pain") versus 2.3 ± 0.9 ("warm, no pain"); $P < 0.01$] and unpleasantness [4.9 ± 1.7 ("distressing") versus 1.0 ± 0.2 ("comfortable"); $P < 0.01$] of somatosensory stimulation. No changes were reported in perceived pain intensity or unpleasantness during the course of the experiment.

FMRI Results. As previously reported (6), we observed clear activation of brain regions associated with the experience of pain (16–18) when contrasting painful with warm stimulation (including insula, anterior cingulate cortex, cerebellum, somatosensory cortices, and thalamus; data not shown).

The regions most commonly found active during mismatch between the expectation and the delivery of painful stimulation were the hippocampal formation and parahippocampal gyrus (collectively termed the hippocampal system), the superior parietal gyrus (GPs), the superior frontal gyrus (GFs), and the cerebellum (≥8 of 12 subjects). Table 1 specifies the frequency and laterality of activations that are consistent with either the

Table 1. Numbers of subjects showing significant activations in conjunction analyses testing predictions derived from either the absolute mismatch model or the signed mismatch model

Region	Total activating with either model	Absolute mismatch model			Signed mismatch model		
		Total either left or right	Left	Right	Total either left or right	Left	Right
Hippocampal system	10	9	7	4	2	1	1
GFs							
Dorsal	10	9	7	4	3	2	2
Ventral	10	6	3	4	8	5	4
Cerebellum	10	10	7	9	2	2	1
GPs	10	4	2	3	9	8	4
Putamen	3	2	1	1	1	1	0
Caudate nucleus	1	1	0	1	0	0	0

Each conjunction analysis combines three independent t tests ($P < 0.05$).

signed or the absolute mismatch model. Data are shown for brain regions activated in at least 8 of 12 subjects, as well as for regions of theoretical interest (19, 20).

Hippocampal System (Fig. 1). FMRI signal change corresponded more frequently to the absolute mismatch model [nine subjects; mean Talairach coordinates: $x = -24$ mm, $y = -22$ mm; $z = -12$ mm (left); $x = 25$ mm, $y = -29$ mm, $z = -10$ mm (right)] than to the signed mismatch model (two subjects, see Table 1). Fig. 1*a* shows hippocampal activation consistent with absolute mismatch (red) for a typical subject: The locus of activation was the hippocampal formation in seven subjects, and the parahippocampal gyrus was the locus in four subjects. Fig. 1*b* depicts the Talairach coordinates of hippocampal activations related to absolute (red triangles) and signed (black squares) mismatch in coronal view for all individual subjects. The mean time course of FMRI signal change (Fig. 1*c*) during P1 to P5 shows no significant trend. Activation areas were specific to mismatch involving pain. In no subject was the identified area activated in a conjunction analysis of the comparisons W1 versus W2 and EW1 versus EW2.

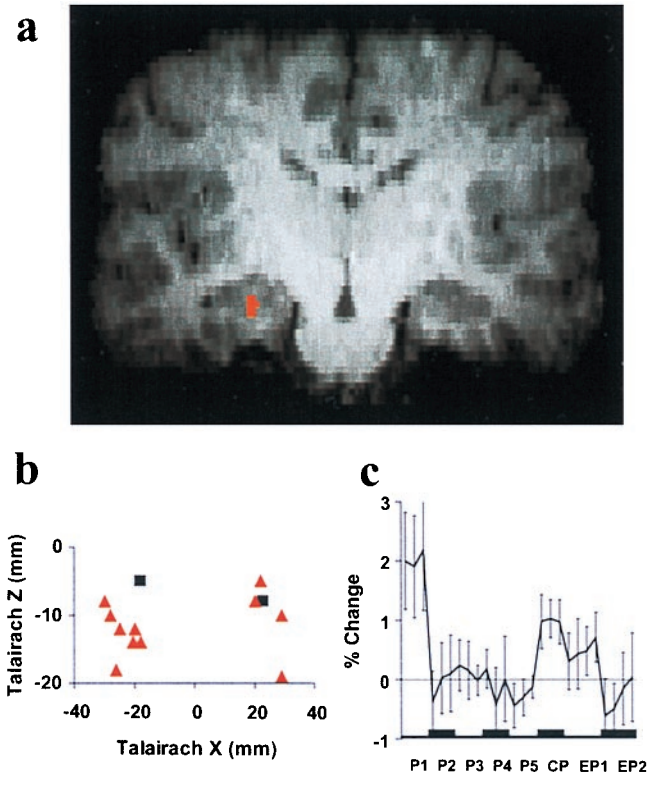

Fig. 1. Hippocampus. (*a*) Typical activation (red) identified from a conjunction analysis combining three independent t tests ($P < 0.05$) of predictions derived from the absolute mismatch model. (*b*) Talairach coordinates of individual subjects' activation centers identified in conjunction analyses testing predictions derived from either the absolute mismatch model (red triangles) or the signed mismatch model (black squares). (*c*) Group-averaged time course of FMRI signal (percentage change, mean ± SEM) with painful stimuli in the acquisition phase (P1 to P5), during CP, as well as after the two extinction trials (EP1 and EP2).

Dorsal GFs [The Intersection of GFs with Brodmann's Areas (BA) 8 and 6; Fig. 2]. FMRI signal change again corresponded more frequently to the absolute mismatch model [nine subjects; mean Talairach coordinates: $x = -17$ mm, $y = 33$ mm, $z = 48$ mm (left); $x = 16$ mm, $y = 36$ mm, $z = 48$ mm (right)] than to the signed model (three subjects, see Table 1). Ventral GFs exhibited a more variable response pattern (Table 1). Fig. 2*a* shows an activation in dorsal GFs consistent with absolute mismatch (red)

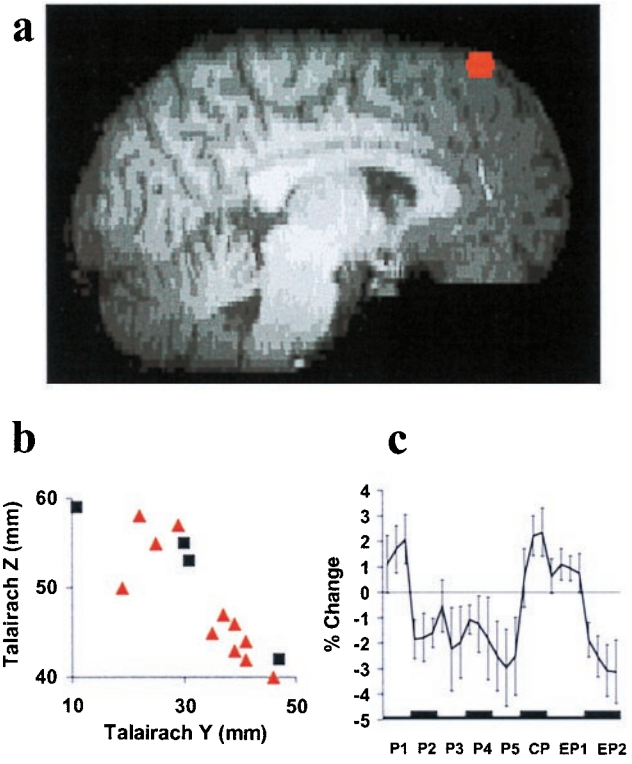

Fig. 2. Dorsal GFs. See Fig. 1 for details of representations *a*–*c*.

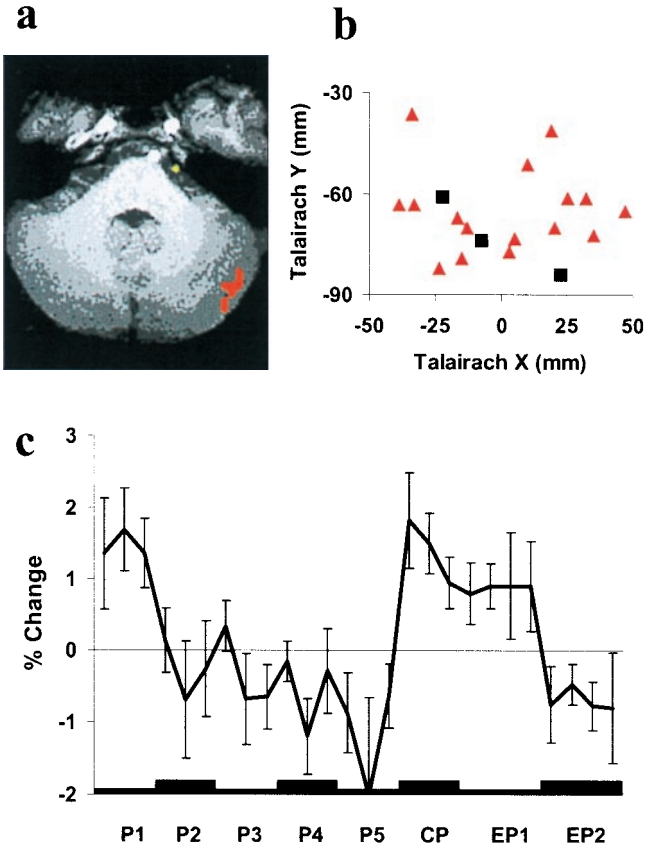

Fig. 3. Cerebellum. See Fig. 1 for details of representations *a–c*.

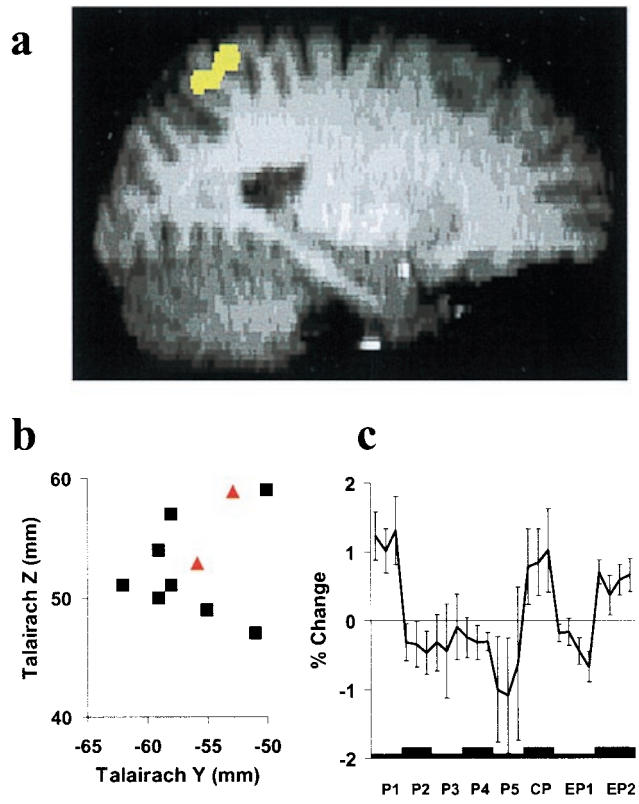

Fig. 4. Left GPs. (*a*) Typical activation (yellow) identified from a conjunction analysis combining three independent *t* tests ($P < 0.05$) of predictions derived from the signed mismatch model. See Fig. 1 for details of representations *b* and *c*.

for a typical subject, and Fig. 2*b* depicts the Talairach coordinates of activations related to absolute (red triangles) and signed (black squares) mismatch in sagittal view for all individual subjects. The mean time course of FMRI signal change in dorsal GFs (Fig. 2*c*) during P1 to P5 shows a significant quadratic trend [$F\,(1, 8) = 10.903$; $P < 0.05$]. Activation areas were again specific to mismatch involving pain. In no subject was the identified area activated in a conjunction analysis of the comparisons W1 versus W2 and EW1 versus EW2.

Cerebellum (Fig. 3). FMRI signal change again corresponded more frequently to the absolute mismatch model [10 subjects; mean Talairach coordinates: $x = -24$ mm, $y = -66$ mm, $z = -26$ mm (left); $x = 23$ mm, $y = -63$ mm, $z = -25$ mm (right)] than to the signed model (two subjects, see Table 1). Fig. 3*a* shows a typical activation in the cerebellum consistent with absolute mismatch (red), and Fig. 3*b* depicts the Talairach coordinates of cerebellar activations related to absolute (red triangles) and signed (black squares) mismatch in horizontal view for all individual subjects. The mean time course of FMRI signal change (Fig. 3*c*) during P1 to P5 differs from the one found in the dorsal GFs in that it shows a significant linear trend [$F\,(1, 9) = 10.483$; $P < 0.05$]. In no subject was the identified area activated in a conjunction analysis of the comparisons W1 versus W2 and EW1 versus EW2.

Left GPs (Fig. 4). Left GPs differs from the other brain areas in that FMRI signal change corresponded more frequently to the signed mismatch model [eight subjects; mean Talairach coordinates: $x = -26$ mm, $y = -56$ mm, $z = 52$ mm] than to the absolute mismatch model (two subjects, see Table 1). The right GPs exhibited a weaker and more variable response pattern (Table 1). Activations in left GPs related to signed mismatch were clustered in BA 7 near the intraparietal sulcus. Fig. 4*a* shows a typical activation in left GPs consistent with signed mismatch (yellow), and Fig. 4*b* depicts the Talairach coordinates of activations related to signed (black squares) and absolute (red triangles) mismatch in sagittal view for all individual subjects. The mean time course of FMRI signal change in the left GPs (Fig. 4*c*) during P1 to P5 exhibits no significant trend. In no subject was the identified area activated in a conjunction analysis of the comparisons W1 versus W2 and EW2 versus EW1.

Discussion

Associative learning about pain is an important adaptive behavior. Like other types of associative learning, it is thought to depend critically on detecting mismatches between expected and actual experience (7–12). The first goal of the present study was therefore to identify brain areas whose activation pattern is consistent with detecting mismatches between expected and actual pain. All subjects became aware of the mismatches presented. Functional imaging during the task revealed areas in the hippocampal system, the GFs, the GPs, and the cerebellum that were activated during mismatches between expected and actual pain only.

The second goal was to compare mismatch-related brain responses to predictions derived from two different theoretical accounts of mismatch. Signed and absolute mismatch differ in their predictions for FMRI signal change during extinction, but not for acquisition or CP. We found that in extinction, the hippocampus, dorsal GFs, and cerebellum activated at EP1 relative to EP2, which is consistent with the absolute mismatch model (11). The left GPs, in contrast, deactivated at EP1 relative to EP2, which is consistent with the signed mismatch model (9). The extinction condition also is relevant because no painful stimuli are delivered, so that simple

alternative accounts based on perceived pain-intensity changes or motor behavior can be excluded.

The FMRI signal from the dorsal GFs and the cerebellum decreased over successive painful stimulations in the acquisition phase. This finding is consistent with mismatch models, because mismatch decreases as subjects learn to expect the pain. The polynomial order of this trend differed for the two brain regions, suggesting different speeds of learning. It is unlikely that habituation or desensitization can account for these slopes, because a painful stimulus of identical characteristics delivered directly after the acquisition phase (i.e., the CP) resulted in a considerably enhanced FMRI signal (see Figs. 1*c*, 2*c*, 3*c*, and 4*c*). In addition, areas typically involved in pain perception showed no slopes in our task (6).

The activation areas related to prediction error for pain were not identified by a conjunction analysis of unexpected warm stimulation and its omission. This result suggests that pain engages an associative learning system distinct from that for biologically less-relevant stimuli.

Hippocampal System. Activation of the hippocampal system during mismatch is consistent with comparator theories of hippocampal function (21–23). These theories maintain that one function of the hippocampus is to compare actual and expected stimuli (i.e., stimuli registered in memory).

Our study shows activation of the same hippocampal regions during three different types of mismatch. The first type of mismatch is novelty. Novelty is presentation of a stimulus in the absence of any particular expectation, as exemplified by unexpected pain (P1). Subjects were familiar with the painful stimulus (from thresholding), but they had no *a priori* knowledge as to with which visual cue the painful stimulus would be paired. Hippocampal involvement in novelty detection has been shown with single-cell recording (21, 24) and hippocampal lesions (25). In functional imaging studies, hippocampal activation was observed when contrasting novel and familiar stimuli, e.g., novel and familiar faces (26), complex pictures (27, 28), line drawings (29), and artificial grammar strings (30). This evidence clearly suggests that the hippocampus is activated when there is sensory input but no memories relevant to it in the given context.

The second type of mismatch is presentation of a stimulus when different or no stimulation is expected, as exemplified by CP. Hippocampal involvement in this type of mismatch has been demonstrated. Vinogradova (21) showed that any change within an otherwise uniform train of stimuli (e.g., decrease in intensity) brings about the reappearance of a previously habituated hippocampal response. Honey and colleagues (31) showed in rats that hippocampal lesions result in a failure to detect mismatches that are generated when an auditory stimulus associated with one visual stimulus is presented with a different (but equally familiar) visual stimulus. This evidence suggests that the hippocampus is activated when there is sensory input and recall of memories relevant to the context, but the two sources of information are conflicting.

The third type of mismatch is absence of an expected stimulus, as exemplified by unexpected omission of pain (EP1). Indirect demonstrations of hippocampal involvement in this type of mismatch come from single-cell recording (32) and from functional brain imaging during pseudoconditioning (33). This involvement suggests that the hippocampus is engaged when there is no sensory stimulation but memories relevant to the context are at variance with this experience.

The present study shows that within each subject, the same hippocampal area is activated during all three types of mismatch, thereby providing evidence for Gray's comparator theory of hippocampal function (22). In contrast to brain imaging studies of novelty using emotionally neutral material (26–30), hippocampal activation areas in our study seem to be specific to mismatch related to an emotionally relevant stimulus, pain. Special processing of emotionally relevant material in certain hippocampal regions may be related to noradrenergic and serotonergic innervation, which seems to be involved in signaling stimulus significance and to be responsible for gating information flow within the hippocampus (22).

Conditioning theories suggest that the mismatch calculation is a necessary condition for associative learning. If only the hippocampus computes the mismatch signal, associative learning should not be possible without it. In fact, hippocampal lesions impair associative learning only under specific memory requirements, e.g., formation of complex associations or rapid switching of associations (23, 34, 35).

Dorsal GFs. Activation in the dorsal GFs during mismatch centered mostly around BA 8 and the supplementary eye field (SEF), which corresponds to area 8B of Walker (36). The SEF has previously been found active in eye movement tasks as well as in covert attentional shift tasks where attention was redirected to a peripheral stimulus without making an eye movement (37). It was concluded that the SEF forms part of a network that controls the allocation of attention in space (37). Our finding could therefore be explained by enhanced attention to the stimulation site during mismatch.

There is previous evidence for mismatch-related neuronal activity in the SEF during associative learning of eye movements. Monkeys learned to associate visual cues with saccadic eye movements in specific spatial directions, and subsequently the cue–direction associations were reversed. Studies of similar learning tasks by Rescorla (38) suggest that during this task, a hierarchical [cue → (eye movement → reward)] association is being learned. Mismatch can therefore be assessed only when all components are known, i.e., after the eye movement that results in delivery of the reward. In line with this argument, activity in a subpopulation of SEF neurons decreased after the eye movement as monkeys learned the associations, reappeared with the reversal, and decreased again as they learned the new associations (39). In addition, lesions of the eye fields of the frontal cortex have been shown to impair associative learning of eye movements (40). The SEF activation during mismatch in our task may therefore support the formation of associations between the visual cues and allocation of attention to the stimulation site.

Cerebellum. It is well established that cerebellar lesions disturb conditioning of cutaneo-muscular reflexes. Studies of eye-blink conditioning suggest that the learning mechanism resides in the cerebellum (41). We therefore speculate that cerebellar activity related to mismatch in our study might support associative learning of a protective movement (e.g., hand withdrawal). Single-unit responses related to mismatch have been found in an input structure to the cerebellum, the dorsal accessory olive (42), suggesting that the mismatch signal might not be generated locally. Instead, it might be provided by projections from the hippocampus (43).

Interestingly, the cerebellar region activated during mismatch is largely identical to the area we found active during the light associated with pain (6). Mismatch-related activity determines new learning or *encoding*, whereas activity during the light most plausibly represents the consequences of *retrieval* from memory. The cerebellar FMRI time courses during encoding (Fig. 3*c*) and retrieval (6) show prolonged, linear change, suggesting a slower learning process than in other areas studied. The identified activation sites were not involved in mismatch related to warm stimulation. This finding is consistent with our speculation of cutaneo-muscular reflex conditioning, as innocuous stimulation does not support such learning.

It may be argued that our observation of cerebellar activation is attributable to unexpected events evoking stronger motor responses than anticipated ones. However, it is not likely that this interpretation can account for our observation, because the

extinction phase did not contain any painful stimuli that could provoke unconditioned motor responses. Yet there was clearly cerebellar activation.

Posterior Parietal Cortex (PPC). The homologue monkey area of BA 7, 7m, or PGm (D. N. Pandya, personal communication; ref. 44) receives equally strong inputs from visual and somatosensory sources (45). The prediction error signal found in BA 7 of the PPC may therefore subserve associative learning between the two modalities by changing their connectivity within BA 7. In support of this argument, Miltner *et al*. (46), in a study of human visual-to-somatosensory associative learning, found increasing γ coherence between pairs of electrodes placed rostrocaudally across medial PPC as learning progressed. Moreover, rats with lesions of the cholinergic input to the PPC failed to show enhanced associative learning about a stimulus when that stimulus coincided with a mismatch (47).

Apkarian *et al*. (48) showed that FMRI signal change in BA 5 and 7 in response to repeated, identical, painful stimulation correlated with mean pain intensity ratings. This observation is consistent with our hypothesis that BA 7 signals prediction error. Subjects are likely to form an expectation of how painful stimuli in the experiment are going to be. As with most quantitative judgements, this expectation will be based on a weighted average of intensities encountered on previous pain trials (49). With a simple predictor function as observed by Apkarian *et al*., increases in pain intensity therefore correspond to positive mismatches, and decreases to negative mismatches.

The PPC has been implicated in attention (50), and the left PPC specifically in covert motor attention (51). It is unlikely that such attentional processes (and likewise, arousal processes) can account for the pattern of BA 7 activation observed in our study. These accounts predict that EP1 evokes more attention and/or arousal than does EP2 because omission of pain is more surprising at EP1 (52). In contrast, we found PPC activation during EP2 relative to EP1 (Fig. 2). Apkarian *et al*. (48) have suggested a role for BA 7 in pain intensity perception (48). Our observation cannot be explained in terms of this account, as one of the comparisons in our conjunction analysis did not involve pain.

Conclusion. The present study revealed brain areas whose activation pattern is consistent with detecting mismatches between actual and expected pain. Our finding supports psychophysiological evidence suggesting that learning theories containing a mismatch term may provide a plausible description of the neural computations during associative learning (53). Comparator theories have repeatedly implicated the hippocampus in mismatch detection, and our study provides compelling evidence in favor of this view. The involvement in mismatch detection of multiple brain loci and their specificity for pain indicates that the capacity for associative learning is not implemented in a unitary neural substrate, but in multiple sites that differ in their learning parameters and may be specific to individual response systems. Searching for interventions to specifically modulate activation of these brain regions offers an approach to identifying new treatments for chronic pain, which often has a substantial associative learning component.

We thank D. N. Pandya and M. F. S. Rushworth for neuroanatomical advice, D. Dobson for help with building the Peltier device, and S. Smith for image analysis advice. A.P. holds a Junior Research Fellowship at Merton College, Oxford, U.K.; and I.T., P.M.M., and the Oxford Centre for FMRI of the Brain are funded by the Medical Research Council (U.K.). This work was supported by the McDonnell-Pew Program in Cognitive Neuroscience.

1. Pavlov, I. P. (1927) *Conditioned Reflexes* (Oxford Univ. Press, Oxford).
2. Mowrer, O. H. (1938) *Psychol. Rev.* **45,** 62–91.
3. Weisenberg, M. (1987) *Behav. Res. Ther.* **25,** 301–312.
4. Chapman, C. R. (1996) *Prog. Brain Res.* **110,** 63–81.
5. Lethem, J., Slade, P. D., Troup, J. D. & Bentley, G. (1983) *Behav. Res. Ther.* **21,** 401–408.
6. Ploghaus, A., Tracey, I., Gati, J. S., Clare, S., Menon, R. S., Matthews, P. M. & Rawlins, J. N. P. (1999) *Science* **284,** 1979–1981.
7. Kamin, L. J. (1969) in *Punishment and Aversive Behavior*, eds. Campbell, B. A. & Church, R. M. (Appleton–Century–Crofts, New York), pp. 279–296.
8. Sokolov, E. N. (1963) *Annu. Rev. Physiol.* **25,** 545–580.
9. Recorla, R. A. & Wagner, A. R. (1972) in *Classical Conditioning II: Current Research and Theory*, eds. Black, A. H. & Prokasy, W. F. (Appleton–Century–Crofts, New York), pp. 64–99.
10. Mackintosh, N. J. (1975) *Psychol. Rev.* **82,** 276–298.
11. Pearce, J. M. & Hall, G. (1980) *Psychol. Rev.* **87,** 332–352.
12. Ploghaus, A. (1997) *Klassische Konditionierung und Schizophrenie* (Logos, Berlin).
13. LaBar, K. S., Gatenby, J. C., Gore, J. C., LeDoux, J. E. & Phelps, E. A. (1998) *Neuron* **20,** 937–945.
14. Duvernoy, H. M. (1991) *The Human Brain: Surface, Three-Dimensional Sectional Anatomy and MRI* (Springer, Berlin).
15. Amaral, D. G. & Insausti, R. (1990) in *The Human Nervous System*, ed. Paxinos, G. (Academic, San Diego), pp. 711–755.
16. Coghill, R. C., Sang, C. N., Maisog, J. M. & Iadarola, M. J. (1999) *J. Neurophysiol.* **82,** 1934–1943.
17. Casey, K. L. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 7668–7674.
18. Bushnell, M. C., Duncan, G. H., Hofbauer, R. K., Ha, B., Chen, J.-I. & Carrier, B. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 7705–7709.
19. Mirenowicz, J. & Schultz, W. (1996) *Nature (London)* **397,** 449–451.
20. Redgrave, P., Prescott, T. J. & Gurney, K. (1999) *Trends Neurosci.* **22,** 146–151.
21. Vinogradova, O. S. (1975) in *The Hippocampus*, eds. Isaacson, R. L. & Pribram, K. H. (Plenum, New York), Vol. 2, pp. 3–70.
22. Gray, J. A. (1982) *The Neuropsychology of Anxiety* (Oxford Univ. Press, Oxford).
23. Schmajuk, N. A. (1997) *Animal Learning and Cognition* (Cambridge Univ. Press, Cambridge, U.K.).
24. Fried, I., MacDonald, K. A. & Wilson, C. L. (1997) *Neuron* **18,** 753–765.
25. Knight, R. T. (1996) *Nature (London)* **383,** 256–259.
26. Haxby, J. V., Ungerleider, L. G., Horwitz, B., Maisog, J. M., Rapoport, S. I. & Grady, C. L. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 922–927.
27. Tulving, E., Markowitsch, H. J., Kapur, S., Habib, R. & Houle, S. (1994) *NeuroReport* **5,** 2525–2528.
28. Stern, C. E., Corkin, S., Gonzales, G. R., Guimares, A. R., Baker, J. R., Jennings, P. J., Carr, C. A., Sugiura, R. M., Vedantham, V. & Rosen, B. (1996) *Proc. Natl. Acad. Sci. USA* **93,** 8660–8665.
29. Gabrieli, J. D., Brewer, J. B., Desmont, J. E. & Glover, G. H. (1997) *Science* **276,** 264–266.
30. Strange, B. A., Fletcher, P. C., Henson, R. N. A., Friston, K. J. & Dolan, R. J. (1999) *Proc. Natl. Acad. Sci. USA* **96,** 4034–4039.
31. Honey, R. C., Watt, A. & Good, M. (1998) *J. Neurosci.* **18,** 2226–2230.
32. Kang, E. & Gabriel, M. (1998) *Hippocampus* **8,** 491–510.
33. Blaxton, T. A., Zeffiro, T. A., Gabrieli, J. D. E., Bookheimer, S. Y., Carrillo, M. C., Theodore, W. H. & Disterhoft, J. F. (1996) *J. Neurosci.* **16,** 4032–4040.
34. Weiskrantz, L. (1978) in *Functions of the Septo-Hippocampal System*, eds. Elliot, K. & Whelan, L. (Elsevier, Amsterdam), pp. 373–387.
35. Weiskrantz, L. & Warrington, E. K. (1975) in *The Hippocampus*, eds. Isaacson, R. L. & Pribram, K. H. (Plenum, New York), Vol. 2, pp. 411–428.
36. Bruce, C. J. (1988) in *Neurobiology of Neocortex*, eds. Rakic, P. & Singer, W. (Wiley, New York), pp. 297–329.
37. Corbetta, M. & Shulman, G. L. (1998) *Philos. Trans. R. Soc. London B* **353,** 1353–1362.
38. Rescorla, R. A. (1990) *J. Exp. Psychol. Anim. Behav. Processes* **16,** 326–334.
39. Chen, L. L. & Wise, S. P. (1995) *J. Neurophysiol.* **73,** 1101–1121.
40. Passingham, R. E. (1993) *The Frontal Lobes and Voluntary Action* (Oxford Univ. Press, Oxford).
41. Krupa. D. J. & Thompson, R. F. (1995) *Proc. Natl. Acad. Sci. USA* **92,** 5097–5101.
42. Sears, L. L. & Steinmetz, J. E. (1991) *Brain Res.* **545,** 114–122.
43. Schmahmann, J. D. & Pandya, D. N. (1993) *J. Comp. Neurol.* **337,** 94–112.
44. Eidelberg, D. & Galaburda, A. M. (1984) *Arch. Neurol.* **41,** 843–852.
45. Cavada, C. & Goldman-Rakic, P. S. (1989) *J. Comp. Neurol.* **287,** 393–421.
46. Miltner, W. H. R., Braun, C., Arnold, M., Witte, H. & Taub, E. (1999) *Nature (London)* **397,** 434–436.
47. Bucci, D. J., Holland, P. C. & Gallagher, M. (1998) *J. Neurosci.* **18,** 8038–8046.
48. Apkarian, A. V., Darbar, A., Krauss, B. R., Gelnar, P. A. & Szeverenyi, N. M. (1999) *J. Neurophysiol.* **81,** 2956–2963.
49. Baron, J. (1988) *Thinking and Deciding* (Cambridge Univ. Press, Cambridge, U.K.).
50. Gitelman, D. R., Nobre, A. C., Parrish, T. B., LaBar, K. S., Kim, Y. H., Meyer, J. R. & Mesulam, M. M. (1999) *Brain* **122,** 1093–1106.
51. Rushworth, M. F., Nixon, P. D., Renowden, S., Wade, D. T. & Passingham, R. E. (1997) *Neuropsychologia* **35,** 1261–1273.
52. Siddle, D. A. (1991) *Psychophysiology* **28,** 245–259.
53. Kopp, B. & Wolff, M. (2000) *Biol. Psychol.* **51,** 223–246.