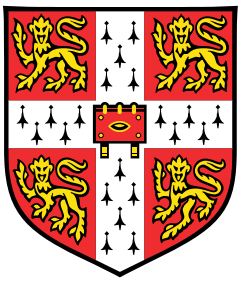
Encoding and Decoding of Pain Relief in the Human Brain



Suyi Zhang

Department of Engineering
University of Cambridge

This dissertation is submitted for the degree of
Doctor of Philosophy

Peterhouse December 2018

Abstract

Encoding and Decoding of Pain Relief in the Human Brain.

Suyi Zhang

The studies in this thesis explored how pain and its relief are represented in the human brain. Pain and relief are important survival signals that motivate escape from danger and search for safety, however, they are often evaluated by subjective descriptions only. Studying how humans learn and adapt to pain and relief allows objective investigation of the information processing and neural circuitry underlying these internal experiences.

My research set out to use computational learning models to provide mechanistic explanations for the behavioural and functional neuroimaging data collected in pain/relief learning experiments with independent groups of healthy human participants.

With a Pavlovian acute pain conditioning task in Experiment 1, I found that 'associability' (a form of uncertainty signal) had a crucial role in controlling the learning rates of different conditioned responses, and can be used to anatomically dissociate underlying neural systems.

Experiment 2 focused on relief learning of terminating a tonic pain stimulus, in which the priority for relief-seeking is in conflict with the general suppression of cognition and attention. I showed that associability during active learning not only controls the relief learning rate, but also correlates with endogenously modulated (reduced) ongoing pain.

This finding was confirmed in Experiment 3 using an independent active relief learning paradigm in a complex dynamic environment. Critically, both experiments showed that associability was correlated with responses in the pregenual anterior cingulate cortex (pgACC), a brain region previously implicated in aspects of endogenous pain control related to attention and controllability. This provided a potential computational account of an information-sensitive endogenous analgesic mechanism.

In Experiment 4, I explored the implications of endogenous controllability for technology-based pain therapeutics. I designed an adaptive closed-loop system that learned to control pain

stimulation using decoded real-time pain representations from the brain. Subjects were shown to actively enhance the discriminability of pain only in the pgACC, and uncertainty during learning again correlated with endogenously modulated pain and were associated with pgACC responses.

Together, these studies (i) show the importance of uncertainty in controlling learning during both acute and tonic pain, (ii) describe how uncertainty also flexibly modulates pain to maximise the impact of learning, (iii) illustrate a central role for the pgACC in this process, and (iv) reveal the implications for future technology-based therapeutic systems.

Declaration

I hereby declare that except where specific reference is made to the work of others, the contents of this dissertation are original and have not been submitted in whole or in part for consideration for any other degree or qualification in this, or any other university. This dissertation is my own work and contains nothing which is the outcome of work done in collaboration with others, except as specified in the text and Acknowledgements. This dissertation contains fewer than 65,000 words including appendices, bibliography, footnotes, tables and equations and has fewer than 150 figures.

All neuroimaging experiments were performed by myself, with contributions and assistance from collaborators and colleagues, at the Center for Information and Neural Networks, Japan, and the Advanced Telecommunications Research Institute, Japan.

Suyi Zhang
December 2018

Acknowledgements

I would like to thank my supervisor Dr Ben Seymour. The work presented here would not exist without his invaluable support and guidance. I would also like to express my gratitude to all my colleagues and collaborators in Cambridge and Japan, especially the imaging teams at the Center for Information and Neural Networks and the Advanced Telecommunications Research Institute for their assistance in performing the studies. Thank you.

Financial support was generously provided by the WD Armstrong Fund and the Cambridge Trust.

Table of contents

List of publications

Research articles

- Zhang S, Mano H, Ganesh G, Robbins T, and Seymour B. Dissociable Learning Processes Underlie Human Pain Conditioning. *Current Biology* **26**.1 (2016), 52-58.
  (Experiment 1)

- Zhang S, Mano H, Lee M, Yoshida W, Robbins T, Kawato M, and Seymour B. The Control of Tonic Pain by Active Relief Learning. *eLife* **7** (2018), e31949.
  (Experiment 2 and 3)

- Zhang S, Mano H, Yoshida W, Yanagisawa T, Shibata K, Kawato M, and Seymour B. Endogenous Controllability of Brain-Machine Interfaces for Pain. *bioRxiv* (2018), 369736.
  (Experiment 4)

Review article

- Zhang S and Seymour B. Technology for Chronic Pain. *Current Biology* **24**.18 (2014), R930-R935.

Collaboration

- Mano H, Yoshida W, Shibata K, Zhang S, Koltzenburg M, Kawato M, and Seymour B. Thermosensory Perceptual Learning Is Associated with Structural Brain Changes in Parietal-Opercular (SII) Cortex. *Journal of Neuroscience* **37**.39 (2017), 9380-9388.

List of figures

List of tables

Nomenclature

Roman Symbols

A	Action
H	Entropy
Q	Value (action / state-action)
R	Reinforcement / Outcome
S	State
T	Time (trial) or State transition function
V	Value (state)

Greek Symbols

α	Associability / Learning rate
δ	Prediction error
ε	Explore probability (ε-greedy rule)
γ	Discount factor
π	Policy
τ	Temperature (softmax rule)

Acronyms / Abbreviations

A-O	Action-Outcome
AAL	Automated anatomical labelling (fMRI atlas)
ACC	Anterior cingulate cortex
BIC	Bayesian Information Criterion
BLA	Basolateral component of amygdala
BMI	Brain-machine interface
BMS	Bayesian model selection
BOLD	Blood oxygenation level dependent
CeA	Central nucleus of amygdala
CHEPS	Contact heat-evoked potential stimulator
CR	Conditional response
CS	Conditional stimulus

DBS	Deep brain stimulation
DLPFC	Dorsolateral prefrontal cortex
EMG	Electromyography
EPI	Echo-planar imaging
fMRI	Functional magnetic resonance imaging
FWER	Family-wise error rate
GLM	General linear model
HMM	Hidden Markov model
HRF	Haemodynamic response function
ITI	Inter-trial interval
MDP	Markov decision processes
MLE	Maximum likelihood estimation
MNI	Montreal Institute of Neurology space (fMRI)
MPRAGE	Magnetisation prepared rapid acquisition gradient echo
MVPA	Multivariate pattern analysis
NAc	Nucleus accumbens
OFC	Orbitofrontal cortex
PAG	Periaqueductal grey
PB	Parabrachial area
PFC	Prefrontal cortex
pgACC	Pregenual anterior cingulate cortex
PH	Pearce-Hall rule
RL	Reinforcement learning
ROI	Region of interest
rtfMRI	Real-time functional magnetic resonance imaging
RMV	Rostroventral medulla
RW	Rescorla-Wagner rule
S-A-O	Stimulus-Action-Outcome
S-R	Stimulus-Response
SCR	Skin conductance response
SEM	Standard error of the mean
SII	Secondary somatosensory cortex
SI	Primary somatosensory cortex
SPM	Statistical parametric mapping
STD	Standard deviation
SUIT	Spatially unbiased infratentorial template (fMRI atlas)

SVC	Small volume correction
TD	Temporal difference learning
TE	Echo time (fMRI)
TMS	Transcranial magnetic stimulation
TR	Repitition time (fMRI)
UR	Unconditional response
US	Unconditional stimulus
VAS	Visual analogue scale
VBA	Variational Bayesian analysis
VLPFC	Ventrolateral prefrontal cortex
VMPFC	Ventromedial prefrontal cortex
VTA	Ventral tegmental area

Chapter 1

Introduction

Pain and pain relief are vital for survival. The sensory component of pain signals the presence of danger, while its alleviation brings relief, which is associated with safety. To act as effective survival aids, pain and relief need additional components to engage responses. The presence of pain as a negative event is inherently aversive, as compared to the preferred state of relief where pain is absent. The information conveyed by these sensations motivate the behaviour of escaping danger and seeking safety respectively. Together, these distinct aspects of pain and relief serve to maximise the chance of survival.

In this thesis, pain and its relief are studied from a motivational perspective. Specifically, they are explored in the context of learning: how behaviour and perception are modulated as the association between environmental signals and pain/relief outcomes changes. The purpose of this approach is twofold: firstly, the subjective internal experiences of pain and relief can be represented as learning problems, where hypotheses can be tested objectively with measurable behavioural and neural responses during learning. Secondly, learning processes can be formalised using learning theory and mathematical models, producing quantifiable predictors to allow the search for specific behavioural and neural correlates. Understanding the ways humans learn and adapt to pain and relief can provide insights into the development of pathological states such as chronic pain. It can also help to identify potential targets for pain therapeutics.

This chapter provides a concise overview of the systems neuroscience of pain and relief in relation to learning. A general introduction to pain, relief, and their motivational nature is given in section 1.1. Section 1.2 introduces reinforcement learning theory and its formalised models, with neurobiological evidence from previous pain and relief learning studies summarised in section 1.3. Section 1.4 briefly discusses clinical implications of pain/relief learning and the final section 1.5 outlines the remainder of the thesis.

1.1 Pain, relief, and motivation

Pain and pain relief as biologically relevant signals *motivate* behavioural changes to maintain bodily integrity (Navratilova and Porreca, 2014; Vlaeyen, 2015). These behaviours may vary in their acquisition and expression, often depending on the knowledge of how pain and relief are associated with the environment. In turn, these associations and their resulting motivational states can affect the hedonic experience of pain and relief, forming a bidirectional relationship between perception and motivation (Fields, 2000; Wiech and Tracey, 2013).

Pain and its relief

Pain is an unpleasant sensory and emotional experience associated with, or described in terms of, actual or potential tissue damage (IASP, 1994). As a subjective experience, pain is multidimensional: its *sensory* dimension is characterised by the stimulus that activates nociceptors (i.e. sensory receptors that transduce and encode noxious mechanical, thermal, or chemical stimulus), often described in terms of stimulus intensity, location, source, and quality (Gold and Gebhart, 2010; IASP, 1994); the *affective* dimension of pain refers to its aversiveness, which is tightly linked to the motivations underlying behaviour adjustment to prevent further harm (Navratilova and Porreca, 2014; Wiech and Tracey, 2013); a *cognitive* dimension influences pain perception, through factors such as attention, expectation, and prior experience (Navratilova and Porreca, 2014; Wiech et al., 2008); and finally, a *social* dimension of pain also exists, whereby communications and support from others may impact an individual's pain experience (KD Craig, 2015). A recently proposed update (Williams and KD Craig, 2016) suggests the definition of pain should facilitate appreciation of these other dimensions:

> Pain is a distressing experience associated with actual or potential tissue damage with sensory, emotional, cognitive and social components.

Consider the example of joint pain from running: the location and quality of the pain may signal to the runner to adjust running distance or posture to avoid further injury. However, the actual injury may not produce obvious pain, for example, in the context where the runner is approaching the finishing line in an intense race. In other cases, the pain may persist even after the injury has healed, such as in chronic pain. This multidimensional nature highlights the complexity and subjectivity of pain.

Compared to pain, pain relief is relatively less well defined. Existing definitions of relief vary in wording from source to source, but there is one common characteristic: relief derives from situations where expected or previously experienced negative stimuli is stopped, reduced

or absent (Fig 1.1a) (Deutsch et al., 2015; Leknes et al., 2011). In the context of this thesis, 'relief' refers to the relief from noxious physical pain specifically, and the terms 'relief' and 'pain relief' are used interchangeably throughout the text.

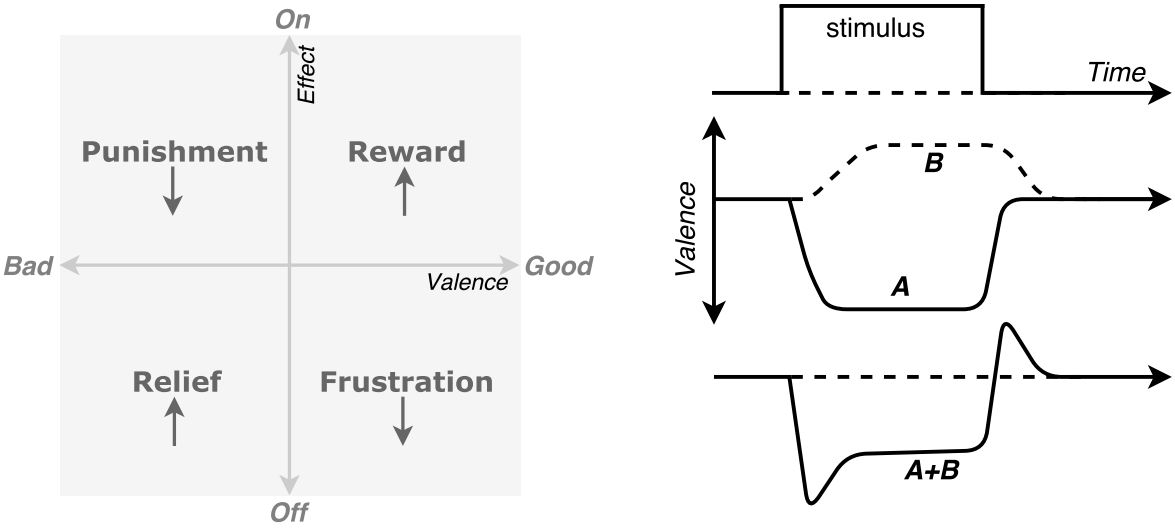
In clinical settings, pain relief in humans is often measured using the reduction in global pain ratings (Leknes et al., 2008). However, evidence suggests global ratings of pain relief represent more than changes in pain intensity – other factors such as satisfaction with change in pain, and previous maximum pain experienced, are shown to influence its perception (Jensen et al., 2005). Relief is also hypothesised to be a contributing factor behind the addictive effects of many analgesic drugs, and the formation of different phobic/avoidance behaviour (Deutsch et al., 2015; Leknes and Tracey, 2008). This suggests that pain relief is more than a simple sensory mapping of reduced pain, it possesses affective and motivational underpinnings that should be considered in a wider context.

Motivational theories

Motivation has two fundamental components: *action* and *learning* (Rescorla and Wagner, 1972). A *reinforcer* can change the strength and probability of a response to be emitted in the future (Skinner, 1938). A reinforcing event, such as pain or relief, is able to make an animal willing to expend energy to elicit actions that will increase or decrease access to it; it also supports learning of cues associated with these events so that the actions leading to them can be repeated earlier in the future (Gerber et al., 2014; Pavlov, 1927; Thorndike, 1911). Secondary reinforcers (e.g. cues) can elicit similar behavioural consequences after repeated pairing with primary reinforcers (e.g. pain).

Many theories in human motivation and learning explore how stimuli reinforce behaviour (Fig 1.1). The classic motivational assumption stipulates that people approach pleasure and avoid pain. Approach/avoidance theories suggest that drive states, such as 'fear', regulate different reactions to match the valence of reinforcing stimuli (Fig 1.1a) (Corr, 2013; JA Gray, 1987; Konorski, 1967; LeDoux, 2014; Mowrer, 1960). However, this dichotomy has been criticised for being oversimplified, prompting the development of the opponent process theory (Fig 1.1b) (Solomon, 1980; Solomon and Corbit, 1974). Opponent process theory suggests that a primary affective or hedonic process A elicited by any deviation from homoeostasis (e.g. pain) typically triggers an opponent process B with the opposite valence (e.g. relief) with a slower temporal latency. Behavioural response is determined by the magnitude and orientation of the net result of two processes $A+B$ – if process A stops, process B will dominate because of its slower rise and decay time. The major difference between the opponent process theory and the drive theories is that the temporal dynamics of the two opposing processes are more complex than the simple omission opponency of a fear state (Konorski, 1967; Mowrer, 1960).

Other theories consider motivation from an adaptive perspective. Regulation theories assume a feedback control loop where actions minimise discrepancies between actual behaviour and desired goals, and affective valence dynamically reflects the progression of goal pursuit (e.g. doing well in avoiding danger would bring relief) (Carver and Scheier, 1990; Higgins, 1997). Cognitive appraisal theory proposes the appraisal of controllability and certainty has impacts on both affect and behaviour in order to cope with a changing environment (Lazarus, 1991). Behavioural responses can also be controlled by habit learning, where a reinforcer serves to strengthen the stimulus-response (S-R) association through over-training, instead of encoding a behavioural goal (Balleine and Dickinson, 1998; Everitt and Robbins, 2005). Deutsch et al. (2015) has suggested that these theories can coexist to account for the different features of motivated behaviour, which may vary depending on the procedural differences in experimental paradigms.



(a) (b)

Fig. 1.1 (a) Affective valence and event effect. The diagram shows reward, punishment, frustration, and relief in the valence-effect coordinate system. Event effects can either be predicted (in the case of Pavlovian learning), or actual consequences of behaviour (instrumental learning). Arrows inside the coordinates indicate approach/withdrawal in Pavlovian sign tracking, or the increase/decrease in behavioural strength in instrumental learning following Thorndike's law of effect (adapted from Bouton, 2007; Gerber et al., 2014). (b) The opponent processes of affective dynamics. Opponent process theory assumes that the onset of a negative stimulus (e.g. pain, top panel) elicits a primary process *A* with negative valence, and an opponent process *B* with positive valence (middle panel). The summation of the two (bottom panel) reflects the positive effect of relief when the negative stimuli is removed (adapted from Solomon and Corbit, 1974).

Pavlovian conditioning

Two learning paradigms, *Pavlovian* and *instrumental conditioning*, have been used extensively to study motivation – specifically, how animals adapt to the environment through experience (Bouton, 2007). Pavlovian conditioning allows a previously neutral cue (*conditional stimulus*, CS) to gain incentive values through association with an outcome (*unconditional stimulus*, US). Food as a US can elicit an *unconditional response* (UR) such as drooling without learning. After a CS-US association is established, the CS on its own can elicit a *conditional response* (CR), allowing animals to act before US appears (Berridge, 2012; Bouton, 2007). While this response may not directly have any impact on the outcome, Pavlovian conditioning is important for animals to learn to predict and prepare for upcoming biologically significant events. Therefore, Pavlovian conditioning can be crudely equated to *prediction learning* (Dayan and Balleine, 2002).

Establishing the cue-outcome association also depends on *contingency*, which refers to the relative probability of an outcome occurring in the presence of a cue, as contrasted with the probability in the absence of the cue (Rescorla, 1968). The temporal and spatial proximity between cue-outcome pairing (*contiguity*) and the *consistency* of these events' co-occurrences have influence over the cue's control over behaviour in learning (R Miller and Escobar, 2004; Pavlov, 1927).

While Pavlovian conditioned responses generally reflect the valence of the CS, for example, approaching good CSs and withdrawing from bad CSs ('sign tracking', Fig 1.1a), CRs can have a wide range of behavioural manifestations. Specific responses may be elicited to prepare for food consumption, taste avoidance, reproduction, and territory defence; while more general, systemic responses, for example, freezing and other fight/flight physiological responses can also occur during fear conditioning (Bouton, 2007). Pavlovian responses to aversive events are relatively more diverse than those to appetitive ones, as more details of the stimulus are taken into account (DC Blanchard and RJ Blanchard, 1988). The dissociation of these different manifestations of conditioned responses and the representations of their learning processes are investigated within the context of Pavlovian pain conditioning in Chapter 3.

Instrumental conditioning

During instrumental conditioning, animals learn associations between actions and outcomes in order to behave in ways that improve their chances of survival (Bouton, 2007). Unlike Pavlovian conditioning, instrumental responses can have a direct impact on the probability of outcome occurrence, which permits animals control of the environment (Skinner, 1938). Instrumental conditioning is therefore considered to be *active learning*, as opposed to *passive*

learning in Pavlovian conditioning (Deutsch et al., 2015). Since repeating learned actions in similar situations in the future helps to accumulate reward or avoid punishment, instrumental conditioning can often be equated to *action learning* (Dayan and Balleine, 2002; Thorndike, 1911, 1932).

Instrumental actions can either be *goal-directed* or *habitual* (Balleine and O'Doherty, 2010). Goal-directed actions involve planning ahead, for example, simulating all possible consequences from all action sequences in an exhaustive tree search for action evaluation. Habitual actions, on the other hand, rely on the past and use previous experience to guide behaviour (Huys, 2007). These two modes of action evaluation can be distinguished using *outcome devaluation* and *contingency degradation*. The goal-directed system will exhibit reduced sensitivity to a particular reward (e.g. sucrose), when the previously rewarding substance is devalued by association with a negative outcome (e.g. illness). Conversely, an association initially learned by the habit system will not be sensitive to such devaluation as it does not involve representation of the specific outcome. Similarly, degrading the contingency between response and reinforcer by including unpaired outcomes in training also promotes habitual actions (Balleine and Dickinson, 1998; Dickinson and Balleine, 2002; Yin et al., 2004).

The above examples showed that an outcome and the actions leading to it can be valued differently depending on changing motivational state, which suggests the process of decision-making can be framed as a value comparison problem. Subjective *values* are assigned to events (e.g. stimuli, actions, outcomes), where they can be compared on an ordinal scale of preference to produce a decision on action, similar to using a 'common currency' for evaluation (McNamee, 2015; Montague and Berns, 2002). This concept has been used extensively in machine learning (Sutton and Barto, 1998) and psychology (Schultz, 2006). Since value allows direct comparison between events of different modalities, and is able to capture the temporal dynamics of motivational shifts, it is hypothesised to have biological relevance in decision-making and learning. The formal representation of value is discussed in section 1.2.

Predictability and controllability

The inherently probabilistic relationships between cues, actions, and outcomes introduce *uncertainty* in an environment, making it difficult to prepare properly for future events (Grupe and Nitschke, 2013). Learning paradigms have been used to study pain and relief motivation by manipulating the *predictability* and *controllability* of pain or relief outcomes, usually through changing the paradigm contingency. Predictability concerns the estimated likelihood of stimuli predicting outcome occurrences, while controllability refers to the perception of responses changing outcome probabilities – both are relevant to the estimation of uncertainty, and can

be distinguished in different contexts of learning (Grupe and Nitschke, 2013; Maier and LR Watkins, 2005).

Pavlovian conditioning mainly focuses on learning about stimuli that *predict* danger and safety. CS+, a stimulus that is repeatedly paired with experimental pain to signal impending aversive events, can act as a cue for escape or avoidance. 'Safety signal' CSs on the other hand can either be associated with the omission of aversive stimuli (e.g. in backward conditioning, a safety CS+ appears after a painful US, Andreatta et al., 2013), or the absence of onset of aversive stimuli (e.g. unpaired CS− as conditioned inhibitor in discrimination training, Fernando et al., 2013). Instrumental conditioning paradigms usually investigate the effects of operant behaviour, where instrumental responses during CS presentation directly *control* escape or avoidance of the US. Paradigm variations include active avoidance, where performing a certain response (e.g. choices in humans, Kim et al., 2006; conditioned place preference in rodents, Navratilova et al., 2012, 2015) results in the omission of the aversive US; passive avoidance, where withholding a specific response leads to US omission (e.g. go/no go actions in humans, Huys et al., 2011; Levita et al., 2012); and active escape from negative tonic stimuli (e.g. self administration of analgesic drugs in rats, Navratilova and Porreca, 2014).

Predictability and controllability of the environment not only affect motivation, but also hedonic experience. The importance of controllability was first demonstrated experimentally through studies on 'learned helplessness', where exposure to uncontrollable shocks during aversive Pavlovian conditioning was shown to impair later instrumental learning of escape, when comparing subjects exposed to escapable and inescapable shocks (Maier and Seligman, 1976; Overmier and Seligman, 1967). Apart from poor escape behaviour, uncontrollable shocks also produced exaggerated fear conditioned responses, reduced appetitive learning, and other behavioural signs of depression and anxiety (Maier and LR Watkins, 2005). The detection and learning of control over aversive events have now been found to be important to overcome the passive failure to escape (Maier and Seligman, 2016). Various studies have manipulated predictability and controllability to show their modulatory effects on pain and relief perception, including perceived control over analgesia (Salomons et al., 2004; Wiech et al., 2006, 2014c), cue-evoked expectation (Atlas et al., 2010), and uncertainty over the probability of pain delivery (Yoshida et al., 2013). In Chapter 4, relief motivation is explored in the context of learning to escape ongoing tonic pain, where we investigated the impact of relief controllability and predictability on pain behaviour and perception by comparing active and passive relief learning in instrumental and Pavlovian conditioning respectively.

Inference and measurement

Inferring motivation underlying behaviour depends largely on the experimental design and research question. In humans, subjective ratings and choices are commonly used to assess pain and relief experience, where experimental variables – such as stimulus intensity and outcome probability – are manipulated to induce motivational change (Leknes et al., 2013, 2008). For non-verbal animals, reinforced behaviour such as avoidance, escape, approach, are attributed to preference and motivation, in lieu of explicit self-reports (Navratilova et al., 2012, 2015).

A combination of overt behaviour, physiological parameters, and neural substrates can be measured to characterise inferred learning. Conditioned responses, either instrumental (e.g. choice preferences and reaction times, Kim et al., 2006; Prevost et al., 2013) or Pavlovian (e.g. startle modulation, facial musculature activations, autonomic/physiological signals, Andreatta et al., 2013; Lang et al., 1990), can reflect changes in the conditioned valence and motivational orientation of both CSs and USs following learning. Neuroimaging techniques, such as functional magnetic resonance imaging (fMRI), can provide insight into the neural mechanisms and circuitry underlying motivation at a systems level, an approach which has been used extensively in pain and decision-making research (Apkarian, 2011; O'Doherty et al., 2007). These methods are described in more detail in Chapter 2.

To summarise, pain and relief motivate action and learning, and motivational changes in turn have impacts on their hedonic experience. Associative learning paradigms have been used to study this reciprocal relationship, with the help of behavioural, physiological, and neuroimaging data for motivational inference. Crucially, the concept of value is able to formalise these experimental paradigms to allow a computational representation of learning.

1.2 Learning through reinforcement

When motivated to change the environment, learned knowledge from previous experience is essential to guide actions. The learning process may include components such as making associations between cue and reinforcers, assigning values to actions, constructing control policies, and prioritising actions. This section introduces reinforcement learning, a computational framework that formally captures the mechanical procedures of information processing during learning, and shows how value and uncertainty abstraction guide decision-making.

Reinforcement learning models

Reinforcement learning (RL) studies how agents learn to maximise reward and minimise punishment (Maia, 2009; Sutton and Barto, 1998). Actions are *reinforced* when they produce higher rewards, or lower punishment, than previously expected, meaning that the same actions are likely to occur again in similar future situations. The opposite happens to actions that produce worse than expected outcomes, leading to their diminished occurrences. It is believed that animals learn by comparing expected and experienced outcomes, and use errors from their predictions to guide future actions (Rescorla and Wagner, 1972). Hence RL involves using this *prediction error* to solve an optimisation problem, either to learn predictions of impending reward/punishment delivery, or to learn to adapt course of actions to achieve maximum reward and/or minimum punishment (Dayan and Balleine, 2002; Sutton and Barto, 1998).

Learning to predict reward/punishment outcomes and control actions to achieve such outcomes belongs to the Pavlovian and instrumental conditioning frameworks respectively (Dayan et al., 2000; O'Doherty et al., 2007; Seymour et al., 2004; AJ Yu and Dayan, 2005). RL uses value abstraction as a formal computational approach to the problems of action control in instrumental learning and reward/punishment predictions in Pavlovian learning. Different RL variations can accommodate the distinction between habit and goal-directed learning in instrumental conditioning. In addition, different types of uncertainty during learning can be incorporated to explain behavioural and perceptual changes. Therefore, RL is an essential framework that formalises conditioning paradigms computationally.

Basic elements of RL

The fundamental elements of RL include:

- State S – set of states referring a combination of external (e.g. context) and internal (e.g. physiological) variables to an agent
- Action A_s – set of actions available in state $s \in S$ that moves an agent to a new state s'
- Reinforcement R – set of reward/punishment the agent may receive in new state s'

It is the agent's goal to learn a control policy (π, a mapping from states to actions) that maximises expected long-term rewards received. This involves discounting future reinforcement in favour of those more proximal to the present, such as with exponential temporal discounting:

$$V(t) = E\left[r_t + \gamma^1 r_{t+1} + \gamma^2 r_{t+2} + ...\right] \tag{1.1}$$

$$V(t) = E\left[\sum_i \gamma^i r_{t+i}\right] \tag{1.2}$$

where t represents time, r_t is the reinforcement at time t, and γ is the discount factor where $\gamma \in [0,1]$. This gives the value $V(t)$ at time t, taking into account the current and future reinforcements (Maia, 2009; McNamee, 2015).

However, in most cases, the reinforcement R has different values depending on contexts and internal states, and the environment in general is more complex, with transitions between states being probabilistic instead of deterministic. Formally, a *Markov decision process* (MDP) defines such a probabilistic sequential decision-making environment. In addition to the elements state S, action A_s, and discount factor γ mentioned above, MDP also has a state transition probability function $T(s,a,s')$ that describes the probability $p(s'|s,a)$ of arriving in new state s' after performing action a in state s, and a reward function $R(s,a,s')$ that describes the reinforcement the agent receives after the state transition of $T(s,a,s')$. Very often in neuroscience experiments, contingencies are designed to be deterministic so that $T(s,a)$ and $R(s,a)$ are sufficient for description, given that s and a unequivocally determines s'. The resulting MDP is also called a deterministic MDP (Maia, 2009; Sutton and Barto, 1998).

Temporal difference (TD) learning

A credit-assignment problem exists within the MDP framework: when a sequence of actions results in an outcome, which one of the actions should get credit for the outcome? The consequences of an action can be spread over time, therefore it is not sufficient to know only the immediate reinforcement, but the expected sum of future reinforcement from that state – the state value. This bypasses the need to look infinite steps ahead for delayed consequences of an action. The state-value function is given by the Bellman equation (Bellman, 1957):

$$V^\pi(s_t) = E\left[\sum_{i=0,\ldots} \gamma^i r_{t+i} \mid s_t = s\right] \tag{1.3}$$

$$V^\pi(s_t) = E\left[r_t + \sum_{i=1,\ldots} \gamma^i r_{t+i} \mid s_{t+1} = s'\right] \tag{1.4}$$

$$V^\pi(s) = E[r_s] + \gamma V^\pi(s') \tag{1.5}$$

The above derivation lays the foundation for *temporal difference* (TD) learning, which is defined as:

$$\hat{V}^\pi(s) \leftarrow \hat{V}^\pi(s) + \alpha\delta \tag{1.6}$$

$$\delta = r_s + \gamma \hat{V}^{\pi}(s') - \hat{V}^{\pi}(s) \tag{1.7}$$

where $\hat{V}^{\pi}$ represents the estimate of state value under current policy π, and δ is called the prediction error, representing the difference between the *estimated* outcome for a particular state $\hat{V}^{\pi}(s)$ and the *actual* outcome received plus the discounted outcome from the next state $r_s + \gamma \hat{V}^{\pi}(s')$.

A positive δ indicates a better than expected outcome, while a negative one signals the opposite. The prediction error can also be regarded as the reinforcement r_s, plus the value difference between two states $\gamma \hat{V}^{\pi}(s') - \hat{V}^{\pi}(s)$, hence if the state value difference is not zero, updating can still occur despite having no reinforcement (e.g. in extinction where $r = 0$), and over time as learning progresses, the error will become smaller as prediction becomes more accurate. When $\delta = 0$, the state value difference also reaches zero, and no more learning occurs since the reinforcement becomes fully predicted (Maia, 2009; Sutton and Barto, 1998).

TD learning and conditioning

In the Pavlovian case, the TD learning rule is often used for state-value learning, summarised as the *Rescorla-Wagner* (RW) rule (Rescorla and Wagner, 1972):

$$V_{t+1}(s) = V_t(s) + \alpha(r_t - V_t(s)) \tag{1.8}$$

where α represents the learning rate ($\alpha \in [0,1]$), V as values and r as reinforcement outcomes.

Generalising to instrumental conditioning, a Q value represents the value of an action in a particular state, and it can be learned in two different ways. The first is *Q learning*, also known as an off-policy TD control algorithm, where the decision error is determined by the Q value of the better option instead of the one actually chosen. In this way, learning is independent of the actual choice, and exploring suboptimal alternatives may speed up convergence to optimal action (CJCH Watkins, 1989). Formally:

$$Q_{t+1}(s,a) = Q_t(s,a) + \alpha(r_t + \gamma \max_a Q(s_{t+1},a) - Q_t(s,a)) \tag{1.9}$$

where $Q(s,a)$ represents approximated optimal action-value of action a in state s.

The alternative, *SARSA* (state-action-reward-state-action), is an on-policy TD control algorithm, where the action value $Q^{\pi}(s,a)$ for the current behaviour policy π is estimated. The Q value here comes from a prediction error calculated using the actual chosen action (Niv et al., 2006; Sutton and Barto, 1998).

$$Q_{t+1}(s,a) = Q_t(s,a) + \alpha(r_t - Q_t(s,a)) \tag{1.10}$$

Action selection aims to decide on an action that maximises both immediate and future rewards. Always selecting the action with the greatest value is referred to as a *greedy* strategy. However, constantly exploiting a single action results in a lack of exploration over other possible options, potentially leading to diminished future rewards. The ε-*greedy* action selection rule exploits the highest value action most of the time, but explores randomly with a probability of ε, in order to balance the exploit-explore trade-off (Sutton and Barto, 1998).

The *softmax* rule ranks actions so that higher value actions have higher probability of being chosen, unlike the random choice in ε-greedy. A softmax distribution is often used to calculate the probability of choosing an action a in state s (Sutton and Barto, 1998):

$$p(a|s) = \frac{exp(Q_t(s,a)/\tau)}{\sum_{b \in A_s} exp(Q_t(s,b)/\tau)} \tag{1.11}$$

where τ is a temperature parameter governing the competition between actions. The policy $\pi(s,a)$ can be represented using the preference $p(s,a)$, updated similarly to state-values using prediction error, with the same softmax rule (Maia, 2009).

Value function and action selection rules are also referred to as *critic* and *actor*, which divides RL algorithms into three groups: actor-only, critic-only, and actor-critic (Grondman et al., 2012). The actor-critic method tries to simultaneously find the best policy and estimate the state-value for the current policy (Dayan and Balleine, 2002). The critic learns and stores the state-values using prediction errors (Equations 1.6 and 1.7), providing the actor with low-variance knowledge of the overall performance. In turn, the actor selects action based on the expected values (e.g. Equation 1.11), updating a parameterised policy to provide a spectrum of likely actions. Actor-critic converges faster and reduces the cost for action-space optimisation from a machine learning perspective (Grondman et al., 2012; Maia, 2009).

Model-free and model-based RL

The TD learning-based algorithms introduced above are examples of the *model-free* RL method. 'Caching' is the foundation of this method, where state or action values are the scalar summary of long-run values in time. The optimal policy is learned directly from these cached values, where accumulated past rewards guide future actions (Daw et al., 2005). Biologically, midbrain dopaminergic neurons exhibit prediction error-like spiking patterns, supporting the feasibility of model-free methods (Schultz, 1998).

In contrast, *model-based* methods attempt to learn the MDP first, i.e. by constructing a cognitive model of the environment that predicts long-run outcomes, in order to find an optimal policy. This usually involves a mental simulation of various possible future outcomes, such as in a tree-search, as well as computing the reinforcement and state transition function (Daw

et al., 2005; Maia, 2009). Predicted values from model-based methods have been found to correlate with prefrontal cortical activities in human decision-making experiments (Daw et al., 2011; Gläscher et al., 2010).

Model-free RL is often associated with habit learning of stimulus-response (S-R) associations, while model-based RL is associated with goal-directed learning of more complex action-outcome (A-O) or stimulus-action-outcome (S-A-O) contingencies (Maia, 2009). Although functionally and computationally separable, the two systems interact, e.g. through Pavlovian-instrumental transfer (where Pavlovian cues strengthen separately trained instrumental actions that lead to the same reward), which has been shown to manifest behaviourally (Holmes et al., 2010; Huys et al., 2011). In Chapter 5, active relief learning is extended from a static to a dynamic environment, where the relief probabilities of potential actions change over time, and model-free and model-based RL models are compared in order to determine whether learning changes as a result of this environmental shift.

Uncertainty

With estimation comes imprecision. Uncertainty forms an inherent part of decision-making, generated within all processing levels ranging from the internal state of the agent to the environment it interacts with. The classification of uncertainty varies according to the criteria chosen, which can be its main sources (sensory, state, rule, or outcome uncertainty, Bach and Dolan, 2012), probabilistic hierarchical levels (uncertainty of environment volatility, observed outcome probability, or outcome category, de Berker et al., 2016; Mathys et al., 2011), reliability (expected or unexpected uncertainty, arises from previously known unreliability of predictive relationships within an environment, AJ Yu and Dayan, 2005), or in the decision-making field, the distinction between ambiguity and risk (not knowing whether a coin is fair or biased, as opposed to knowing the coin is fair but cannot predict the outcome of a particular coin toss, also referred to as reducible or irreducible uncertainty, Kobayashi and Hsu, 2017).

Depending on the design of a learning experiment, uncertainty can be computationally represented using either *entropy* (e.g. when the uncertain variable is nominal and varies across discrete states), distribution *variance* (e.g. when the uncertain variable can be described with a probability distribution), or prediction *surprise* (e.g. when the uncertain variable is subjected to bootstrap updating in RL) (Bach and Dolan, 2012; Le Pelley, 2004; Mathys et al., 2011). It is likely that uncertainty estimates are used to guide behaviour, possibly through directing attention, so that the learning of predictive associations in the environment can be boosted, while irrelevant information is suppressed. *Associability* estimates the uncertainty of associative predictions during learning and adjusting learning rate accordingly – high uncertainty during

learning commands increased attention, so that learning can occur more rapidly (Dayan et al., 2000; Le Pelley, 2004; Pearce and Hall, 1980; AJ Yu and Dayan, 2005).

In summary, value abstraction provides the basis for RL frameworks, in which decision-making processes can be represented with value function updating, action selection rules, and uncertainty estimation. These computational measures can be used to better dissect learning associated with pain and relief, by identifying the behavioural and neural correlates of these internal model qualities.

1.3 The neurobiological basis of pain/relief motivation

Combining learning paradigms and the RL framework allows the computational mapping from pain/relief stimuli to responses in the brain (referred to as *neural encoding*, as opposed to its reverse of *neural decoding*) (Dayan and Abbott, 2001). Localising this functional mapping neurally and behaviourally is important for understanding how information about pain and relief is represented and processed. Neurobiological evidence related to pain/relief motivation identified in previous studies, with emphasis on the central nervous system, are outlined in this section.

The afferent nociceptive pathways

The afferent nociceptive pathways are involved in the reception, processing and transmission of nociceptive input that signals potential tissue damage. Peripheral nociceptors are widely distributed in the body, both externally (e.g. the skin, mucosa) and internally (e.g. muscles, joints, tendons, connective tissues of visceral organs) (Almeida et al., 2004). Nociceptors are activated when mechanical, thermal, or chemical stimuli exceed their threshold for response. However, they can also be *sensitised* as a result of tissue insult and inflammation, which causes their excitatory threshold to reduce and response magnitude to increase, and in some cases, eliciting spontaneous activation without stimulus (Gold and Gebhart, 2010). The action potentials generated from these receptors are conducted through the nociceptive afferent nerves – mainly the larger, thinly myelinated, fast-conducting A-delta fibres, or the smaller, unmyelinated, slow C fibres. A-Delta fibres mostly correspond to the acute, first-phase 'sharp' pain from high intensity mechanical and heat/cold stimuli, while C fibres are responsible for the slower, second-wave 'dull' ache from most stimulus types as well as inflammation (Millan, 1999). These primary afferents conducting nociceptive information reach the dorsal horn of the spinal cord, before ascending to higher cerebral structures in the brain through different spinal pathways, which include the spinothalamic, spinoparabrachial-amygdaloid,

spinoreticulo-thalamic, and spinomesencephalic (periaqueductal grey, PAG) tracts (Fig 1.2) (Almeida et al., 2004; Bushnell et al., 2013).

The primary processing of nociceptive information and the organisation of these afferent tracts in the spinal cord are outside the scope of this thesis, and have been reviewed in details elsewhere (Almeida et al., 2004; Millan, 1999; Schaible and Grubb, 1993).

The 'pain matrix'

A network of brain regions responds consistently to transient or acute painful stimuli across human neuroimaging studies, according to meta-analysis (Apkarian et al., 2005). Often referred to as the 'pain matrix', these regions involve both cortical and subcortical structures encompassing sensory, limbic, and associative areas (Bushnell et al., 2013; Tracey and Mantyh, 2007). The common areas revealed in pain studies include the primary and secondary somatosensory cortices (SI and SII), insula, thalamus, anterior cingulate cortex (ACC), prefrontal cortex (PFC), and periaqueductal grey (PAG); while more variable activation has been observed in the striatum (dorsal and ventral), orbitofrontal cortex (OFC), the amygdala and cerebellum (Fields, 2004; Geha and Waxman, 2016; Navratilova and Porreca, 2014; Wiech and Tracey, 2013). Recently, a 'neurologic signature of pain' derived from the weighted fMRI activity pattern of many regions from the pain matrix has been shown to predict pain intensity in test subjects, as well as differentiating thermal pain from innocuous warm stimulus and social rejection pain (Wager et al., 2013). However, the assumption that the pain matrix mediates the conscious experience of pain is challenged by other evidence – for example, activations of similar brain regions have been identified using other non-nociceptive sensory modalities (Mouraux et al., 2011), in patients who have congenital insensitivity to pain (Salomons et al., 2016), and paradoxically, in pain relief studies (Andreatta et al., 2012; Leknes et al., 2011; Navratilova et al., 2012). These findings lead to the view that the pain matrix could be involved in the processing of multimodal, salient sensory information, or perhaps reflects the results of pain-related learning (Geha and Waxman, 2016). The lack of a specialised 'pain cortex' in the brain also suggests that simple stimulus mapping is unlikely to sufficiently inform the central processing of pain.

Impact of chronic pain

Neuroimaging studies of chronic pain patients have revealed distinct brain activity patterns from transient pain in the clinical population (Apkarian et al., 2005). Pain is considered to be chronic when it persists past the healing phase following injury, usually 3-6 months after its initial onset depending on the injured site (Apkarian et al., 2009; IASP, 1994). Brain

correlates of acute/subacute (6-12 weeks) pain shift towards the amygdala and medial PFC as the pain becomes chronic (Hashmi et al., 2013). In addition, prefrontal-striatal connectivity has been found to predict pain persistence in a longitudinal study (Baliki et al., 2012), and significant anatomical alterations and disruption in functional connectivity have been identified in the prefrontal and limbic regions in chronic pain patients (Fritz et al., 2016; R Yu et al., 2014). Many of the changes observed in chronic pain involve the corticolimbic circuitry responsible for encoding motivational, affective, and cognitive aspects of pain and relief (Porreca and Navratilova, 2017; Wiech and Tracey, 2013). Brain regions including the striatum, amygdala, ACC and PFC, have been previously found relevant to the encoding and evaluation of motivational and affective information in decision-making, particularly for reward processing. Hence they are also collectively known as the reward motivation and decision mesocorticolimbic pathways (*motivation/decision pathways* for short, Fig 1.2, Navratilova and Porreca, 2014; Rushworth et al., 2011).

The motivation/decision pathways

The motivation/decision pathways have been linked to pain and relief motivation and processing. Previous evidence suggests the striatum contributes directly to decision-making involving reward and punishment (Salamone, 1994). The ventral striatum, which include the nucleus accumbens (NAc), ventral caudate and putamen, is often considered related to appetitive/aversive motivations and behaviour. The NAc receives projections from dopaminergic neurons in the ventral tegmental area (VTA), as well as inputs from the PFC, ACC, amygdala, hippocampus, and spinal dorsal horn neurons. This makes it a potential site for motivational and affective signals to translate into adaptive behaviour by ways of dopamine neurotransmission (Levita et al., 2009; Navratilova and Porreca, 2014). The dorsal striatum is suggested to be involved in action selection and initiation, integrating sensorimotor, cognitive, and motivation information from other subdivisions of the striatum (Balleine et al., 2007).

NAc activation has been identified in pain relief neuroimaging studies (Andreatta et al., 2012; Leknes et al., 2011; Navratilova et al., 2012), which led to the interpretation that relief from pain can be rewarding. While mainly associated with reward processing, the striatum also encodes aversive prediction signals (Schultz, 2013; Seymour et al., 2004), as well as general motivational relevance or stimulus salience (Becerra and Borsook, 2008; Deutsch et al., 2015; Levita et al., 2012). Furthermore, relief related cues have previously been found no more reinforcing than a control stimulus (Fernando et al., 2013), or even perceived to be aversive when these cues signal passive prevention relief (Andreatta et al., 2013; Mallan and Lipp, 2007). Based on existing evidence, pain relief therefore does not simply appear to be equal to reward or the direct opposite of pain.

The ACC has been implicated in the encoding of affective features of pain and relief (Bushnell et al., 2013; Vogt, 2005). Heavily connected with premotor, prefrontal, and limbic regions, the ACC also expresses high levels of opioid receptors and endogenous opioid activity that mediate pain suppression (Zubieta et al., 2005). Dissociation between conditioned place aversion and pain-evoked startle responses has been shown in ACC lesion studies, stressing its role in pain related affect evaluation (Johansen et al., 2001). In human neuroimaging studies, the ACC has been found to represent pain unpleasantness (Roy et al., 2009), showing reduced activation during relief compared to a control pain context (Leknes et al., 2013), and modulating responses to reward anticipation in the presence of pain (Wiech and Tracey, 2013), which acts to integrate the 'cost' of pain into decision-making. Despite consistent appearances in pain and pain modulation studies, the ACC exhibits bivalent activations to pain and relief (Becerra et al., 2013; Etkin et al., 2011), as well as other survival-critical processes including attention, foraging, motor and emotional functions, which cautions against the simplistic assumption of the ACC's selectivity for pain (Lieberman et al., 2016; Lieberman and Eisenberger, 2015; Wager et al., 2016).

Closely connected to the adjacent ACC, the PFC and OFC have important roles in value encoding that guides decision-making, as reviewed in Kringelbach (2005) and Rushworth et al. (2011). These regions are also important for action-outcome contingency detection, instrumental behaviour selection, and goal valuation in reward and avoidance learning (Cardinal et al., 2002; Kim et al., 2006). In addition, the medial PFC plays an important role in threat controllability, as its inhibitive control of the dorsal raphe nucleus is necessary to protect against passive helplessness (Amat et al., 2005; Maier and Seligman, 2016). The lateral PFC has also been implicated in analgesic effects related to perceived pain controllability (Salomons et al., 2007; Wiech et al., 2006; Zubieta et al., 2005).

The insula is believed to be involved in many somatosensory and emotional processes (AD Craig, 2009; Geuter et al., 2017; Wager and Barrett, 2017). The anterior/posterior divisions of insula appear to reflect different aspects of pain processing. The anterior insula has strong connection with prefrontal regions and is shown to encode pain predictions and evaluations (Geuter et al., 2017; Ploner et al., 2010). The posterior insula receives direct spinothalamic input and is heavily connected to SI/SII, and appears to encode pain perception such as intensity (AD Craig, 2009; Wiech et al., 2014a).

Closely connected to the ventral striatum and PFC/OFC, the amygdala also contributes to emotion, motivation, learning and attention (Murray, 2007). Previously associated mainly with danger/fear learning (Andreatta et al., 2012; McHugh et al., 2014) and negative affect processing (Balleine and Killcross, 2006), current evidence supports the amygdala's role in encoding reward, punishment, and relief (e.g. safety stimuli, Genud-Gabai et al., 2013;

avoidance cues, Schlund et al., 2010; or relief signals, Sangha et al., 2013). The amygdala has a central role in Pavlovian conditioning, as its basolateral component (BLA) is necessary for a CS to access the motivational values of its predicted US, influencing both instrumental behaviour and autonomic conditioned responses which are normally controlled by its central nucleus (CeA) (Cardinal et al., 2002). In particular, the close connection and interaction between amygdala and the PFC are important for animals to make advantageous choices when facing multiple competing cues, possibly through amygdala updating and PFC/OFC storing of the expected values of reward outcomes (Holland and Gallagher, 2004; Murray, 2007).

Apart from receiving nociceptive inputs, the cerebellum has been suggested to have a role in nociceptive processing (Moulton et al., 2010). fMRI activations in the cerebellum have been consistently identified with both experimental and pathological pain, as well as in placebo and learning studies (Ploghaus et al., 1999; Seymour et al., 2004; Wager et al., 2004). Furthermore, the primary motor cortex and the DLPFC are parts of the cerebrocerebellar loop, suggesting the cerebellum may be involved in sensorimotor integration in addition to pain sensory processing (Kelly and Strick, 2003; Moulton et al., 2010).

The descending modulatory pathways

The brain regions introduced above are not only involved in the motivation/decision pathways, but also the *descending pain modulatory pathways* – a set of systems consisting of cortical and brainstem projections to the spinal cord to regulate or inhibit ascending nociceptive information (Tracey and Mantyh, 2007). These systems involve multiple neural pathways and neurotransmitters, and they jointly regulate the affect, perception and behaviour of pain. Amongst them, the endogenous opioid system occupies a central role, and has been directly implicated in placebo analgesia (the pain reduction that can be attributed to treatment contexts, referring to discrete elements or their combinations that signify the occurrence of pain relief) (Levine et al., 1978; Wager and Atlas, 2015). Important brainstem structures in the opioidergic pathway include the periaqueductal grey (PAG) and rostroventral medulla (RVM), both of which govern the route to/from the spinal cord, in addition to receiving inputs from the VMPFC, VLPFC, amygdala, NAc and hypothalamus (Fields, 2004; Linnman et al., 2012; Wager and Atlas, 2015). Other neurotransmitters, including dopamine, serotonin, cholecystokinin and oxytocin, have also been implicated in placebo analgesia (Benedetti et al., 2005; Wiech and Tracey, 2013).

This complex prefrontal-subcortical circuitry appears to directly modulate pain perception, influencing the sensory aspect of pain via descending modulation of nociceptive processing in the dorsal horn, and the motivational/affective aspect of pain through interactions with the amygdala and striatum (Tracey and Mantyh, 2007; Wager and Atlas, 2015). Conversely,

alterations in pain perception have impacts on this neural network – placebo effects, emotional appraisal/regulation of pain, and the development of chronic pain, all appear to change the functional connectivity among the principal structures in these pathways (Bushnell et al., 2013; Etkin et al., 2011; Wager and Atlas, 2015).

A complete review of other relevant brain structures and pathways in the network is outside the scope of this thesis. For a more comprehensive review see (Bushnell et al., 2013; Etkin et al., 2011; Navratilova and Porreca, 2014; Tracey and Mantyh, 2007; Wager and Atlas, 2015; Wiech and Tracey, 2013). Critically, the converging pathways of pain/relief motivation and modulation further illustrates the importance of computational characterisation of these brain structures in order to understand their individual contribution and interactions.

Reinforcement learning in the brain

In more general decision-making studies, RL has been used to formalise information processing and to identify correlates of different learning processes in the brain. First demonstrated with reward conditioning in monkeys, dopamine neurons have been found to exhibit phasic activations that resemble reward prediction errors (δ) – with a CS predicting reward evoking dopamine neuron firing before the appearance of US, but a CS followed by the omission of US evoking lower than baseline firing at the time of the expected US (Schultz, 1998; Schultz et al., 1993). Typically, a better than predicted reward ($\delta > 0$) elicits dopamine activations, a fully predicted reward ($\delta = 0$) elicits no response, and a worse than predicted reward ($\delta < 0$) causes depressions. This pattern can be found in 60-75% of dopamine neurons across the brain, with even higher percentage in the midbrain (Schultz, 2013).

The neural correlates of RL prediction errors have also been observed in human neuroimaging studies (Garrison et al., 2013; Schultz and Dickinson, 2000). Brain regions with dense connection to dopaminergic neurons, especially the striatum and surrounding structures, have been shown to encode reward prediction errors (O'Doherty et al., 2004). Similar activities have also been observed in other areas within the motivation/decision network, including prefrontal regions OFC, medial PFC and ACC, as well as insula, amygdala, PAG, and cerebellum (Garrison et al., 2013; Geuter et al., 2017; Roy et al., 2014). However, some of these reported neural correlates might reflect their dopaminergic input, warranting the need for the more stringent axiomatic test to ensure the encoding of actual prediction errors (Maia, 2009; Niv and Schoenbaum, 2008). Neural signals in candidate regions cannot in principle encode prediction errors if they falsify one or more axioms. Specifically, the three testable axioms are: (i) consistent outcome ordering (i.e. response to reward is larger than that of no reward), (ii) consistent prediction ordering (response decreases as predicted probability of reward increases),

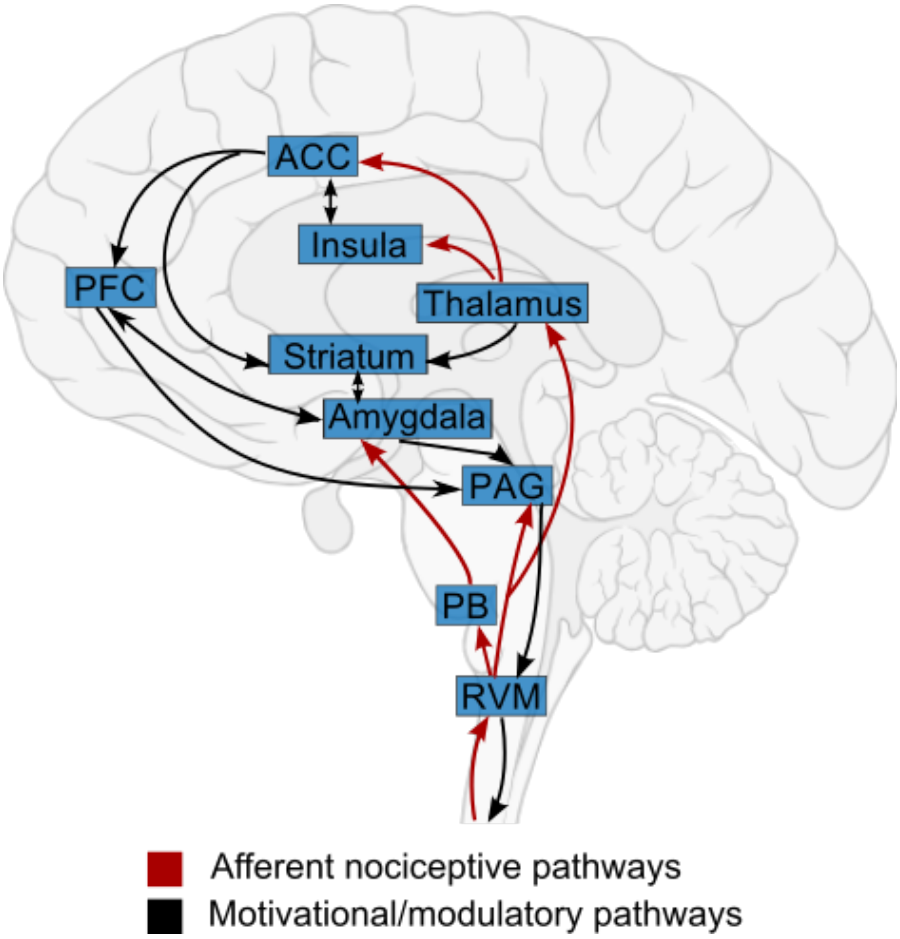


Fig. 1.2 The neural pathways involved in pain and relief information representation and processing span a wide range of cortical and subcortical structures. The afferent nociceptive information enters the brain from the spinal cord, through connections to the rostroventral medulla (RVM), parabrachial area (PB), periaqueductal grey (PAG), amygdala, and the thalamus, which is further projected to the anterior cingulate cortex (ACC) and insula (red arrows), as well as to the primary and secondary somatosensory cortices (SI/SII) and cerebellum (not shown). The motivation/decision pathways in the corticolimbic areas, involving the prefrontal cortex (PFC), ACC, amygdala, striatum, and insula, integrate motivationally salient information, e.g. pain and pain relief, and makes relevant action decisions. The descending modulatory systems share the many structures of the motivation/decision pathways, along with the PAG and RVM, this system contribute to affective, emotional and cognitive control of pain perception and decision-making (black arrows) (adapted from Bushnell et al., 2013; Navratilova and Porreca, 2014; Wager and Atlas, 2015).

and (iii) no surprise equivalent (response to fully predicted reward is equivalent to no reward) (Roy et al., 2014; Rutledge et al., 2010).

Apart from reward, aversive prediction errors derived from punishing reinforcers have been shown to evoke robust activations in the insula and habenula, and diminished activations in the ventral striatum and ACC (Garrison et al., 2013; Matsumoto and Hikosaka, 2007). Aversive prediction errors have also been linked with the serotonin system, with some theories proposing that serotonin and dopamine underpin the motivational opponency between the appetitive and aversive systems (Daw et al., 2002; Maia, 2009; Seymour et al., 2012). Recent experimental evidence with sub-second *in vivo* measurement of serotonin concentration has provided initial support for this theoretical model (Moran et al., 2018).

Variations of RL algorithms have been derived to account for the computations underlying different learning systems in the brain. Model-free approaches, especially the actor-critic model (section 1.2), have been associated with habit (S-R) learning, since the model-free 'caching' is not flexible enough to reflect immediate changes in action or state-values when outcomes are being manipulated (Daw et al., 2005; Maia, 2009). Imaging evidence has suggested a functional differentiation between the actor and critic components: the ventral striatum tracks prediction errors in both instrumental and yoked Pavlovian reward conditioning, while activity in the dorsal striatum correlates only with those from instrumental learning, where action evaluation and selection are required (O'Doherty et al., 2004). However, a recent meta-analysis has found that the ventral and dorsal striatum were activated similarly in both instrumental and Pavlovian conditions (Garrison et al., 2013), and other brain regions such as OFC and amygdala are also proposed as potential critics (Maia, 2009). The SARSA algorithm has been found to describe dopamine neuron firing in monkeys during a learning task, as compared to other more sophisticated actor-critic or Q learning models (G Morris et al., 2006; Niv et al., 2006). Model-based RL approaches have been associated with more complex goal-directed learning, where reward and state transitions are tracked, and have been linked to the evolutionarily more advanced prefrontal cortex (Daw et al., 2005; Gläscher et al., 2010). Findings from these studies support the coexistence of distinct model-free and model-based learning processes in the brain, with prediction reliability being a potential basis for arbitration between the two systems (Daw et al., 2005; Lee et al., 2014).

To conclude, brain regions associated with both the motivation/decision and descending pain modulatory pathways have been consistently implicated in pain and relief, as well as in reinforcement learning. A clearer theoretical and computational characterisation of pain and relief learning can increase the understanding of this particular neural network, and its role in pain motivation, perception, and potential therapeutics.

1.4 Clinical implications and therapeutic potentials

The increased engagement of the motivational and emotional circuits of the brain when transitioning from acute to chronic pain has long been suspected to contribute to pain 'chronification' (Apkarian, 2011). Chronic pain has an average prevalence of 20% in the adult population of Europe and North America, and the current medications available fail to meet the need of patients – with two thirds of the European respondents reporting unsatisfactory pain control (Andrew et al., 2014; Breivik et al., 2006). Therefore, understanding the mechanisms behind pain chronification, including the involvement of the motivation/modulatory neural network, and by extension, pain and relief motivation, may help to provide insights for novel therapeutics.

Numerous studies comparing chronic pain patients to healthy controls have revealed functional, anatomical, and molecular changes to the brain, with many of these changes observed in structures of the motivational circuit described above, detailed in the previous section (Apkarian et al., 2009). Accumulating evidence points to a bidirectional relationship between the morphological reorganisation of this circuit and pain chronification, possibly also the many affective and cognitive comorbidities of chronic pain. The computational characterisation of pain/relief learning in the brain could therefore help elucidate the mechanism behind this transition.

The challenge of objectively and reliably measuring pain experience is an important issue surrounding translational research and drug development. Since reflexive pain behaviour measured in rodents cannot readily capture human affective pain, motivational measures involving learning paradigms (e.g. conditioned place preference measured in rodents) have shown promise as an alternative, with rat pain models being used to validate analgesic outcomes in humans (Navratilova et al., 2013). Indeed, motivation/affective components may be useful supplements to fMRI-based decoders for human pain measurement ('neural pain signature', Wager et al., 2013), a technique currently restricted to experimentally induced pain but not clinical pain (Salomons et al., 2016). Being able to computationally characterise pain/relief also serves to deepen the understanding of pain-related behaviour and brain activities, so that their limitations as pain measures can be evaluated more clearly.

Existing treatments for patients with intractable chronic pain have targeted brain regions within the motivation/modulatory network – either through surgical lesions, internal stimulation such as deep brain stimulation (DBS), or external stimulation such as transcranial magnetic stimulation (TMS) (for a review see Zhang and Seymour, 2014). The latest methods, for example the non-invasive decoded neurofeedback based on fMRI measured activations or connectivity (Megumi et al., 2015; Shibata et al., 2011), and genetic-based control techniques such as optogenetics (Deisseroth, 2011), can provide controllable intervention with high temporal and spatial precision. In addition, they have the potential to form part of a brain-machine interface, 'closed-loop' pain control system, which aims to adjust intervention according to

the real-time pain-related brain activities – hence minimising the hazards of over-stimulation and the need for battery replacement (Zhang et al., 2013; Zhang and Seymour, 2014). Apart from identifying appropriate targets for therapeutic stimulation, studying how the brain learns and adapts during pain and relief can also reveal its impacts on the performance of these technology-based therapeutics. In Chapter 6, a proof of concept closed-loop system is built to control experimental pain delivery based on real-time decoded pain-evoked fMRI brain activity. This is the first attempt to assess the influence of such systems on the representation and subjective experience of pain in human.

In summary, characterising pain and relief motivation and understanding the role of the motivation/decision neural circuit involved can be useful for elucidating the mechanisms behind pain chronification and improving the design of technology-based pain therapeutics.

1.5 Thesis structure

This chapter introduced the idea of studying pain and relief from a motivational perspective, and reviewed theoretical, computational, and neurobiological evidence that support its importance. Specifically, the reinforcement learning computational framework emerges as a potential approach to bridge experimental learning paradigms with the resulting behavioural and neural responses. This also raises practical research questions regarding the nature of pain and relief experience:

- How do humans learn and adapt behaviourally and neurally in the presence of pain and relief, from a computational perspective?
- How does learning influence the subjective experience of pain and relief, as well as their representations in the brain?
- How does the understanding of pain/relief motivation inform the development of future technology-based pain therapeutics?

The remainder of this thesis aims to explore these questions with a series of human functional neuroimaging experiments. In Chapter 2, the methods for behavioural and neural data collection and analysis are briefly reviewed. The main body of the thesis consists of four experiments focusing on pain and relief learning. Chapter 3 uses a Pavlovian acute pain conditioning task to dissociate the different learning processes underlying the different conditioned responses from learning. Chapter 4 and 5 focus on relief learning through terminating tonic pain, investigating how this modulates the motivational and hedonic aspects of pain, and whether

learning differs for active and passive learning paradigms. Chapter 6 explores the feasibility and potential impacts of a brain-machine interface pain control system where real-time decoded information from the brain is used to control future experimental pain delivery. Finally, a summary of the contributions from these experiments is given in Chapter 7, with a discussion on the potential directions for future research.

Chapter 2

Methods

The main objective of computational neuroscience is to derive mathematical models that formalise the mechanical procedure of information manipulation and processing implemented through the collective activities of a population of neurons (Bermúdez, 2014). For a system like the brain, any mathematical models proposed need to strike a balance between biological plausibility and computational tractability. Specifically, these models need to conform with the electrical and chemical properties of neurons, and should be simple and efficient in dealing with the given information-processing tasks (Bermúdez, 2014; RC O'Reilly and Munakata, 2000).

The problem of pain and its relief can be approached from a motivational perspective. Using pain and/or relief as reinforcing outcomes in conditioning experiments allows the complex behaviour and subjective perception to be analysed in the context of learning. Hypothesised computational learning models then generate dynamic predictions that can be compared with actual neural and behavioural data acquired during these experiments. This comparison serves to assess the explanatory power of the proposed models for realistic biological processes, as well as discriminating between competing hypotheses (Cohen et al., 2017; Daw and Doya, 2006; O'Doherty et al., 2007).

In this chapter, an overview of the methods used to achieve the above goal and their relevant background are given. A variety of neural and behavioural responses in experiments can be recorded non-invasively in healthy human participants in the forms of functional neuroimaging and physiological measurements. Overviews of data collection, processing, and analysis are described in sections 2.1 and 2.2. Stimulation used to elicit pain and relief is reviewed in section 2.3. Finally, the procedures of computational model fitting and comparison are summarised in section 2.4.

2.1 Functional magnetic resonance imaging (fMRI)

Functional magnetic resonance imaging (fMRI) has become the most commonly used method for studying human brain function over the past two decades. It is proposed to measure indirectly the local excitatory neuronal activities. Through a series of signal processing and statistical analyses, raw fMRI data can be transformed into statistical maps showing responses in particular brain regions related to experimental stimuli, which are often manipulated to alter mental or perceptual functions (Penny et al., 2011). fMRI has become a prominent technique in cognitive neuroscience due to its demonstrated safety, non-invasiveness, and relatively high spatial resolution (Cohen et al., 2017; Poldrack et al., 2011).

Theoretical overview

The most common method of fMRI is measuring the *blood oxygenation level dependent* (BOLD) signal change, which reflects indirectly the changes in local neural activity through changing blood oxygenation, often as a result to external stimulus (Kwong et al., 1992; Ogawa et al., 1992; Ogawa et al., 1990). Determinants of the BOLD signal include: (i) the neuronal activity that responds to an external stimulus, (ii) the process that converts neuronal activity to haemodynamic responses ('neurovascular coupling'), (iii) the haemodynamic response driven by the increased local blood flow and oxygen metabolism, and (iv) the MR signal change detected by the MRI as a result of haemodynamic responses altering the magnetic field properties (Arthurs and Boniface, 2002; Buxton, 2013). Specifically, metabolically active tissues have increasing demand for oxygen, which increases local cerebral blood flow and therefore regional concentration of freshly oxygenated blood. As oxygenated and deoxygenated haemoglobin have different magnetic properties, the BOLD signal effectively uses this as an endogenous contrast to reflect tissue differences in concentration of these molecule types. The decreasing concentration of deoxygenated blood observed in local vasculature decreases paramagnetism and increases the MR signal strength observed locally (Huettel et al., 2014).

fMRI BOLD measures the haemodynamic activity associated with a population of neurons. Since BOLD relies on detecting blood oxygenation, its spatial resolution is limited by the lower density of microvasculature compared to neurons (Arthurs and Boniface, 2002). The typical size of a voxel (the building block of a 3D image) in fMRI studies is a cube with 2-3 mm edges. A voxel can span multiple cortical columns and vasculature elements, containing at least 10^6 neurons. Numerous studies have shown that the BOLD signals correlate strongly to local field potentials (LFPs) and evoked potentials, which represent population synaptic activity instead of neuronal spikes (Arthurs and Boniface, 2002; Goense et al., 2016; Logothetis et al., 2001). Comparing to other functional imaging methods such as positron emission tomography

(PET), fMRI has superior spatial resolution and allows activity localisation without invasive procedures or ionising radiations (Buxton, 2013).

As the cerebral blood flow response has a relatively slow rise and recovery time, the temporal resolution of BOLD fMRI typically spans 1-3 seconds, significantly slower than that of actual neural activity and hence many other electrophysiological modalities (Buxton, 2013). As the delay between sensory input and haemodynamic response is relatively constant, the BOLD response can be approximated as a linear time invariant system for events that are separated sufficiently in time. This property allows the use of statistical techniques such as *statistical parametric mapping* (SPM) to assess the significance of correlations between sensory input stimuli and haemodynamic responses measured with MRI, both spatially and temporally, which forms the basis of activation mapping studies in the human brain (Friston et al., 1994a; Friston et al., 1994b; Poldrack et al., 2011).

Experimental designs

To identify the brain region responding to a particular sensory stimulus, fMRI experiments are designed to assess the modulatory effects of a stimulus as it varies in time (Buxton, 2013). A typical *block design* alternates periods of time in which different stimuli are given – for example, a participant receives painful thermal stimuli for 30s, and then non-painful warm stimuli for 30s, repeated multiple times in *sessions* for statistical power. In comparison, an *event-related design* has different stimuli being presented interleaved in time, taking advantage of fMRI's relatively fine temporal precision. This type of design allows changing responses to the same stimulus to be captured, for example, to detect the effects of learning (Dale, 1999). In addition to determining the response to specific stimuli, fMRI can also be used to explore *functional connectivity* between various brain regions based on their temporal signal correlation, either with or without applying external stimuli (task- or resting-state functional connectivity) (Friston, 2011; JX O'Reilly et al., 2012). With increasing computational power and algorithm efficiency, *real-time fMRI* is able to acquire experimental images approximately in real-time, providing the option of incorporating advanced design features such as providing real-time feedback or exerting external controls based on the detected brain activations (deCharms, 2008).

Data analysis

Producing statistical parametric maps of brain activations from raw fMRI data obtained in an experiment usually involves (i) preprocessing, and (ii) statistical modelling. Preprocessing aims to correct any spatial or temporal confounds from data acquisition, before transforming the images into a standard space for activation reporting. Statistical modelling estimates the

statistical significance of brain responses within and across subjects to produce statistical maps (Penny et al., 2011).

Preprocessing

Motion is a major source of artefacts for fMRI images. Despite the use of head restraints, movement in subjects introduces a spatial mismatch between images acquired sequentially in time. *Spatial realignment* performs a rigid body transformation that minimises the difference between the reference image (usually the first one) and all others subsequently acquired (Friston et al., 1995). The rigid body transformation is defined in three translational and three rotational directions, and the realignment process produces a six-parameter estimate for each image, forming a matrix that can be incorporated in statistical modelling and for data quality inspection. In addition, artefacts from cyclic physiological fluctuations (e.g. cardiac or respiratory) can be corrected using principal component analysis focusing on the noise susceptible regions (Behzadi et al., 2007).

In order to extrapolate results from across subjects, brain images are often transformed to a common geometrical space. *Spatial normalisation* transforms images to fit a standard brain-space template, for example the Montreal Institute of Neurology (MNI) space derived from hundreds of healthy subjects. The process usually involves (i) coregistering fMRI images to higher resolution anatomical images so that both are in the subject space, (ii) normalising subject's anatomical image to match the template image through affine transformation (adding scaling and shearing to rigid body transformation), (iii) applying the warping parameters obtained from (ii) to the fMRI images (Holden, 2008). To increase signal to noise ratio, *smoothing* can be applied to images through convolution with a 3D Gaussian kernel, usually with 8-12 mm diameter, to reduce noise and the inter-subject differences that cannot be accounted for by normalisation (Penny et al., 2011).

Various software packages are available for fMRI data processing, including SPM (Statistical Parametric Mapping; http://www.fil.ion.ucl.ac.uk/spm/software/), FSL (https://fsl.fmrib.ox.ac.uk/), and many others. Preprocessing pipelines may vary for different software in terms of implementation and additional processing steps. Recent software packages such as fmriprep (https://github.com/poldracklab/fmriprep, Esteban et al., 2017) aim to incorporate preprocessing steps from different software to establish a generalised pipeline in order to increase reproducibility.

Statistical modelling

The purpose of statistical modelling of fMRI data is to detect statistically significant and regionally specific correlations between stimulus input and physiological responses in the brain.

The *general linear model* (GLM) approach to fMRI (Friston et al., 1994a; Friston et al., 1994b) models observed data (dependent variable, e.g. voxel time-series from images) as a linear combination of predictor variables (independent variables, e.g. experimental stimuli or conditions). Specifically, a GLM for i observations modelled by j predicting variables can be summarised as:

$$y_i = \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots + \beta_j x_{ij} + \varepsilon_i \tag{2.1}$$

where y is the observed time series of voxel i, x_{ij} the ith value of predictor variable j, β_j the parameter estimate for predictor variable j, and ε_i the error of observation i. The GLM can also be expressed in matrix and vector form:

$$\mathbf{Y} = \mathbf{X}\beta + \varepsilon \tag{2.2}$$

where $\mathbf{X}$ (referred to as the *design matrix*) is an $i \times j$ matrix where each column corresponding to a predictor x_j, β is a $j \times 1$ column vector of coefficients, and $\mathbf{Y}$ is the observation column vector of length i (Poldrack et al., 2011). $\mathbf{X}$ can include experimental manipulations proposed to modulate brain activities (i.e. regressors of interest), and other nuisance predictor variables such as subject movement (regressors of no interest). The design matrix is usually represented as a series of boxcar functions (ones at events with varying durations and zeros otherwise), and convolved with a canonical *haemodynamic response function* (HRF) to account for the 4-6s delayed peak time of BOLD effects.

The model parameters β are then estimated from the BOLD signals to test the hypothesis of whether an experimental manipulation has significant effect on a particular brain region. β values can be estimated using the classical or the Bayesian approach by minimising residuals (Penny et al., 2011). Iterating over all voxels yields a series of beta images, each representing the partial correlation between a predictor variable and the brain response, taking into account all other predictors specified in the model. Finally, the model parameters estimated are tested using t or F statistics to produce statistical parametric maps. A contrast vector $\mathbf{c}$ consisting of weights for predictors of interest is used to compute t statistics:

$$t = \frac{\mathbf{c}^\intercal \hat{\beta}}{\sqrt{\hat{\sigma}^2 \mathbf{c}^\intercal (\mathbf{X}^\intercal \mathbf{X})^{-1} \mathbf{c}}} \tag{2.3}$$

where $\hat{\beta}$ is the estimated model parameters, $\hat{\sigma}^2$ the variance of data residuals, $\mathbf{X}$ the design matrix (Pernet, 2014). F statistics are computed similarly with a contrast matrix $\mathbf{c}$ to test effects and/or differences across conditions.

GLMs typically use the *massive univariate* approach by fitting a separate model for each individual voxel. The large number of voxels and subsequent hypothesis tests can lead to a high risk of false positives (Type I errors). For example, a typical brain with ∼100,000 voxels, and a voxel-wise Type I error rate of 0.05 implies potentially 5000 false positive results. False positives can be controlled by thresholding the height of t or Z statistics for voxel activation with more stringent approaches. Random field theory is used for *multiple comparison correction* in the software SPM, which considers the likelihood of individual voxel exhibiting spatial correlation with each other, or the smoothness of the statistical map. This approach can reduce the effective number of independent observations across voxels when determining appropriate thresholds (Worsley et al., 2004), as compared to the overly conservative Bonferroni correction where all voxels are assumed independent. Given that in many neuroimaging studies there exists *a priori* hypotheses regarding the brain regions activated, it is possible to restrict multiple comparison correction within a smaller volume of interest instead of using the whole brain. *Small volume correction* (SVC) adjusts the activation p value taking into account the shape and size of the mask of volume provided, which is usually defined with anatomical atlases or coordinates from related activations in previous studies, when considering the activation's statistical significance (Worsley et al., 1996).

Model-based fMRI incorporates specific predictors generated from candidate computational models as GLM regressors by including them as parametric modulators (O'Doherty et al., 2007). Instead of examining activations directly resulting from experimental stimuli, the model-based method aims to probe how these activations might be *generated*, usually by testing competing hypotheses regarding their underlying computations. This method allows the use of fMRI to explore the higher level aspects of cognition such as learning and decision-making, with the added advantage that a parametric design also increases statistical power compared with the simple subtractive methods across different experimental conditions (Cohen et al., 2017).

Finally, *multivoxel pattern analysis* (MVPA), or *'decoding'*, examines the spatial activation patterns from ensembles of voxels to determine the information they collectively represent (Haxby et al., 2001). Classifier- or similarity-based MVPA learns different neuronal patterns corresponding to different cognitive states, potentially revealing how information can be represented distributedly across the whole brain (Cohen et al., 2017). MVPA classifiers are routinely used in neuroimaging-based *neurofeedback*, a technique that has seen increased attention with the development of the field of brain-machine interface. Using real-time fMRI, measured neural activity patterns can be decoded in real-time, and immediate perceptual

feedback are presented to participants based on their classified brain states (deCharms et al., 2004, 2005). Real-time feedback supposedly facilitates or guides self-regulation of the neural substrates underlying a particular behaviour or cognitive process (Sitaram et al., 2017). While the mechanism behind this regulation remains unclear, a neural network including the basal ganglia and insula has been found to be involved in the process, which suggests that learning may play an important role (Emmert et al., 2016).

To summarise, fMRI provides a safe, non-invasive way of studying brain function in actively behaving human participants. The unique combination of temporal recording and anatomical localisation in fMRI is particularly useful for investigating learning effects and their underlying computations.

2.2 Physiological and behavioural measurements

Physiological and behavioural measures are routinely used for validating and fitting computational models, helping to bridge the gap between functional theories and brain mapping (O'Doherty et al., 2007). In this thesis, additional physiological and behavioural data is acquired alongside fMRI measures to study the multiple facets of pain/relief learning.

Skin conductance responses (SCRs)

Skin conductance responses (SCRs) have been widely used as peripheral indicators of sympathetic arousal through sweat gland activities (Staib et al., 2015). Similar to a set of variable resistors connected in parallel, the sweat ducts are filled with varying amount of sweat depending on the sweat gland activity, which is controlled by the activation level of the sympathetic nervous system. The variable resistors become more *conductive* with higher amount of sweat produced, and the change in conductance can be measured by passing a small current through two electrodes attached to the surface of the skin (Dawson et al., 2007).

SCRs have been used in aversive conditioning studies as an indicator of learning success, and preparatory responses to upcoming events (Boll et al., 2013; Li et al., 2011; JS Morris and Dolan, 2004; Schiller et al., 2008). There are two ways to infer conditioning from SCRs. First, it is possible to look at the anticipatory SCR following a conditioning cue (CS), which is typically done during the trials in which the US, such as pain, is omitted, because of the slow rise time of the SCR (time-to-peak is usually several seconds). Alternatively, one can look at the SCR to the US to see how it is modified as a direct measure of conditioning.

SCR analysis involves extracting data features from the continuous recording. The conventional analysis approach involves band-pass filtering to reduce observation noise, defining a

response time window after stimulus events (0-5s used in studies cited above), and detecting the peak amplitude within this window following predefined criteria (Bach et al., 2009; Dawson et al., 2007). A more recent model-based approach is implemented in the software PsPM (Psycho-physiological Modelling; http://pspm.sourceforge.net/). Its linear model approach assumes sympathetic nervous responses follow stimulus onset with constant latency, and estimates SCR amplitudes with a GLM approach similar to that of fMRI analysis; or alternatively, it estimates SCR amplitudes with a non-linear model, assuming variable onset and latency for event-related SCRs, such as the anticipatory effects in conditioning (Bach et al., 2010, 2009). SCR amplitudes are usually transformed for normalisation before statistical testing since its distribution skews towards lower values (Schiller et al., 2008).

Electromyography (EMG)

Surface electromyography (EMG) offers the opportunity to study a range of motor-related conditioned responses. Limb EMG has been used to study limb withdrawal reflexes during aversive conditioning, allowing analysis of response incidence and magnitude on a trial-by-trial basis with millisecond precision (Kaulich et al., 2010; Timmann et al., 2000). Facial EMG measures are sensitive to changes in spontaneous facial expressions, which has been used to assess response to nociception during anaesthesia and emotional conditioning experiments (Criswell, 2011; Littlewort et al., 2009). Choosing the specific muscles to record EMG from often requires considerations of the movement range of possible conditioned responses, muscle function and size, the ease of electrode attachment, as well as posture constraints (Criswell, 2011).

Heart rate

Changes in heart rate compared to baseline can also be used to assess conditioning. Specifically, cardiac deceleration is thought to be an index of perceptual processing of sensory information through parasympathetic activity, and unpleasant stimuli are associated with more pronounced deceleration (Bradley, 2009). Cardiac acceleration has been interpreted as evidence of mobilisation for avoidance, and is related to muscle preparation (Hamm et al., 1993). These characteristics make heart rate a suitable measure in aversive conditioning, however, it has been observed to be less sensitive to changes in stimulus information compared to SCRs (Peri et al., 2000).

Subjective ratings

For internal events that cannot be directly observed, self-reporting as a means of communicating sensory experience is frequently used for assessment (Turk and Melzack, 2011). To examine the influence of learning on perceptual experience, ratings of pain and relief are typically collected repeatedly in conditioning experiments to capture their fluctuations in magnitude over a period of time. *Visual analogue scales* (VAS) are often used for this purpose, where the subject is asked to rate their pain based on a linear scale from minimum to maximum representing 'no pain' to 'unbearable pain' ('no relief' and 'very pleasant relief' for relief ratings). The validity of VAS has been demonstrated via its positive correlation to other forms of pain self-report measures and observed pain behaviour (Jensen et al., 1986; DD Price et al., 1983). However, VAS can be relatively more time-consuming and cognitively demanding compared to other measures (Turk and Melzack, 2011).

Choices

During instrumental conditioning, participants are usually given the choice of multiple responses, each being reinforced according to a separate schedule operating concurrently ('concurrent schedule', Bouton, 2007). Choices can reflect *which*, *when*, and *how much* of a reinforcer one prefers, which can be modelled within the RL framework. Apart from being realistic and intuitive, choice reaction time can also reveal additional information for inference, such as uncertainty or expectation violation in decision-making, as well as impulsiveness and self-control in personality traits (Bouton, 2007; Meyniel et al., 2016).

2.3 Pain stimulation

Methods for eliciting noxious experimental pain safely and reliably in humans become more limited when compatibility with MRI needs to be considered. The experiments in this thesis used thermal and electrical pain stimulations during fMRI scanning, delivered using MRI-safe stimulators with special physical filters for noise reduction.

Stimulation methods

Thermal stimulation activates pain-specific A-delta and C fibres in the ascending nociceptive pathways using a contact heat-evoked potential stimulator (CHEPS) (Chen et al., 2001). Thermal pain thresholds are relatively consistent across and reproducible within individuals (Byrne and Dafny, 1997; Chakour et al., 1996). Uniquely, thermal stimuli can be used to deliver tonic

pain to mimic a clinically valid pain relief experience, through the rapid heating and cooling capabilities of CHEPS (70 and 40 °C/s respectively). Strict temperature and duration limits are programmed in thermal stimulator software to ensure safe delivery of tonic thermal pain without the risk of burns (Medoc, 2017).

Electrical stimulation is also widely used to elicit a painful, aversive state in experimental psychology studies (Li et al., 2011; Schiller et al., 2008; Seymour et al., 2004). By activating both A-delta and non-nociceptive A-beta fibres, electrical stimulation produces a sharp, focal painful sensation, especially when delivered through a concentric surface electrode. This type of electrode consists of a central anode and a concentric circular cathode of 10mm diameter, and was initially designed to study facial pain, but has since been used in many experimental pain studies (Seymour, 2010). Electrical stimulation is more appropriate for tasks with repetitive stimulation because it has lower peripheral habituation comparing to thermal stimulation (Seymour et al., 2004).

Calibration procedures

As individuals have different pain thresholds, calibration is conducted to set threshold stimulation parameters (temperature or electrical current) for each participant before each experiment. This typically consists of a staircase procedure, followed by a randomised sequence of stimuli for participants to rate in terms of their subjective intensities. The staircase method increases stimulation intensities by a fixed amount each time to establish a suitable range, in a way that participants can anticipate and stop upcoming trials if needed. The randomised sequences that followed remove potential anticipatory effects of the staircase method (Derbyshire et al., 1997; Seymour et al., 2004). A sigmoid or Weibull function can then be fitted to the stimulation parameters and the collected ratings to determine the appropriate settings for the desired subjective pain level (Seymour, 2010).

2.4 Computational modelling

A generative model proposed to explain the pain/relief learning process should explain experimental data well while structurally remaining as simple as possible, striking a balance between biological plausibility and computational tractability (Claeskens and Hjort, 2008). *Model fitting* refers to the procedure of estimating model parameters to produce predictions that fit the experimental data best (Myung, 2003). The goodness of model fit can be assessed with various summary statistics, which are often used for *model comparison* amongst competing models in order to determine the most plausible theoretical explanation.

Model fitting

Trial-by-trial analysis is particularly suitable for learning models as it can capture the temporal dynamics and complexity of the data. For example, in a conditioning experiment, a trial typically consists of a visual cue, followed by the participant's chosen action, and finally the outcome of reward or punishment. The *evolution function* of a learning model produces a prediction at each of these trial time points (e.g. cue value, action value, outcome prediction error), which combines aggregated experience from past trials with actual response in the current trial. These model predictions are assumed to reflect the experimental data (e.g. action chosen, evoked SCR, subjective rating) with additive noise, described by the *observation function*. Together, these functions map the components underlying learning processes to the observable neural and/or physiological data (Daunizeau et al., 2014; Daw et al., 2011).

Model parameters are used to specify a model, and to characterise factors of experimental interest (e.g. learning rate). Model fitting following *maximum likelihood estimation* (MLE) seeks to find the free parameter estimates $\hat{\theta}_M$ that maximises $P(D|M,\theta_M)$, the probability of observing current dataset D given a model M, as described by Bayes' rule:

$$P(\theta_M|D,M) \propto P(D|M,\theta_M) \cdot P(\theta_M|M) \tag{2.4}$$

For example, one can estimate the learning rate in an RL model that maximises the probability of choosing an action given a visual cue. Typically, free parameters of the evolution and observation functions are estimated within a predetermined range to maximise the likelihood estimate aggregated over the entire dataset. This process can be accomplished by discretising the parameter space for a grid search, or using the nonlinear optimisation functions available in scientific computing software (e.g. *fmincon* in MATLAB) (Daw, 2011). A variational Bayesian approach is also implemented in the software VBA toolbox (Variational Bayesian Analysis; https://mbb-team.github.io/VBA-toolbox/) to maximise the approximation of the log model evidence (Daunizeau et al., 2014).

Model parameter estimates are likely to differ between individuals, potentially reflecting certain population-level characteristics. The *summary statistics* approach treats each participant's parameter estimate as a random variable, and the means of group parameters are used for population-level inference (Friston et al., 2005). In a *fixed effect* analysis, the same model is assumed to generate the data for all individuals of a group. Alternatively, for a *random effect* analysis, individual subject's model parameters are assumed to be drawn from an unknown population distribution, where theses parameters are allowed to differ across individuals. Fixed effects analysis does not consider between-subject variability, which contributes to higher risk of false positive results (Poldrack et al., 2011). The group-level summary statistics can be

assumed to define subject-level parameters hierarchically, acting as constraints for the likely parameter range. Group-level parameters can then be estimated by updating group parameter summary statistics from all subjects' parameter posterior and iterating until convergence. This *mixed-effect modelling* approach has the advantage of reducing the likelihood of producing outliers in individual parameter estimates because of the group constraints (Daunizeau et al., 2014; Daw, 2011; Penny et al., 2004).

Model comparison

Model comparisons address the question of to what extent the experimental data supports a collection of proposed models. The data likelihood $P(D|M,\hat{\theta}_M)$ assigned for model parameter optimisation can be used as model evidence to compare among different models. However, these measures need to be corrected to account for model complexity, usually by taking into account the number of free model parameters, in order to avoid overfitting (Daw et al., 2011). For learning models, the *Bayesian Information Criterion* (BIC, Schwarz, 1978) is often used:

$$\text{BIC} = \frac{k}{2} log\text{n} - log(P(D|M,\hat{\theta}_M)) \tag{2.5}$$

where n is the number of data points (e.g. choices from all trials) and k is the number of free parameters in the model. However, the penalty of BIC for model complexity ignores the different admissible range of individual model parameters (i.e. the priors). This problem can be circumvented by Bayesian model selection (BMS), where parameter priors are factored into the calculation (MacKay, 2003).

When considering model evidence from multiple participants, model identity could be considered as a random effect that varies across individuals, where the best fitting model can be different for each participant (Daunizeau et al., 2014; Stephan et al., 2009). Random effect BMS assumes one model out of a collection was drawn from a distribution for each individual, in addition to the variations in model parameters and observed data. This inference therefore reflects how likely a model fits the individuals within a population. For fixed effect BMS, on the other hand, the same model is assumed to generate the data of all individuals, and different models are compared based on the summed model evidence from these individuals.

In summary, collecting fMRI and other physiological and behavioural data during pain/relief learning experiments allows us to test and compare various possible reinforcement learning models. This informs us about the generative processes underlying the different conditioned responses, and the effects of learning on the pain/relief hedonic experience. The main workflow used in the studies in this thesis is summarised in Fig 2.1.

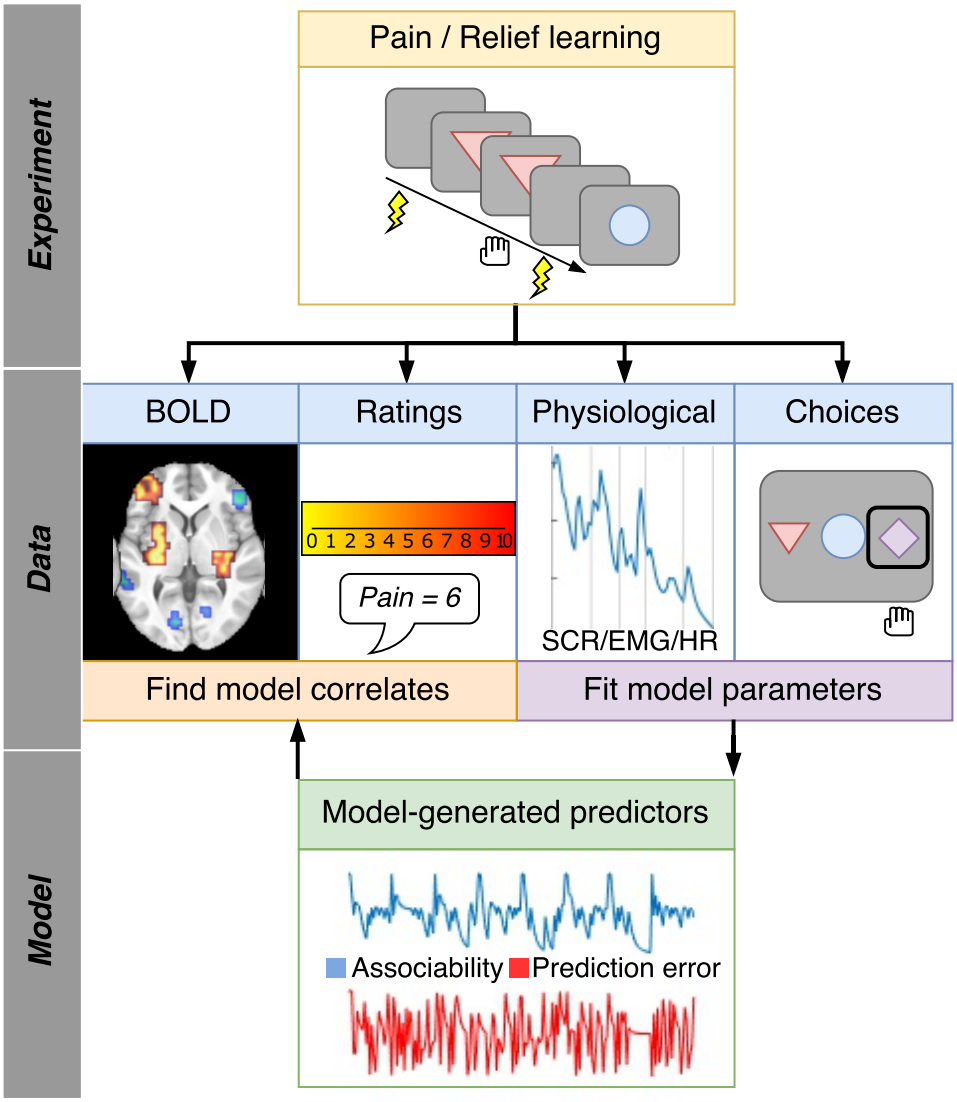


Fig. 2.1 Methods summary. The main purpose of the studies in this thesis is to understand pain and relief representation and processing in the brain. A series of data, including fMRI BOLD signals, subjective ratings, physiological recordings, and choices can be collected during the pain/relief learning experiments designed to study their motivational effects. Generative models proposed to explain pain/relief information processing are usually fitted with behavioural data (e.g. physiological data or choices) from individuals, and the winning model emerges from model comparison is considered the most likely candidate underlying the computation. Model-generated predictors are then used to localise brain regions with model correlated activations, and to assess the impact of learning on hedonic experiences.

Chapter 3

Experiment 1: Dissociable learning processes underlie pain conditioning

Pavlovian conditioning underlies many aspects of pain behaviour, from the execution of protective motor responses to the elicitation of fear. However, it remains unclear whether there are multiple learning mechanisms involved in these behaviour, and how they might be represented in the brain. Using a parallel conditioning paradigm in which different Pavlovian cues predicted pain to either the left or right forearm, we show that conditioning involves two distinct learning processes. First, a preparatory system learns autonomic responses and is correlated with brain activity in amygdala-striatal circuits (the classic 'fear-learning' circuit). Second, a 'consummatory' system learns limb-withdrawal responses, detectable with limb electromyography and correlated with ipsilateral cerebellar activity. These results define a distinct computational role for the cerebellum in pain, and show that the overall phenotype of conditioned pain behaviour depends on two parallel reinforcement learning circuits.

Adapted from Zhang et al. (2016)

3.1 Introduction

Pavlovian conditioning underlies many aspects of pain behaviour, including fear and threat detection (LeDoux, 2014), escape and avoidance learning (Gerber et al., 2014), and several types of endogenous analgesia (Wager and Atlas, 2015). Although a central role for the amygdala is well established (Phelps and LeDoux, 2005), both human and animal studies implicate other brain regions in learning, notably ventral striatum and cerebellum (Seymour et al., 2004). It remains unclear whether these regions make different contributions to a single aversive learning process, or represent independent learning mechanisms that interact to generate the expression of pain-related behaviour.

We designed a human parallel aversive conditioning paradigm in which different Pavlovian visual cues probabilistically predicted thermal pain primarily to either the left or right arm, and studied the acquisition of conditioned Pavlovian responses using combined physiological recordings and functional magnetic resonance imaging (fMRI). Using computational modelling based on reinforcement learning theory, we found that conditioning involves two distinct types of learning processes. First, a 'preparatory' system learns non-specific conditioned responses such as aversive facial expressions and autonomic responses including skin conductance. The associated learning signals – the learned associability and prediction error – were correlated with fMRI brain responses in amygdala-striatal regions, corresponding to the classic aversive (fear) learning circuit. Second, a specific lateralised system learns 'consummatory' limb-withdrawal responses, detectable with electromyography of the arm to which pain is predicted. Its related learned associability was correlated with responses in ipsilateral cerebellar cortex, suggesting a novel computational role for the cerebellum in pain. In conclusion, our results show that the overall phenotype of conditioned pain behaviour depends on two dissociable reinforcement learning circuits.

3.2 Methods

Subjects

15 healthy human subjects (one female) participated in a Pavlovian first-order delay conditioning experiment (Fig 3.1). All subjects gave informed consent prior to participation, and the study was approved by the Ethics and Safety committee of the National Institute of Information and Communications Technology, Japan.

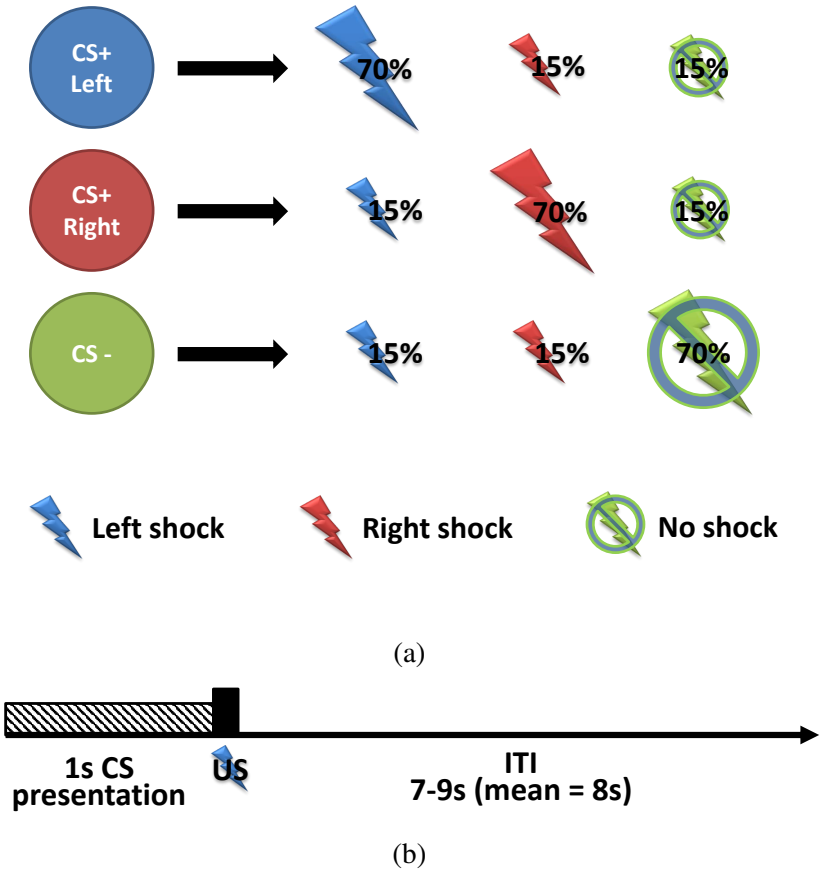


Fig. 3.1 Experimental design. (a) Each trial involved one of three Pavlovian CS cues, each of which primarily predicted (70%) either left pain (blue symbol), right pain (red), or no pain (green), and infrequently predicted the other outcomes (15%). (b) On each trial, a 1s CS cue was followed immediately by pain or no pain (US) in a delay conditioning procedure, followed by a variable 7-9s inter-trial interval (ITI).

Experimental design

Subjects learned conditioned associations between different visual cues (abstract coloured images presented on a computer screen) and brief painful heat stimuli delivered either to the left forearm, the right forearm, or not at all. A relatively short CS-US interval of 1s was used to optimise detection of reflex-like conditioned muscle activities, similar to the design of eye-blink conditioning studies (Daum et al., 1993). Ultra-brief painful heat stimuli at 55°C were delivered through two 27mm diameter contact heat-evoked potential stimulators (CHEPS, Medoc Pathway, Israel), to the subject's left or right inner forearm. The CHEPS thermodes

can heat up rapidly at 70°C/s to 55°C, followed by immediate cooling at 40°C/s to baseline temperature of 30°C.

Physiological measurement and analysis

Physiological signals were continuously recorded using MRI compatible BrainAmp ExG MR System with specialised electrodes and sensors (Brain Products, Munich, Germany, Fig 3.4). Off-line processing and analysis were implemented in MATLAB (The MathWorks Inc., MA, USA).

Skin conductance responses (SCR) were assessed as the peak-to-peak amplitude difference in a time window of 0.5-4.5s after cue onset (pain-omitted trials), and 0.5-5.5s (pain trials). Raw SCR magnitudes were square root transformed for normalisation (confirmed with Kolmogorov-Smirnov test and visual inspection of histogram), and scaled to individual subject's mean square root transformed US response (Li et al., 2011; Schiller et al., 2008).

Upper-limb electromyography (EMG) recordings were taken from the brachioradialis and biceps-brachii muscles on both arms. MRI artefacts were removed by using a custom-made filtering program (Ganesh et al., 2007). The resultant EMG signals were band-pass filtered at 10-150Hz, full wave rectified, and baseline adjusted. The signals from 1s CS-US interval were extracted and sorted according to trial types for further analysis. Moreover, conditioned EMG response (CR) was defined as where ISI EMG activity reached 30% of the EMG maximum of that trial, staying above that with a minimum duration of 200ms, and a minimum integral of $1\text{mV} \cdot \text{ms}$ (Thieme et al., 2013). The percentage of EMG CR incidence was averaged across left and right.

Facial EMG (corrugator muscle) and heart rate were collected in behavioural study only (Fig 3.5a and 3.5b). Facial EMG during the 1s CS-US duration was averaged within 0-500ms and 500-1000ms bins for statistical comparison across subjects. Due to hardware constraint, SCRs were recorded on left side only, as there is no definitive evidence of laterality difference between electrodermal activity recorded on left or right hand (Dawson et al., 2007). It should be noted that null effects presented in results might be significant given a larger sample size.

Computational model analysis

We constructed reinforcement learning models, fitted trial-by-trial model value / associability to SCR data for parameter estimation and model comparison, and then used obtained learning signals to probe brain activity (Boll et al., 2013; Li et al., 2011; Schiller et al., 2008). In this way, the brain responses are specifically related to the behaviourally fitted learning model. These models can be used to test competing hypotheses about the neural representation of

preparatory (i.e. laterality non-specific) and consummatory (i.e. laterality specific) learning processes.

Standard temporal difference model

This model is the simple 'real-time' instantiation of the Rescorla-Wagner (RW) model (Rescorla and Wagner, 1972). The value V of trial $n+1$ for a given cue j is updated based on the value of current trial n and the prediction error, the difference between current value V_j and outcome stimulus value R at trial n, weighted by a constant learning rate α :

$$V_j(n+1) = V_j(n) + \alpha \cdot (R(n) - V_j(n)) \tag{3.1}$$

where the learning rate α ($0 \leq \alpha \leq 1$) is a free parameter.

Hybrid temporal difference model

The hybrid model combines both Rescorla-Wagner and Pearce-Hall (PH) models, where the RW rule is used for error-driven value update and PH associability is used as a dynamic learning rate for RW to modulate predictive learning (Li et al., 2011). The value of associability decreases if the conditioned stimuli become correctly predictive of the stimuli outcome (Pearce and Hall, 1980). The values for the hybrid model were updated as follow:

$$V_j(n+1) = V_j(n) + \kappa \cdot \alpha_j(n) \cdot (R(n) - V_j(n)) \tag{3.2}$$

$$\alpha_j(n+1) = \eta \cdot \left|R(n) - V_j(n)\right| + (1-\eta) \cdot \alpha_j(n) \tag{3.3}$$

where free parameters α_0 (initial associability, $0 \leq \alpha_0 \leq 1$), $\kappa (0 \leq \kappa \leq 1)$, $\eta (0 \leq \eta \leq 1)$, are determined by fitting to behavioural data.

Assuming the preparatory learning system cannot distinguish lateralised outcomes, then $R(n) = 1$ for all pain trials regardless of laterality. While the consummatory learning system tracked outcomes ipsilateral to its side only, ignoring the opposite side, then for the left system, $R(n) = 1$ for left pain, or $R(n) = 0$ for both right pain and no pain, and vice versa for the right system. The three learning systems were assumed to share the same learning parameters, but were sensitive to different outcome types.

For individual sessions, model parameters were fitted by maximizing likelihood for individual subject's sequence of SCRs, modelled as the normal distribution around a mean determined by the predicted value V (or associability α, or the sum of both), computed by the model on trial n, scaled by free parameters β_1, β_2 and shifted by a constant term β_0, with distribution

variance σ (Li et al., 2011).

$$SCR_n \sim N(\beta_0 + \beta_1 V_n, \sigma) \tag{3.4}$$

$$SCR_n \sim N(\beta_0 + \beta_1 \alpha_n, \sigma) \tag{3.5}$$

$$SCR_n \sim N(\beta_0 + \beta_1 V_n + \beta_2 \alpha_n, \sigma) \tag{3.6}$$

To avoid contamination by pain over CS-predictive responses, only SCRs of no pain (i.e. unreinforced) trials were fitted, but all trials were used in the computation of value and associability. We obtained population free parameters using a hierarchical model fitting approach (a mixed effect method) for subsequent imaging analysis (Daw, 2011). Bayesian information criterion (BIC) value was calculated for each model with optimal individual parameters to quantitatively compare goodness of fit (Table 3.1).

fMRI data analysis

Functional MRI imaging data was acquired on a 3T Siemens Magnetom Trio scanner with Siemens standard 12 channel phased array head coil.

Functional images were collected using a single-shot gradient echo EPI sequence (repetition time TR=2500ms, echo time TE=30ms, field of view=240mm, flip angle=80°). Thirty seven contiguous oblique-axial slices (3.75 × 3.75 × 3.75 mm voxels) parallel to the AC-PC line were acquired. Whole-brain high resolution T1-weighted structural images were obtained. Preprocessing of imaging data was conducted using SPM8 following standard procedures (Wellcome Trust Centre for Neuroimaging, UK; http://www.fil.ion.ucl.ac.uk/spm/).

We conducted a parametric analysis, in which the computational model generated learning signals were used as parametric regressors at the time of CS (visual cue) and US (pain outcome) presentation for each trial (O'Doherty et al., 2007). These trial time points were modelled with stick functions to represent events as finite impulses with zero durations. The best fitting hybrid model from the SCR-based analysis was used to generate the following regressors with population free parameters:

At outcome time:

- Preparatory associability α_{general},
- Left-sided consummatory associability α_{left},
- Right-sided associability α_{right},

At cue and outcome time (i.e. 'full' prediction error as a biphasic response):

- Preparatory prediction error VD_{general},
- Left-sided predicted error series VD_{left},
- Right-sided prediction error series VD_{right},

Regressors of no interest:

- Left pain delivery,
- Right pain delivery,
- Motion parameters ($\times 6$) from affine realignment in preprocessing.

All these regressors were compiled into one single general linear model (GLM) for first-level analysis for individual subject in SPM8. Note that associability and prediction errors were relatively uncorrelated (see demonstration in Fig 2.1), their inclusion in a single regression model is unlikely to be confounded by shared variance (Li et al., 2011). Resulting contrasts were used in second-level one-sample t-tests to make population inference (Fig 3.3). Small volume correction (SVC) for multiple comparison was conducted within anatomically defined 8mm diameter spherical masks built around hypothesised structure coordinates of the amygdala, ventral putamen and cerebellum (Table 3.2).

Functional region of interest (ROI) analysis of the cerebellum was conducted using the Spatially Unbiased Infratentorial Template (SUIT) atlas (Diedrichsen, 2006). Masks of the cerebellum were created using T1-weighted structural scans for each subject, spatially normalised to the SUIT template. Resultant contrasts from first-level analyses were then resliced into SUIT atlas space using previously generated SUIT normalisation parameters. Spatial smoothing of the functional data was omitted in order to avoid contaminating activation from the visual cortex. The SUIT probabilistic MRI atlas of human cerebellum was used to locate cerebellar lobules (Diedrichsen et al., 2009). In addition, post-hoc analyses of all ROIs were conducted by extracting beta estimates for each subject from the functional clusters of interest as they appear in given contrasts using MarsBaR toolbox (http://marsbar.sourceforge.net/). They were then averaged across subjects according to model or trial types without parametric modulation.

3.3 Results

The brain is acutely tuned to detecting a variety of threats, especially pain, and elicits a set of appropriate responses as soon as potential harm is detected. This classic ‘fear’ response is critical for survival, and the way in which cues in the environment are used to predict harm

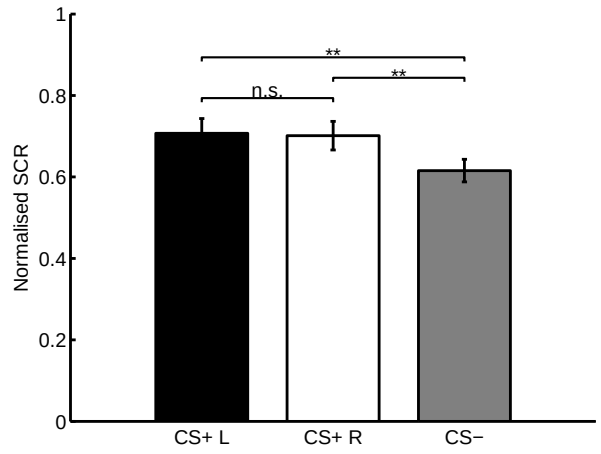
(Pavlovian conditioning) represents one of the most important and evolutionarily conserved learning systems in animals. However, it is not clear whether the overall phenotype of the pain-based fear response represents a single process, or the sum of partially independent processes.

Physiological responses

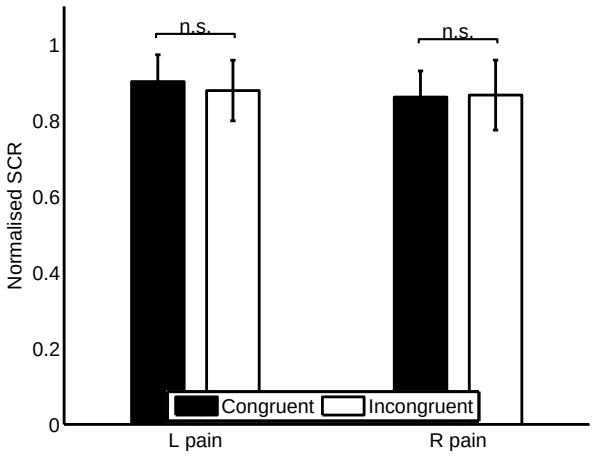
We recorded a number of different physiological responses to evaluate the acquisition of conditioned responses. SCRs did not distinguish the laterality of predicted or received pain, consistent with a preparatory response. Specifically, SCRs showed comparable conditioning to cues that predicted left (CS+ L) or right (CS+ R) arm pain, in comparison to control (CS−) (Fig 3.2a, data represented as mean ± SEM). SCRs to the pain itself were also comparable regardless of whether the pain was delivered to the predicted (congruent) or unpredicted (incongruent) side (Fig 3.2b). We could not identify any significant laterality differences in early or late learning periods during each session, from either normalised SCR magnitude or rise time to peak (Fig 3.5c and 3.5d).

Facial EMG also followed a preparatory pattern. The EMG was recorded from the corrugator muscle, a characteristic muscle of aversive expression, during a behavioural version of the task (Fig 3.4). The response during the 1s CS-US interval averaged across trials showed a significant increase in 500-1000ms time window, for both CS+ L and CS+ R trials compared to CS− trials (combined CS+ L/R vs. CS− paired t-test $p<0.05$ in 500-1000ms), but not significant between CS+ L and R groups ($p>0.1$ for all sample points, Fig 3.2c). Comparing pain-evoked responses for congruent and incongruent prediction trials during 1s duration after painful US delivery revealed no statistically significant differences, consistent with a preparatory response (both $p>0.5$, Fig 3.2d).

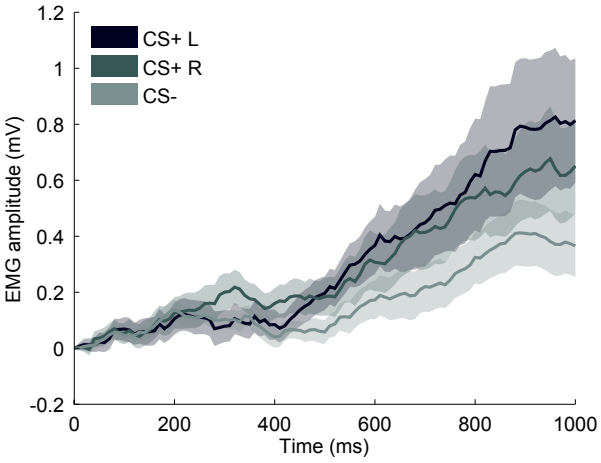
In contrast, EMG responses from each arm (recorded from brachioradialis and biceps-brachii, which are involved in upper limb withdrawal) showed lateralised 'consummatory' patterns. We recorded activity in the 1 second CS-US interval, and compared it to pre-CS baseline activity. We found that responses were significantly greater in the arm in which pain was predicted (ipsilateral) as opposed to the contralateral side (Fig 3.2e and 3.2f). We did not test for facial/limb EMG interaction with laterality as they were collected from different versions of the same experiment. Note that because of the proximity of the stimulating thermode and the EMG electrodes, US responses (to look for congruency effects) are unavoidably too corrupted by electrical artefact for analysis.


Normalised SCR
1
0.8
0.6
0.4
0.2
0
**
**
n.s.
CS+ L
CS+ R
CS−

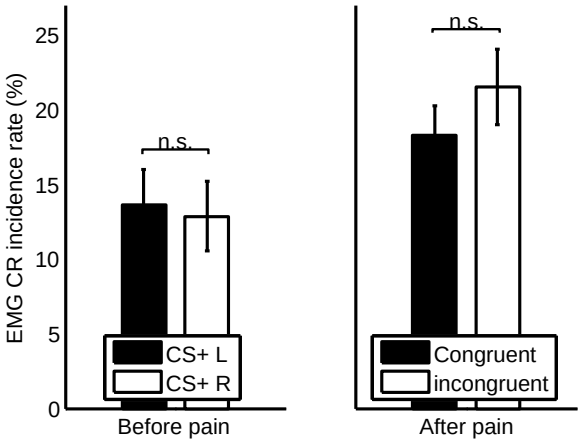
(a)


Normalised SCR
1
0.8
0.6
0.4
0.2
0
n.s.
n.s.
Congruent
Incongruent
L pain
R pain

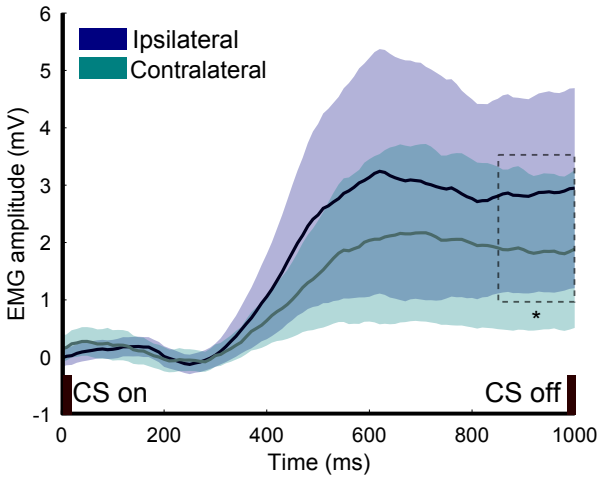
(b)


EMG amplitude (mV)
1.2
1
0.8
0.6
0.4
0.2
0
-0.2
CS+ L
CS+ R
CS-
0
200
400
600
800
1000
Time (ms)

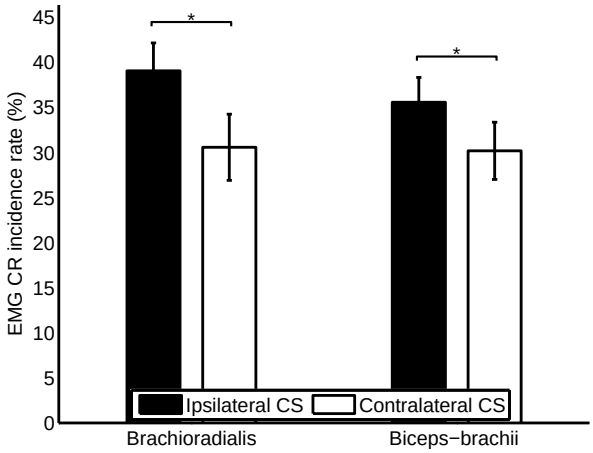
(c)


EMG CR incidence rate (%)
25
20
15
10
5
0
n.s.
n.s.
CS+ L
CS+ R
Congruent
incongruent
Before pain
After pain

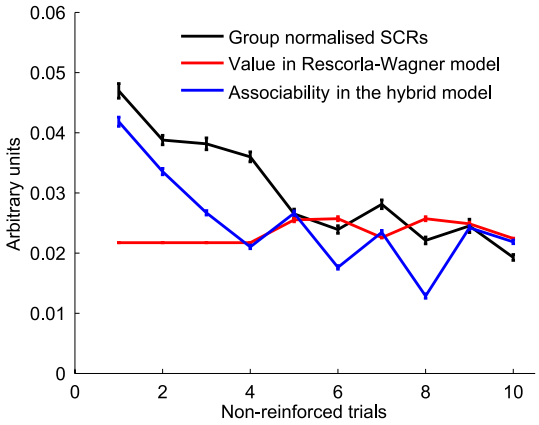
(d)


EMG amplitude (mV)
6
5
4
3
2
1
0
-1
Ipsilateral
Contralateral
*
CS on
CS off
0
200
400
600
800
1000
Time (ms)

(e)


EMG CR incidence rate (%)
45
40
35
30
25
20
15
10
5
0
*
*
Ipsilateral CS
Contralateral CS
Brachioradialis
Biceps−brachii

(f)



(g)

Fig. 3.2 Behavioural results. (a) CS-evoked SCRs in 'unreinforced' trials show significant differences between CS+ L/R and CS− ($T_L(41) = 2.78$, $T_R(41) = 2.99$, both p<0.01), but not between CS+ L and CS+ R ($T(41) = 0.14$, $p = 0.89$). (b) SCRs for reinforced pain trials with congruent / incongruent predictions, separated into L/R pain groups, showing no significant differences. (c) Facial EMG traces during 1s CS-US interval show CS+ L/R $>$ CS− in amplitude (combined CS+ L/R vs CS− p<0.05 in 500-1000ms), but not significantly different between CS+ L/R (all time points p>0.1). (d) Average facial EMG conditioned response (CR) incidence show no significant difference between CS+ L/R during 1s CS-US interval before, or during 1s after pain delivery, between congruent / incongruent trials (both p>0.5). (e) Time-course of upper-limb EMG during 1s CS-US interval averaged across L/R, with ipsilateral $>$ contralateral response amplitude (p<0.05 in 850-1000ms). (f) Average upper-limb EMG CR incidence in brachioradialis and biceps-brachii muscles, significantly greater for ipsilateral trials (both p<0.05). (g) Trial-by-trial model fit of associability (blue) and value (red) to group normalised SCRs (black) of non-reinforced trials in 1 session (first 10 trials). Data represented as mean $\pm$ SEM. *p < 0.05; **p < 0.01; n.s., not significant.

Neuroimaging results

Reinforcement learning theory proposes that acquisition of conditioned responses from trial-by-trial experience utilises two key measures: a prediction error term that records the difference between pain expectations and outcomes (Rescorla and Wagner, 1972); and an 'associability' term that keeps track of the uncertainty of predictions (Pearce et al., 1981). These two measures are then integrated to update CS values that provide the prediction for the next trial. Accordingly, the larger the prediction error the greater the update in CS value. The associability term acts as the learning rate of value, with higher associability representing greater uncertainty and hence more rapid learning.

SCRs were of sufficient fidelity to permit trial-by-trial analysis using a computational statistical model fitting procedure. In agreement with previous reports (Boll et al., 2013; Li et al., 2011), we found the SCR sequences were best described by a preparatory associability term, illustrated in Fig 3.2g, as confirmed with a model comparison procedure based on model fitting BICs (Table 3.1).

We then used the estimated model parameters in a linear regression with brain responses recorded by concurrent fMRI, to identify whether anatomically distinct learning signals related to preparatory and left / right consummatory learning signals could be dissociated. We used the computational parametric regressors for all learning signals (associability and prediction error for both preparatory and consummatory temporal difference models) in a single GLM (see methods). These values were generated using population free parameters with the best fitting model, the hybrid model, obtained from the behavioural data (SCRs) fitting procedure mentioned earlier.

We found that bilateral ventral putamen and amygdala BOLD signals correlated with a preparatory temporal prediction error and associability signal respectively (Fig 3.3a and 3.3b). In contrast, left and right consummatory associabilities correlated with ipsilateral cerebellar responses. Associability signal clusters were located symmetrically in lobule left V extending into left VI, and spanning the border between lobules right V and right VI (Fig 3.3c). The peak coordinates of these cerebellar activations were in grey matter, as identified by the SUIT atlas (Table 3.2). In addition, post-hoc analyses of functional ROIs support the hypothesised roles of structures identified by computational models. Beta estimates were extracted for each subject from the functional clusters of interest as they appear in the contrasts presented in the results. They were averaged across subjects according to model or trial types without parametric modulation, where amygdala, putamen and cerebellum showed differential responses to preparatory and consummatory model outputs (Fig 3.5e).

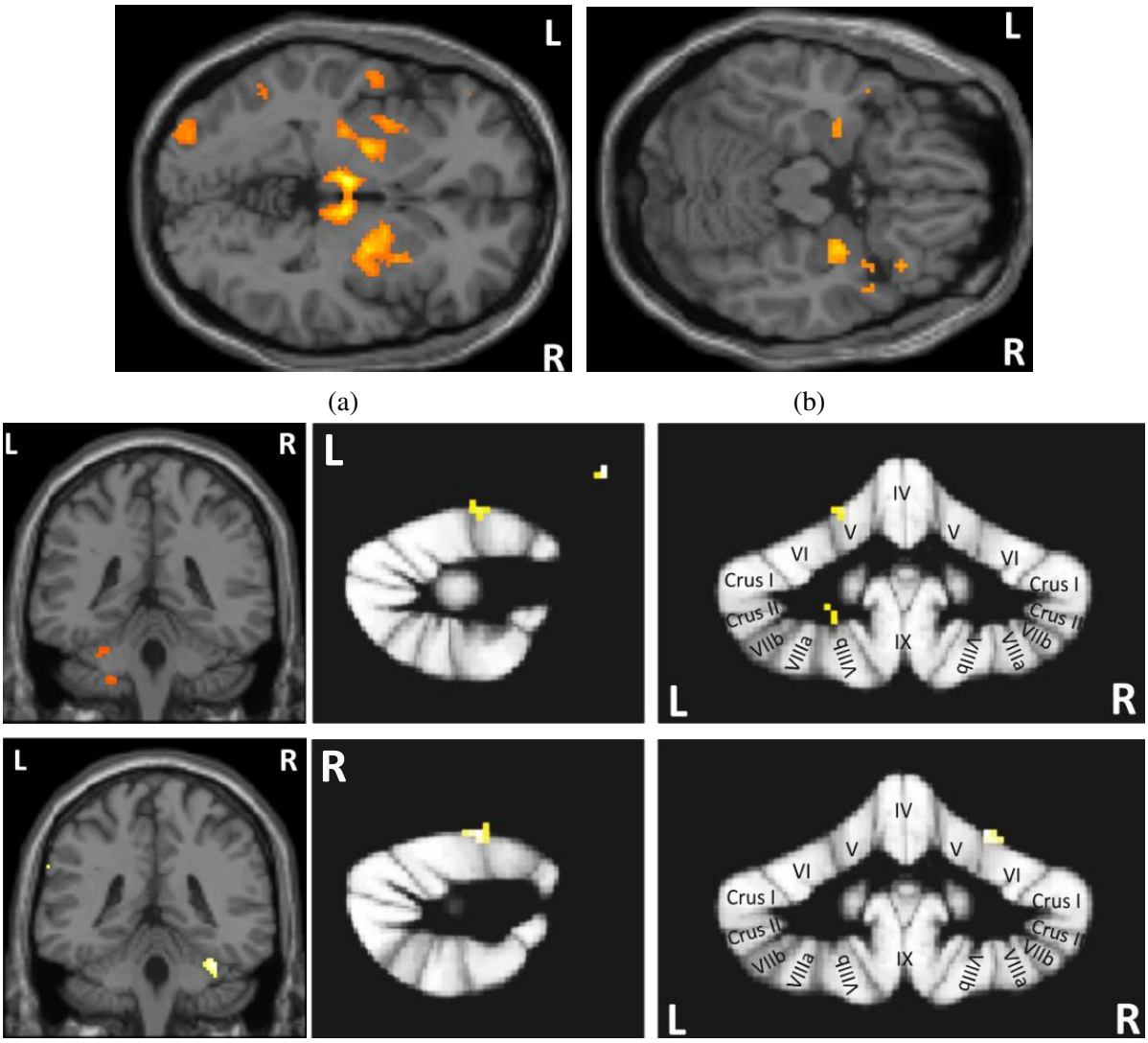


Fig. 3.3 Statistical parametric maps showing: (a) Preparatory prediction error in bilateral ventral putamen (MNI z=-2, p<0.001 unc.). (b) Preparatory associabilities in bilateral amygdala (z=-18, p<0.01 unc.). (c) Ipsilateral activations to consummatory associabilities (y=-40, p<0.001 unc., all p<0.05 in small volume correction (SVC) using anatomically defined 8mm diameter spherical ROI masks built around hypothesised structure coordinates, Table 3.2). ROI analysis of cerebellum using SUIT probabilistic atlas template showing: Top: left anterior cerebellum activations in the border between lobule V and VI (SUIT space coordinates: 24, -52, -15), and in lobule VIII (-22, -50, -41, p<0.004 unc.); bottom: right anterior cerebellum activation in the border between lobule V and VI (-18, -52, -13, p<0.001 unc.). unc.: uncorrected threshold.

3.4 Discussion

In summary, our results dissociate two distinct response-learning systems underlying human pain. An amygdala-striatal system was identified with predictors from a general computational reinforcement learning model that largely ignores information about the laterality of pain, and the same model predictions also explain preparatory conditioned responses, including autonomic responses and facial expression. In contrast, a cerebellar system was identified with a laterality-specific learning model, which also corresponds with specific 'consummatory' limb withdrawal responses appropriate to the anatomical site of predicted pain.

The role of the amygdala in preparatory conditioning is well established. For instance, amygdalar lesions impair autonomic responses, freezing, potentiated startle, and active avoidance (Gerber et al., 2014; LeDoux, 2014). Our data shows that a preparatory associability signal correlates with activity at the level of the fMRI BOLD, consistent with previous studies in both humans and rodents (Boll et al., 2013; Holland and Gallagher, 1993, 2006; Li et al., 2011). It is important to note, however, that aversive prediction errors have been identified at a neuronal level in rodents (Johansen et al., 2010; McHugh et al., 2014). Although there exist species and methodological differences in comparison to our study, it illustrates the differences in methodology between BOLD responses and neuronal physiological recordings. In particular, since the BOLD signal could be conveying the average signal of a potentially computationally heterogeneous group of neurons, some caution is needed against over-interpretation of the results. On the other hand, it is still unclear how some computational quantities might be encoded by distributed activity of a population of neurons. In addition, while the amygdala contributes to SCR control (Davis, 1992), the associability signal was fitted to SCRs from only a subset of non-reinforced trials (1/3 of all trials) for model parameter estimation. Therefore the amygdala BOLD response parametrically correlated with associability is unlikely to be driven by SCRs alone.

Results from other studies also argue against any simplistic single model of amygdala function. For example, amygdala responses have been shown contralateral to the shock laterality in unilateral eye-blink conditioning (Blair et al., 2005), and to exhibit non-symmetrical activations in a range of fear paradigms (Apergis-Schoute et al., 2014), in contrast to the results here which lacked laterality dissociation. Other factors such as motivational state (Balleine and Killcross, 2006), and sensitivity to inferred ('model-based') cue-outcome contingency (Prévost et al., 2011) have also been demonstrated. Therefore, whilst our computational model-based analysis showed that the expression of preparatory responses did not appear to distinguish laterality, we certainly cannot exclude the possibility that neuronal processing within the amygdala may incorporate information about outcome identity, including laterality.

The involvement of the putamen in aversive conditioning was discovered much later than amygdala, and its function has been less clear. Since the putamen receives cortical somatotopic pain projections (Bingel et al., 2004), it is possible that it might have carried a consummatory or sensory-specific error signal (Roy et al., 2014; Torrecillos et al., 2014). However, the non-lateralised nature of the signal seen here instead provides good evidence to suggest that it is primarily part of a preparatory system.

Most significantly, the results provide a formal account of one of the roles of the cerebellum in pain. Previous research, including using human fMRI, has shown cerebellum responses to noxious stimuli, however defining a specific role in pain processing has been difficult (Moulton et al., 2010). Stimulation of the cerebellum can alter nociceptive thresholds and reflexes in animals (Saab and Willis, 2003), suggesting it may engage in pain modulation along with various brain stem structures involved in the cerebrocerebellar loop (Kelly and Strick, 2003; Moulton et al., 2010). Evidence from human studies indicate cerebellum may be activated by other processes related to, but not exclusive to pain sensory processing, for example motor withdrawal (Dimitrova et al., 2003), anticipation to pain (Ploghaus et al., 1999), and negative emotions (Singer et al., 2004). This has led to the proposal that the cerebellum may act as an integrator of various effector systems of pain such as sensorimotor integration, pain modulation, and affective processing (Moulton et al., 2010).

Our results provide evidence of an uncertainty-sensitive associative learning process for ipsilateral conditioned motor responses. Anatomically, the major activation was localised in the anterior lobe, bordering lobule V and VI, which concurs with the sensorimotor area of previous functional topographic studies (Stoodley and Schmahmann, 2009). Conditioned postural limb activation during electrical shock conditioning are known to depend on an intact anterior and superior cerebellum (Timmann et al., 2000). Electrical shocks, however, also recruit ascending proprioceptive fibres that project to cerebellum and support motor learning. Here, our use of thermal pain stimulation – which should selectively activate A-delta and c-fibres afferents, provides evidence of a primary nociceptive-driven learning process.

This result suggests parallels with eye-blink conditioning, a prototypical consummatory response. Anatomically, both animal and human lesion experiments have identified an association between lobe V and VI with impairment or disruption of eye-blink conditioning (Lavond and Steinmetz, 1989; Thieme et al., 2013). Computationally, cerebellar climbing fibre activity has been shown to represent prediction error magnitude (Schultz and Dickinson, 2000), from which associability might be calculated. Previous eye-blink studies have suggested a distinction between preparatory and consummatory learning processes. Although both excitatory and inhibitory conditioning on one eye can transfer to the other (Pearce et al., 1981); cues predicting unilateral airpuff do not block acquisition of contralateral blink responses, but they do block

autonomic responses (Betts et al., 1996). This suggests preparatory and consummatory learning systems are distinct, but interact.

Together, our data shows that the expression of learned pain behaviour is the sum of multiple, distinct neural processes. This has important implications for how we evaluate pain and its treatment, especially in animals where motor responses such as paw withdrawal and tail flick are the predominant outcome measures by which pain is inferred. Our data shows that different emitted responses may correspond to different underlying neural subsystems of responses to pain, which may help explain difficulties in translating animal-to-human results.

3.5 Tables

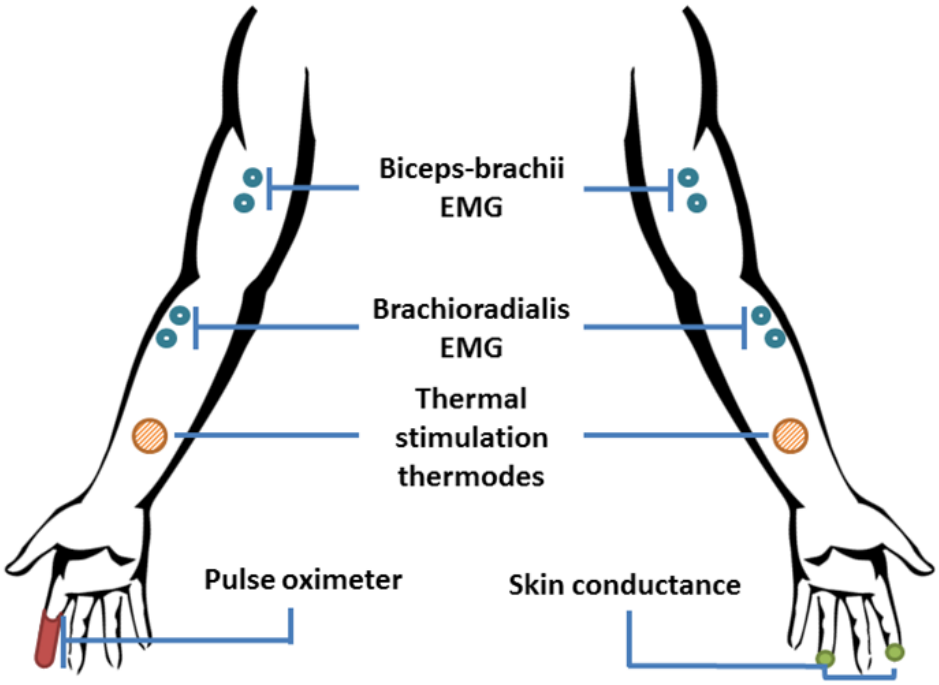
Table 3.1 Goodness of fit to SCRs for individual models (scanning, 3 sessions 15 subjects, 34 non-reinforced trials per session per subject. bold: winning model, V: model values, α: associabilities).

Model	Mean BIC	# of free parameters	Fitted free parameters
RW (V)	-193.34	4	α, β_0, β_1, σ
Hybrid (V)	-186.69	6	α, κ, η, β_0, β_1, σ
Hybrid (α)	-198.73	6	α, κ, η, β_0, β_1, σ
Hybrid (V + α)	-197.36	7	α, κ, η, β_0, β_1, β_2, σ

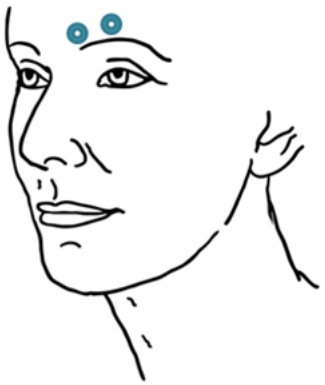
Table 3.2 Neuroimaging ROI analysis. Small volume correction (SVC) for multiple comparison within anatomically defined 8mm spherical masks of hypothesised structure coordinates (we did not find evidence of consummatory prediction errors on either side, see text).

Model / Region	peak p(FWE-corr)	cluster size	t statistics	Z statistics	x, y, z {mm}
Associability (preparatory)					
Amygdala (L)	0.089	13	2.96	2.57	-28 -6 -18
	0.112		2.78	2.44	-32 -4 -16
Amygdala (R)	0.038	36	3.58	2.97	30 -4 -18
Prediction error (preparatory)					
Ventral putamen (L)	0.002	43	5.77	4.06	-32 2 2
	0.003	43	5.67	4.02	-24 -4 -6
Ventral putamen (R)	0.003	91	5.50	3.95	26 -4 -6
	0.014		4.49	3.48	34 -4 -2
Associability (consummatory left)					
Cerebellum (L)	0.017	15	4.19	3.32	-28 -40 -30
Associability (consummatory right)					
Cerebellum (R)	0.011	7	4.55	3.51	32 -38 -30

3.6 Supplementary figures

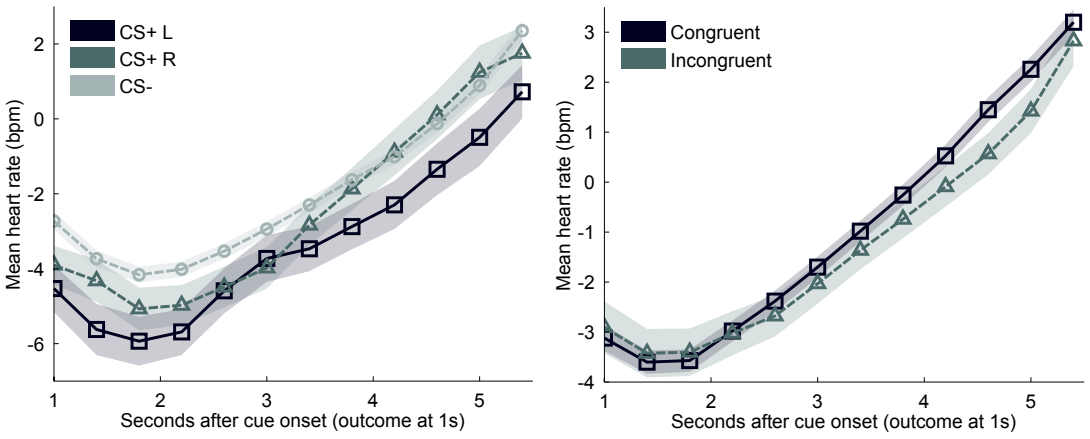


(a)

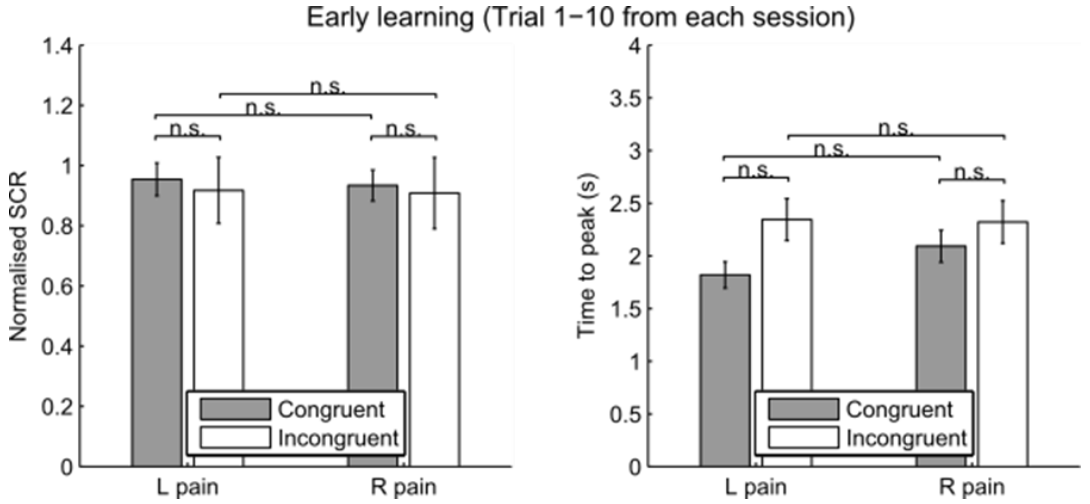


(b)

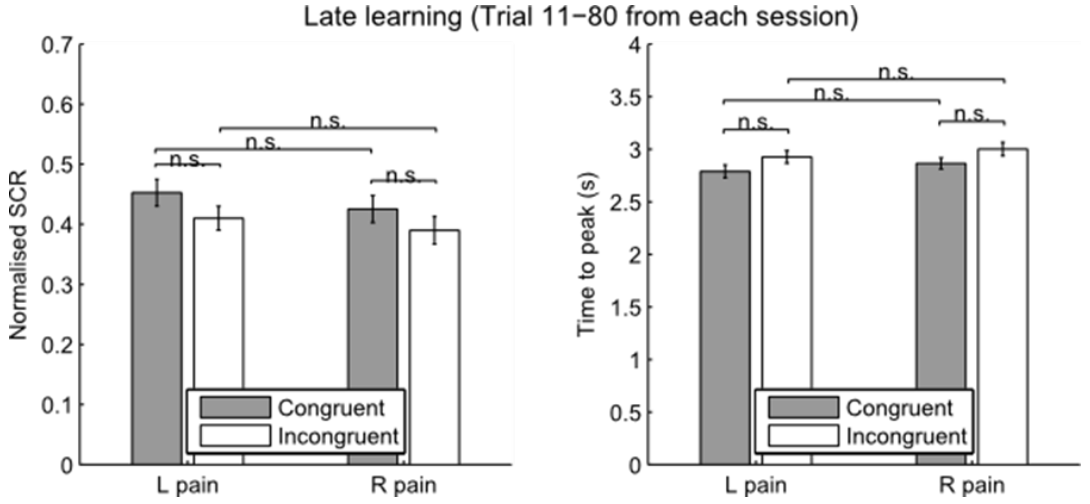
Fig. 3.4 Recording and stimulating apparatus placement. (a) Placement of bilateral upper-limb EMG electrodes (brachioradialis and biceps-brachii), pulse oximeter (heart rate), skin conductance electrodes, and stimulation thermodes. (b) Placement of facial EMG (corrugator) electrodes.


CS+ L
CS+ R
CS-
Mean heart rate (bpm)
Seconds after cue onset (outcome at 1s)
Congruent
Incongruent
Mean heart rate (bpm)
Seconds after cue onset (outcome at 1s)

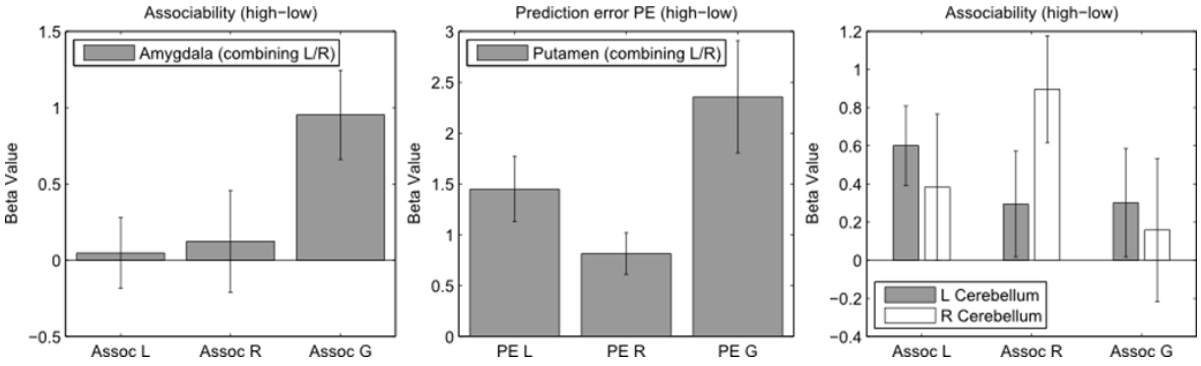
(a) (b)


Early learning (Trial 1−10 from each session)
Normalised SCR
n.s.
Congruent
Incongruent
L pain
R pain
Time to peak (s)
n.s.
Congruent
Incongruent
L pain
R pain

(c)


Late learning (Trial 11−80 from each session)
Normalised SCR
n.s.
Congruent
Incongruent
L pain
R pain
Time to peak (s)
n.s.
Congruent
Incongruent
L pain
R pain

(d)



(e)

Fig. 3.5 Supplementary results (heart rate was acquired only in 2 behavioural sessions). (a) Heart rate changes of 'unreinforced' trials in time window 1-5.5s after cue onset (0-1s not shown due to signal contaminated by thermal stimulation). No difference between CS+ L/R (all time points $p>0.4$), but both significantly different from CS− (vs. CS+ L all time points and vs. CS+ R 1-3s, $p<0.05$), with more pronounced deceleration (1-3s) and acceleration (4-5.5s).
(b) Heart rate changes in pain trials. Incongruent prediction trials showed less pronounced acceleration over congruent trials (but all sample points $p>0.1$). This is consistent with previous results when subjects were given novel (surprising) aversive stimuli (Bradley, 2009), and raises the possibility that some autonomic responses may not entirely blind to the laterality of pain, and hence may not be purely a preparatory response.
(c) SCR were separated into early (trials 1-10) and late (trials 11-80) learning periods in each session (scanning, 3 sessions). Early learning trials showed no significant laterality differences in either SCR magnitude (peak-to-peak magnitude at time window 0.5-5.5s after CS onset), or the time taken to rise to peak SCR (time difference from minimum to maximum SCR magnitude, n.s.: $p>0.05$).
(d) Late learning trials were similar to early learning trials.
(e) Beta values of ROIs identified sorted into high/low associability or prediction error, then aggregated into high minus low trials. Amygdala and putamen ROIs showed higher values for general system outputs (Assoc/PE G), while cerebellum ROIs showed differential high value for ipsilateral system outputs (Assoc/PE L/R). The overall BOLD activity pattern may not fully capture the dynamic predictions of the RL models, therefore we showed only the trend of activities and did not include full statistics in this analysis.

Chapter 4

Experiment 2: Comparing active and passive relief learning

Tonic pain after injury characterises a behavioural state that prioritises recovery. Although generally suppressing cognition and attention, tonic pain needs to allow effective relief learning so that the cause of pain can be reduced if possible. Previous evidence showed that uncertainty and attention modulate pain, however, whether and how this might relate to relief learning remains unknown. Here, we compared active and passive relief from tonic pain, implemented as instrumental and yoked Pavlovian escape conditioning tasks respectively, to investigate whether learning to obtain relief have influence over the motivational and hedonic aspects of pain. We showed that active relief-seeking involves a reinforcement learning process manifest by error signals observed in the dorsal putamen. Critically, this system also uses an uncertainty ('associability') signal detected in pregenual anterior cingulate cortex (pgACC) that both controls the relief learning rate, and endogenously modulates the level of tonic pain.

Adapted from Zhang et al. (2018a)

4.1 Introduction

Tonic pain is a common physiological consequence of injury and results in a behavioural state that favours quiescence and inactivity, prioritising energy conservation and optimising recuperation and tissue healing. This effect extends to cognition, and decreased attention is seen in a range of cognitive tasks during tonic pain (Crombez et al., 1997; Lorenz and Bromm, 1997). However, in some circumstances, this could be counter-productive, for instance if attentional resources were required for learning some means of relief or escape from the underlying cause of the pain. A natural solution would be to suppress tonic pain when relief learning is possible. Whether and how this is achieved is not known, but it is important as it might reveal central mechanisms of endogenous analgesia.

Two observations provide potential clues as to how a relief learning system might modulate pain. First, in some situations, perceived controllability has been found to reduce pain (Becker et al., 2015; Salomons et al., 2004, 2007; Wiech et al., 2014b), suggesting that the capacity to seek relief can engage endogenous modulation. Second, instructed attention has commonly been observed to reduce pain (Bantick et al., 2002). Therefore, it may be that attentional processes that are *internally* triggered when relief is learnable might provide a key signal that controls reduction of pain.

In general, learning involves distinct processes of prediction ('state learning') and control ('action learning') (Mackintosh, 1983), although relief learning during tonic pain has not been thoroughly investigated. But a quantitative model of relief learning – one that describes the computational processes that are implemented in learning centres in the brain – would allow interrogation of how an attentional process might operate to modulate tonic pain. In the case of phasic pain, learning can be described by reinforcement learning (RL) models – a well-studied computational framework for learning from experience. RL models describe how to predict the occurrence of inherently salient events, and learn actions to exert control over them (maximising rewards, minimising penalties) (Seymour et al., 2004). RL models aim to provide a *mechanistic* (beyond a merely descriptive) account of the information processing operations that the brain actually implements (Dayan and Abbott, 2001), and have a solid foundation in classical theories of animal learning (Mackintosh, 1983). In such models, an agent learns state or action value functions through outcomes provided by interacting with the world. These functions can be learned by computing the error between predicted and actual outcomes, and using the error to improve future predictions and actions (Sutton and Barto, 1998). Experimentally, the validity of these models can be tested by comparing how well different model-generated predictors fit the actual behavioural and/or neural data (O'Doherty et al., 2007).

During learning, attention is thought to boost learning of predictive associations and suppress other irrelevant information. Computationally, this can be achieved by estimating the

uncertainty as predictive associations are learned, and using this as a metric to control learning rates. Accordingly, high uncertainty corresponds to high attention and leads to more rapid learning (Dayan et al., 2000; AJ Yu and Dayan, 2005). One well-recognised way of formalising uncertainty in RL is by computing a quantity called the *associability*, which calculates the running average of the magnitude of recent prediction errors (i.e. frequent large prediction errors implies high uncertainty / associability). The concept of associability is grounded in classical theories of Pavlovian conditioning (the 'Pearce-Hall' learning rule, Holland and Schiffino, 2016; Le Pelley, 2004; Pearce and Hall, 1980), and provides a good account of behaviour and neural responses during Pavlovian learning (Boll et al., 2013; Li et al., 2011; Zhang et al., 2016). In this way, associability reflects a *computational* construct that captures aspects of the *psychological* construct of attention.

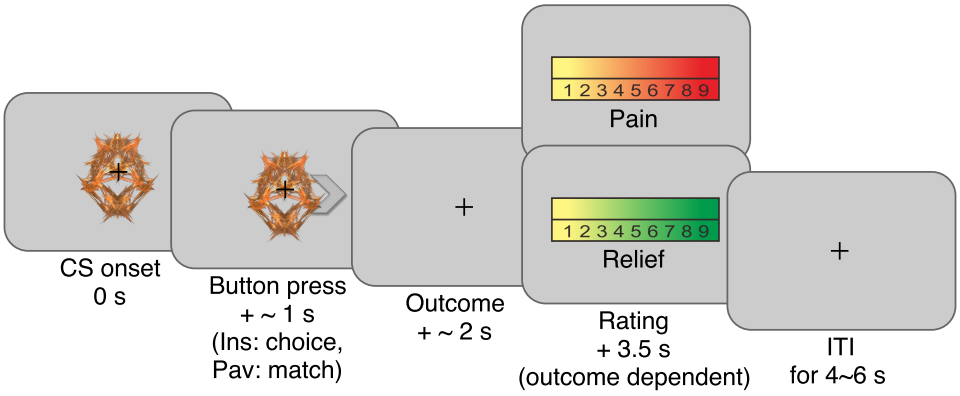
If it is the case, therefore, that attention can be understood as an uncertainty signal that drives learning during relief-seeking, it can then be tested whether it modulates tonic pain in parallel. Standard models of RL do not include any mechanism by which the subjective experience of outcomes is under control, although in principle endogenous modulation of tonic pain could arise from any component of the learning system, including an associability signal. Using an associability signal in this way would make intuitive sense, because it would reduce ongoing pain when requirement for learning was high.

The study has two goals: to delineate the basic neural architecture of relief learning from tonic pain (i.e. pain escape learning) based on a state and action learning RL framework; and to understand the relationship between relief learning and endogenous pain modulation i.e. to test the hypothesis that an attentional learning signal reduces pain. We studied behavioural, physiological and neural responses during an active (instrumental) and yoked passive (Pavlovian) relief learning paradigm. These tasks were designed to place a high precedence on error-based learning and uncertainty, as a robust test for learning mechanisms and dynamic modulation of tonic pain. Using a computationally motivated analysis approach, we aimed to identify whether behavioural and brain responses were well described as state and/or action RL learning systems and examined whether and how they exerted control over the perceived intensity of ongoing pain.

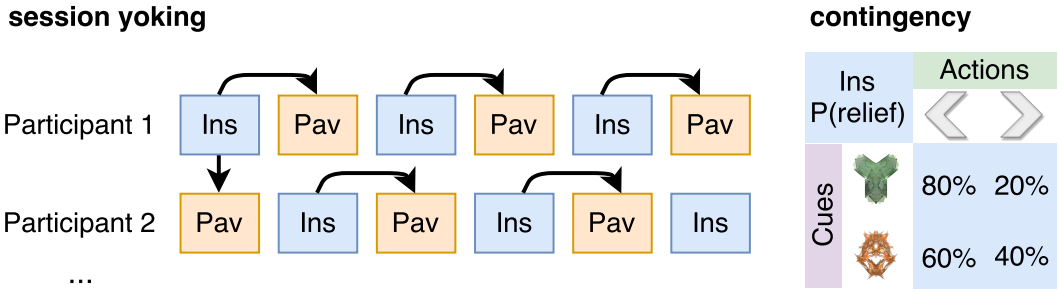
4.2 Methods

Subjects

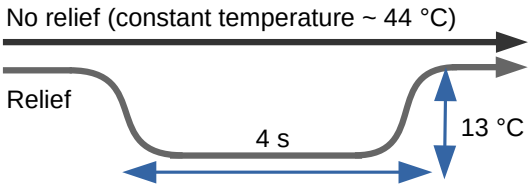
19 healthy subjects participated in the neuroimaging experiment (six female, age 26.1±5.1 years). All subjects gave informed consent prior to participation, had normal or corrected to



(a)



(b)



(c)

Fig. 4.1 Experiment paradigm. (a) Example trial, which was an instrumental relief learning task (Ins) with fixed relief probabilities, yoked with identical Pavlovian task (Pav) within subject. In instrumental trials, subjects saw one of two images ('cues') and then chose a left or right button press, with each action associated with a particular probability of relief. In the yoked Pavlovian session, subjects were simply asked to press button to match the action shown on screen (appearing 0.5s after CS onset). (b) Instrumental/Pavlovian session yoking and cue-outcome contingency, arrows represent identical stimulus-outcome sequence. Note in contingency table, left and right button presses were randomised for both actions and cues. (c) Relief and no relief outcomes, individually calibrated, constant temperature at around 44°C were used to elicit tonic pain, a brief drop in temperature of 13°C was used as relief outcome for 4s, but temperature did not change for the duration in no relief outcomes.

normal vision, and were free of pain conditions or pain medications. The experiment was approved by the Ethics and Safety committee of the National Institute of Information and Communications Technology (NICT), Japan.

Experimental design

Subjects participated in interleaved instrumental conditioning and yoked Pavlovian relief conditioning sessions in which they actively or passively escaped from pain, respectively (Fig 4.1). Tonic pain was maintained by constant thermal stimulation to the left inner forearm (see 'Stimulation' for details), and relief was induced by temporarily cooling the heat stimulus, which abolishes pain and causes a strong, pleasant sense of relief.

In the instrumental conditioning sessions, subjects learned to select responses (button press left or right) given visual stimuli (abstract fractal images on a computer screen) to try and obtain an outcome of a brief cooling (relief) period from the tonic painful heat (Fig 4.1a). There were two types of visual cue: an 'easy' cue with high probability of relief when paired with a particular response (80% relief chance with one of the button press responses and 20% chance with the other response), and a 'hard' visual cue with a lower probability of relief with a particular response (60%/40% relief chance for the two response actions). These different outcome probabilities were used to induce experimental variability in the uncertainty of relief prediction. On each trial, the visual cue (condition stimulus, CS) appeared on screen for 3s, during which subjects were asked to make the left or right button press response. An arrow corresponding to the chosen direction was superimposed on the cue after the decision was made until the 3s display period ended. The disappearance of the cue and response arrow was followed immediately by the outcome of a temporary decrease in temperature of the painful heat stimulus (temporary reduction of temperature by 13°C from the tonic level for 4s), or no change in temperature such that the constant pain continued straight on into the next trial (Fig 4.1c). The next trial started after a jittered inter-trial interval (ITI) of 4-6s (mean=5s) after outcome presentation concluded. There were 20 trials per session, with equal number of 'easy' and 'hard' cues (n=10 each). Each session lasted about 5 min.

The yoked Pavlovian conditioning task was identical to the instrumental task, except subjects did not have control over the outcomes through their responses. Instead, the sequence of cues and outcomes from the previous instrumental session were used (or the first instrumental session from the previous subject, for subjects who started with a Pavlovian session, Fig 4.1b), although subjects were not aware of the yoking process. A different set of fractal images was used for the yoked Pavlovian sessions, so learning on an instrumental session could not be transferral to its corresponding Pavlovian session. To control for motor responses in both sessions, subjects were asked to press the response button according to the randomised

indicator arrow shown, which appeared on screen 0.5s after CS presentation. This is common in neuroimaging studies of Pavlovian and instrumental learning, and it was clearly explained to subjects that these actions bore no relationship to outcomes.

Each subject repeated instrumental and yoked Pavlovian sessions three times (six in total). They were clearly instructed whether it was a Pavlovian or instrumental session. To remove any order confounds, the session order were alternated within and between subjects (i.e. order ABABAB, or BABABA), with half the subjects started with the instrumental and the other half with the Pavlovian task. A short break was taken every two sessions to allow the experimenter to change the location of the heat stimuli probe, to minimise effects of habituation/sensitisation across the whole experiment.

Subjective ratings of perceived trial outcomes (pain relief, or ongoing pain) were collected near the beginning, middle, and end of each session, in identical order for instrumental and its yoked Pavlovian counterpart. A linear rating scale appeared 3.5s after outcome presentation (0.5s overlap with relief duration if any), where the scale ranged from 0 (no pain at all) to 10 (unbearable pain) for no relief outcome (red scale in Fig 4.1a), and 0 (no relief at all) to 10 (very pleasant relief) for relief outcome (green scale). On average 8 pain and 8 relief ratings per paradigm i.e. total 16 for each subject were collected. Details are summarised in Table 4.1. Although it is the case that ratings are inherently subjective, their modulation reflects an objective process that may explain a component of this apparent subjectivity. This does raise the issue of whether the subjective relief ratings influence the outcome values when learned in the RL model, but this (presumably subtle) effect is something that is beyond the experimental power of this experiments to resolve.

Stimulation

Painful tonic thermal stimuli were delivered to the subject's skin surface above the wrist on the left inner forearms, through a contact heat-evoked potential stimulator (CHEPS, Medoc Pathway, Israel). The 27mm diameter CHEPS thermode is capable of rapid cooling at 40°C/s, which made rapid temporary pain relief possible in an event related design.

The temperature of painful tonic stimuli was set according to the subject's own pain threshold. Before the task, two series of 6 pre-set temperatures were presented in random order (set 1: mean $\pm$ standard deviation (std) 43.7$\pm$1.7°C; set 2: 44.6$\pm$0.6°C), with each temperature delivered for 8s, after which the subject determined whether the stimulation period was painful or not (ISI=8s). The higher of the two lowest painful temperatures from the two tests was used as the tonic stimulation temperature.

Final temperature used was 44.3$\pm$0.2°C, supported by previous evidence that a similar temperature/location for prolonged exposure (>7s) corresponds to medium pain in ratings

(Atlas et al., 2010). Relief temperature was set constant at 13°C below threshold temperature for all subjects. Thermode temperature stayed at threshold level unless relief outcome occurs, where rapid cooling at 40°C/s to relief temperature and stayed for 4s, before ramping back to threshold level at 40°C/s.

Physiological measures

Skin conductance responses (SCRs) were measured using MRI-compatible BrainAmp ExG MR System (Brain Products, Munich, Germany) with Ag/AgCl sintered MR electrodes, filled with skin conductance electrode paste.

SCR data was recorded on volar surfaces of distal phalanges of the second and fourth fingers on the left hand (tonic pain side with thermode attached). The signals were collected using BrainVision software at 500Hz with no filter.

Off-line processing and analysis were implemented in MATLAB (The MathWorks Inc., Natick, MA, USA), with the PsPM toolbox (http://pspm.sourceforge.net/). The data was down-sampled to 10Hz, band-pass filtered at 0.0159-2Hz (1st order Butterworth). Given the variable nature of SCR onset and duration in a learning experiment, the non-linear model in PsPM was used. Boxcar and delta regressors were constructed at cue onset (duration=3s, cue presentation), and at outcome onset (duration=3s, during relief period) respectively. These regressors were convolved with the canonical skin conductance response function, to estimate event-related response amplitude, latency, and dispersion (only SCR amplitude were used in modelling).

Sessions with more than 20% trials (4 out of 20 trials) with cue-evoked SCR amplitude below the threshold of 0.02 were labelled as not having enough viable event related SCRs. 15 subjects and 50 sessions remained. Trial SCRs were log-transformed within subject before model fitting.

Other behavioural measures

Trial-by-trial choice data (button press indicating choices) and reaction times (length of time taken from CS onset to choice button press) of subjects were recorded as part of behavioural measurements. All behavioural data including raw SCRs, choices, and ratings can be found in the supplementary data attachment online (Zhang et al., 2018a).

Computational learning models

To capture relief learning we fitted behavioural responses using different learning models from previous studies (Table 4.2). Free Energy (F) is the variational Bayesian approximation of a

model's marginal likelihood, and the sum of F for all participants provides model absolute fit evaluation. Actual model comparison was conducted based on random-effect analysis. For instrumental learning, the reinforcement of subjects' responses based on experienced relief can be modelled using reinforcement learning models (Sutton and Barto, 1998). For Pavlovian learning, physiological responses can be used for model fitting (Boll et al., 2013; Li et al., 2011; Zhang et al., 2016).

Win-Stay-Lose-Shift (WSLS) model

WSLS assumes a subject has fixed pseudo Q values for each state-action pair, where a relief outcome always produces a positive value for the chosen state-action pair (i.e. win-stay), while the remaining state-action combinations had negative values (i.e. lose-shift). A no relief outcome flipped the sign of all values. Two free parameters p_1 and p_2 ($0 \leq p_{1,2} \leq 1$) scaling the pseudo Q values for the two cues presented were used in model fitting, which were assumed fixed through out the experiment but varied for individuals.

Temporal Difference (TD) model

The predicted state-action value Q given particular state s and action a between successive trials is updated using an error-driven delta rule with learning rate α ($0 \leq \alpha \leq 1$) (Gläscher et al., 2010; G Morris et al., 2006; Sutton and Barto, 1998):

$$Q_{t+1}(s,a) = Q_t(s,a) + \alpha \cdot (r_t - Q_t(s,a)) \tag{4.1}$$

where r_t is the outcome of the trial (relief=1, no relief=0). The probability of choosing action a from a set of all available actions $A_s \in \{a,b,c...\}$ in trial t is modelled by a softmax distribution,

$$p(a|s) = \frac{exp(\tau \cdot Q_t(s,a))}{\sum_{b \in A_s} exp(\tau \cdot Q_t(s,b))} \tag{4.2}$$

where τ is the inverse temperature parameter governing the competition between actions ($\tau > 0$).

Rescorla-Wagner (RW) model

For Pavlovian learning, where choice decisions are not available, the standard TD model based on the Rescorla-Wagner learning rule updates the state value $V(s)$ for a given cue following a

similar prediction error based updating process:

$$V_{t+1}(s) = V_t(s) + \alpha \cdot (r_t - V_t(s)) \tag{4.3}$$

Hybrid model

The hybrid model incorporated an associability term as a changing learning rate for a standard TD model in state value learning (Le Pelley, 2004; Li et al., 2011). The associability term is also referred to as Pearce-Hall associability, an equivalent measure of uncertainty, which is modulated by the magnitude of recent prediction error (Pearce and Hall, 1980). The varying learning rate can be used in Pavlovian state-learning:

$$V_{t+1}(s) = V_t(s) + \kappa \cdot \alpha_t(s) \cdot (r_t - V_t(s)) \tag{4.4}$$

$$\alpha_{t+1}(s) = \eta \cdot |r_t - V_t(s)| + (1-\eta) \cdot \alpha_t(s) \tag{4.5}$$

where η, κ are free parameters limited within the range of [0,1].

The model can also be extended to instrumental action-learning:

$$Q_{t+1}(s,a) = Q_t(s,a) + \kappa \cdot \alpha_t(s,a) \cdot (r_t - Q_t(s,a)) \tag{4.6}$$

$$\alpha_{t+1}(s,a) = \eta \cdot |r_t - Q_t(s,a)| + (1-\eta) \cdot \alpha_t(s,a) \tag{4.7}$$

Model fitting and comparison

Model fitting

Model fitting was performed with the Variational Bayesian Analysis (VBA) toolbox (https://mbb-team.github.io/VBA-toolbox/). The toolbox seeks to optimise free energy within the Bayesian framework, analogous of maximum likelihood. Behavioural data (choices, SCRs) were fitted separately for each individual resulting in different sets of parameters, and model fitting performance was measured by aggregating individual subject fitting statistics. The mean of all subject parameters were used to generate regressors for fMRI analysis following conventions (Table 4.3).

The VBA toolbox takes in an evolution function that describes the learning model (e.g. value updating rule), and an observation function that describes response mapping (e.g. softmax action selection). For choice fitting, data was split into multiple sessions to allow between-session changes in observation function parameters, but evolution function parameters and initial states were fixed throughout all sessions.

For SCR fitting, multi-session split were the same as choice fitting. The first two trials from each session were excluded from fitting to avoid extreme values from startle effects, which also served to reduce the confound from general habituation of SCRs. Trials with insufficient event-related responses were also excluded (see 'Physiological measures' above). The observation function for SCR fitting was:

$$gx = \text{predictor} + b \tag{4.8}$$

with b as a free parameter. The predictor was not scaled to avoid overfitting.

Parameter prior setting for models followed previous studies. TD, RW and hybrid models all have initial values as 0, and initial associability as 1. All free parameters of the evolution function were assumed to have prior variance of 1, with the exception of SCR fitting where it was assumed to be 0.05 to reduce flexibility. Fitted learning parameters were similar to previous studies (Li et al., 2011; Zhang et al., 2016).

Model comparison

Model comparison was implemented with random-effect Bayesian model selection in the VBA toolbox. The best fitted model for each individual was allowed to vary, and model frequency in population (i.e. in how many subjects the model was the best-fit model) and model exceedance probability (how likely the model is more frequent than other models compared) were estimated from model fitting evidence (free energy from learning models in choice and SCR fitting, or log likelihood from regression models in rating fitting) (Daunizeau et al., 2014; Stephan et al., 2009).

We also calculated the protected exceedance probabilities, which corrected for the possibility that model differences observed are due to chance (Rigoux et al., 2014). Results were shown in figure supplements in the same way as in the original exceedance probabilities in the results section (Fig 4.5). See http://mbb-team.github.io/VBA-toolbox/wiki/BMS-for-group-studies/#rfx-bms for details of its calculation.

fMRI acquisition

Neuroimaging data was acquired with a 3T Siemens Magnetom Trio Tim scanner, with the Siemens standard 12 channel phased array head coil.

Functional images were collected with a single echo EPI sequence (repetition time TR=2500ms, echo time TE=30ms, flip angle=80, field of view=240mm), 37 contiguous oblique-axial slices (voxel size 3.75 $\times$ 3.75 $\times$ 3.75 mm) parallel to the AC-PC line were acquired. Whole-brain

high resolution T1-weighted structural images (dimension 208 $\times$ 256 $\times$ 256, voxel size 1 $\times$ 1 $\times$ 1 mm) using standard MPRAGE sequence were also obtained.

fMRI preprocessing

Functional images were slice time corrected using SPM12 (http://www.fil.ion.ucl.ac.uk/spm/software/spm12/) with individual session's slice timing output by the scanner. Resulting images were then preprocessed using the fmriprep software (build date 09/03/2017, freesurfer option turned off, https://github.com/poldracklab/fmriprep), a pipeline that performs motion correction, field unwarping, normalisation, field bias correction, and brain extraction using a various set of neuroimaging tools available. The normalised images were smoothed using a Gaussian kernel of 8mm using SPM12. The confound files output by fmriprep include the following signals: mean global, mean white matter tissue class, three FSL-DVARS (stdDVARS, non-stdDVARS and voxel-wise stdDVARS), framewise displacement, six FSL-tCompCor, six FSL-aCompCor, and six motion parameters (matrix size 24 $\times$ number of volumes).

fMRI GLM model

All event-related fMRI data was analysed with GLM models constructed using SPM12, estimated for each participant in the first level. Model generated signals used as parametric modulators were generated with one set of group-mean model parameters, obtained with behavioural data fitting as described. We used the mean of the fitted parameters from all participants in the imaging analysis as this provides the most stable estimate of the population mean (taking into account the fact that individual fits reflect both individual differences and noise). For completeness, however, we also ran the analyses with individually fitted values, which led to similar results (i.e. no change in significance level of each result). All regressors were convolved with a canonical hemodynamic response function (HRF). We also include regressors of no interest to account for habituation and motion effects. Specifically, to keep the analyses from the two relief learning studies (Experiment 2 and 3) the same, we included the following habituation regressors to regress out potential change in tonic pain perception simply due to prolonged stimulation: the number of trials since last receiving a relief outcome ('Relief' term in rating regression model in Equation 5.7), and the log of trial number within session ($log(\text{Trial})$ term). The resulting GLM estimates were entered into a second-level one-sample t-test for the regressors of interest to produce the random-effect statistics and images presented in the Results section.

TD softmax

Regressors of interest:

- CS onset (duration=3s, cue presentation): Q values of chosen cue,
- US onset (duration=3s): prediction error,

Regressors of no interest:

1. CS onset (duration=3s, cue presentation): number of trials since last relief,
2. CS onset (duration=CS onset to US offset, entire trial exclude ITI): within session log trial number,
3. choice press (duration=0),
4. rating press (duration=rating duration),
5. CS offset (duration=0),
6. 24 column confounds matrix output by fmriprep.

Hybrid model associability

Regressors of interest:

- choice press time (duration=0, cue button press): associability (generated for individual session with new V_0/A_0 to match SCR fitting procedure),

Regressors of no interest:

- same as GLM above,
- adding relief onset (duration=0),
- removing choice press regressor.

We note for completeness that it is theoretically possible to model the learning process as a continuously valued function that exactly matches the time-course of the temperature changes. In the context of the current study, the effect of this would be largely orthogonal to the experimental manipulations. However, representation of the baseline temperature as a continuous function is clearly important in real-life contexts in which the baseline level determines homeostatic motivation and phasic reward functions (Morville et al., 2018), and hence future studies could directly manipulate this.

For multiple comparison, we used anatomical binary masks generated using the Harvard-Oxford Atlas (Desikan et al., 2006) for small volume correction. Atlases are freely available with the FSL software (https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases). We thresholded the probability maps at 50%, focusing on ROIs defined a priori (learning related: amygdala, accumbens, putamen, caudate, pallidum, ventromedial prefrontal cortex [VMPFC], dorsolateral prefrontal cortex [DLPFC]; controllability-induced analgesia related: cingulate gyrus – anterior division, insular cortex, VLPFC). We used the frontal medial cortex for VMPFC, the frontal orbital cortex for VLPFC, and the middle frontal gyrus for DLPFC respectively. We reported results with $p<0.05$ (FWE cluster-level corrected). Masks were applied separately, not combined (Table 4.4).

fMRI model comparison

To determine whether state-based and action-based learning involve the same brain regions during instrumental learning, we used Bayesian model selection (BMS) with the instrumental sessions imaging data. We ran Bayesian first level analysis using two separate GLMs containing the PE signals from TD and Hybrid models (at US onset time, durations=3s) using unsmoothed functional imaging data, with the same regressors of no interest as other GLMs described. To reduce computation time, this was restricted to voxels correlated to PEs from previous parametric modulation analysis results, within a mask of conjunction clusters from TD and hybrid PE analysis (cluster formation at $p<0.01$, $k<5$). Resulting log-model evidence maps produced from each model for individual participant were first smoothed with a 6mm Gaussian kernel, then entered into a random-effect group analysis (Stephan et al., 2009). Voxel-wise comparison between models produced posterior and exceedance probability maps to show whether a particular brain region is better accounted for by one model or the other. Posterior probability maps were overlaid on subject-averaged anatomical scans using MRIcroGL (https://www.nitrc.org/projects/mricrogl/).

Axiom analysis for prediction errors

To determine whether ROI activations to prediction errors were responding outcomes or prediction errors, we carried out ROI axiomatic analysis (Roy et al., 2014; Rutledge et al., 2010). For neural signals in an ROI to be considered as relief prediction errors, it should have: (i) consistent outcome ordering regardless of expectations (response to relief is larger than no relief), (ii) consistent expectation ordering regardless of outcome (response decreases with increased predicted relief). Although the axiomatic analysis is useful for delineating outcome and prediction responses in previous reward or aversive PE studies, the continued presence of

tonic pain in our study differs from the 'no stimulation' conditions in these studies. Therefore we did not test for the third axiom of no surprise equivalent (section 1.3), and showed the overall BOLD activity pattern without including full statistics in this analysis.

Trials were separated into relief or no relief outcomes, then into equal-size bins of ascending sorted expected relief values, calculated from TD model as we were primarily interested in instrumental sessions. A median split for values was chosen to ensure sufficient trials were included from each session (relief outcomes per session: 10.8 trials). This produced 4 regressors (2 outcomes $\times$ 2 value bins), to be estimated at outcome time (duration=3s) when PE was generated. GLMs include button presses for choice or rating, and movement related regressors of no interest mentioned above. ROI masks of 8mm spheres were generated from peak coordinates from TD model prediction error exceedance probability map calculated by BMS above (ventral and dorsal striatum, amygdala, VMPFC and DLPFC).

4.3 Results

Behavioural results

Choice In instrumental learning, participants can learn which actions maximise the chance of relief. We assessed the ability of RL models to explain subjects' choice data, in comparison to a simple WSLS decision-making rule. We compared two basic RL models that have been widely studied in neurobiological investigations of reward and avoidance – a TD action learning model with a fixed learning rate, and a version of the TD model with an adaptive learning rate based on action associabilities (hybrid TD model). The associability reflects the uncertainty in the action value, where higher associability indicates high uncertainty during learning, and is calculated based the recent average of the prediction error magnitude for each action. In a random-effects model comparison procedure (Daunizeau et al., 2014), we found that choices were best fit by the basic TD model (model frequency=0.964, exceedance probability=1, Fig 4.2a). Thus, there is no evidence that associability operates directly at the level of *actions*.

Skin conductance responses (SCR) To investigate physiological indices of learning, we examined trial-by-trial SCRs during the 3s cue time, before outcome presentation. SCRs obtained in instrumental sessions were higher compared to yoked Pavlovian sessions (Fig 4.2b, n=15, see Methods for session exclusion criteria, paired t-test $T(14)=2.55$, $p=0.023$), with the average SCR positively correlated between paradigms across individuals (Pearson correlation $\rho=0.623$, $p=0.013$, $n=15$). Raw traces and cue-evoked responses of SCRs can be found in Figure supplements Fig 4.4a and 4.4b.

In Pavlovian aversive (fear) learning, SCRs have been shown to reflect the associability of Pavlovian predictions (Boll et al., 2013; Li et al., 2011; Zhang et al., 2016). Here, associability is calculated as the mean prediction error magnitude for the *state* (i.e. regardless of actions) (Le Pelley, 2004). In instrumental learning, Pavlovian learning of state-outcome contingencies still proceeds alongside action-outcome learning, distinct from instrumental choices, so Pavlovian state-outcome learning can be modelled in both instrumental and Pavlovian sessions. Consistent with previous studies of *phasic* pain, model-fitting revealed that the hybrid model with state-based associability best fit the SCR data in both Pavlovian and instrumental sessions (Fig 4.2c and Fig 4.2d, instrumental sessions: model frequency=0.436, exceedance probability=0.648, Pavlovian sessions: model frequency=0.545, exceedance probability=0.676), when tested against a competing simple Pavlovian Rescorla-Wagner model (akin to a TD model with only one state and a fixed learning rate). However, using the more stringent protected exceedance probability analyses, the advantage of associability over other models were less conclusive (Fig 4.5). Together with the choice results, these analyses suggest that subjects use an associability-based RL mechanism for learning state values during both Pavlovian and instrumental pain escape, and a non-associability based RL mechanism for learning action values in instrumental sessions. This divergence in learning strategies indicates that parallel learning systems coexist, consistent with their dissociation under dopamine antagonists (Dickinson et al., 2000). It is likely that the two systems differ in their way of incorporating information about uncertainty in learning, as well as the nature of their behavioural responses.

Ratings Subjective ratings of pain and relief were taken intermittently after outcomes during the task, to explore how pain modulation might depend on relief learning. Ratings were taken on a sample of trials, so as to minimise disruption of task performance (see Table 4.1 for timing details). Based on the fact that both controllability and attention are implicated in endogenous control, we hypothesised that pain would be reduced when the state-outcome associability was high, reflecting an attentional signal associated with enhanced learning. However, other types of modulation are possible. For instance, pain might be non-specifically reduced in instrumental, versus Pavlovian learning, reflecting a general effect of instrumental controllability. Alternatively, pain might be reduced by the expectation of relief that arises during learning, as it is known that conditioning alone can support placebo analgesia responses (Colloca et al., 2008) (although the extent to which this occurs might depend on the acquisition of contingency awareness during learning) (Locher et al., 2017; Montgomery and Kirsch, 1997). In this case, pain would be positively correlated with the relief prediction error, since it reports the difference between expectation and outcome.

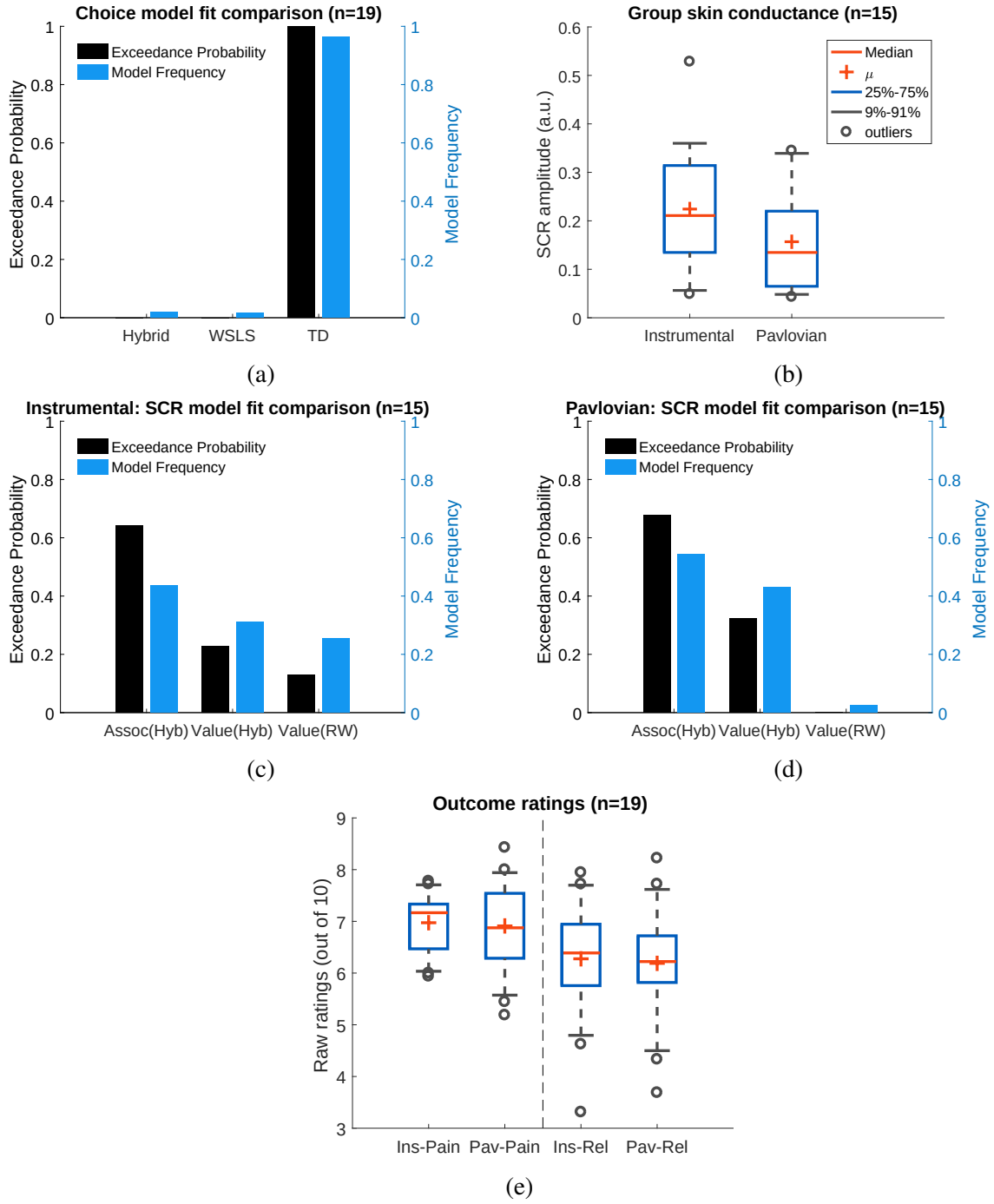


Fig. 4.2 Behavioural results. (a) Choice-fitted model comparison, TD model fit instrumental sessions choices best (TD: action-learning model with fixed learning rate, Hybrid: action-learning model with associability as changing learning rate, WSLS: win-stay-lose-shift model). Model frequency represents how likely a model generate the data given a random participant, while exceedance probability estimates how one model is more likely compared to others (Stephan et al., 2009). (b) Instrumental vs Pavlovian sessions SCRs (n=15, sessions with over 20% trials <0.02 amplitude excluded). (c) Associability from hybrid model fitted trial-by-trial SCRs best in instrumental sessions (Assoc: associability, Hyb: hybrid model, RW: Rescorla-Wagner model). (d) Associability also fitted SCRs from Pavlovian sessions best. (e) Both pain and relief ratings did not differ significantly between instrumental and Pavlovian sessions (Participants' ratings were averaged for each of the four categories shown, mean=8 ratings per person per category, figure legend identical to Fig 4.2b).

To test these competing hypotheses, we first compared the mean ratings of both pain (following a 'no relief' outcome) and relief (following a relief outcome) between Pavlovian and instrumental sessions, and found no significant difference (Mean ± standard error (SEM), n=19, instrumental pain: 6.97±0.13, Pavlovian pain: 6.91±0.20, instrumental relief: 6.46±0.24, Pavlovian relief: 6.33±0.27, between paradigm paired t-test both ratings p>0.5, Fig 4.2e). Hence, there is no support for a general effect of instrumental controllability on subjective pain and/or relief experience. We noted that mean pain and relief ratings were correlated with each other across individuals (ratings averaged across paradigms, Spearman's correlation ρ=0.73, p<0.001), indicating that higher perceived tonic heat pain was associated with higher cooling-related relief.

Next, we correlated pain ratings with the state-based associability and TD prediction error. In accordance with our hypothesis, in instrumental sessions associability was found to be negatively correlated with pain ratings (mean Spearman's $\bar{\rho}$=-0.177, one-sample t-test of Fisher's z-transformed correlation coefficients T(18)=-2.125, p=0.048). In Pavlovian sessions, however, we did not find a correlation ($\bar{\rho}$=-0.114, T(18)=0.758, p=0.458). There was no significant interaction between associability and paradigm (repeated measure ANOVA F(1,18)=1.247, p=0.279). This suggests that although associability is associated with pain modulation, this effect is not necessarily *specific* to instrumental sessions.

We found that the prediction errors were negatively correlated with pain ratings in Pavlovian sessions ($\bar{\rho}$=-0.356, T(18)=-3.198, p=0.005), but not instrumental sessions ($\bar{\rho}$=-0.154, T(18)=0.720, p=0.481). That is, when relief was omitted (i.e. as was always the case on the pain rating trial), a larger frustrated (i.e. negative) relief prediction error was associated with an increase in pain – in contrast to the prediction of a placebo expectation hypothesis. Interaction between prediction error and paradigm was not significant (repeated measure ANOVA F(1,18)=3.706, p=0.0702). Finally, we also looked at relief ratings, but failed to find any significant correlation with either associability or prediction error in either instrumental or Pavlovian sessions.

Neuroimaging results

The behavioural findings support the hypothesis that an associability signal that arises during state-based learning is associated with reduction of pain. Therefore, we then sought to identify (i) neural evidence for an error-based relief learning process, and (ii) the neural correlates of the associability signal associated with tonic pain modulation. We implemented the TD action-learning model and associability-based hybrid TD state-learning model as determined from the behavioural data, using group-mean parameters (learning rate in TD model, and free parameter

κ and η in hybrid TD model) to re-estimate trial-by-trial prediction errors/associability values for each subject as parametric modulators of fMRI BOLD time-series in general linear models.

Prediction errors The prediction error represents the core 'teaching' signal of the reinforcement learning model, and we specified *a priori* regions of interest based on the areas known to correlate with the prediction error in previous reinforcement learning studies of pain and reward (ventral and dorsal striatum, VMPFC, DLPFC, and amygdala (FitzGerald et al., 2012; Garrison et al., 2013; Seymour et al., 2005)).

First, we looked for brain responses correlated with the action prediction error from the TD model in instrumental sessions. This identified responses in bilateral putamen, bilateral amygdala, left DLPFC, and VMPFC (Fig 4.3a, Table 4.4).

Since action-outcome learning and state-outcome learning co-occur during instrumental sessions, we next modelled the state prediction error from the hybrid model in a separate regression model. In instrumental sessions, this revealed responses in similar regions to the TD action prediction error: in the striatum, right amygdala and left DLPFC (figure not shown, Table 4.4), consistent with the fact that state and action prediction errors are highly correlated.

To test which regions were better explained by each, we conducted a Bayesian model selection within the prediction error ROIs (a conjunction mask of correlated clusters to both prediction error signals). This showed that the action-learning TD model had higher posterior and exceedance probabilities in the dorsal putamen and VMPFC (Fig 4.3b warm colour clusters). The state-learning (hybrid) model better explained activities in the amygdala, ventral striatum, and DLPFC (Fig 4.3b cool colour clusters). Applying the same hybrid model prediction error signal in Pavlovian sessions only identified much weaker responses that did not survive multiple correction, in regions including the left amygdala (figure not shown).

To further illustrate the nature of the outcome response, we calculated a median split of the preceding cue values (based on the TD model), and looked at the outcome response for relief and no-relief outcomes. A prediction error response should be (i) higher for relief trials, and (ii) higher when the preceding cue value was low (i.e. when relief was delivered when it was not expected) (Roy et al., 2014). As illustrated in Fig 4.3c, this 'axiomatic' analysis reveals some features of the prediction error, but lacks the resolution to illustrate it definitively.

Associability Since the behavioural data showed that the state-based associability correlated negatively with tonic pain ratings, we examined BOLD responses correlated with trial-by-trial associability from the hybrid model, by using the associability as a parametric regressor at the choice time (see Methods for details of GLMs). We specified *a priori* ROIs according to regions previously implicated in attention and controllability-related endogenous analgesia, notably

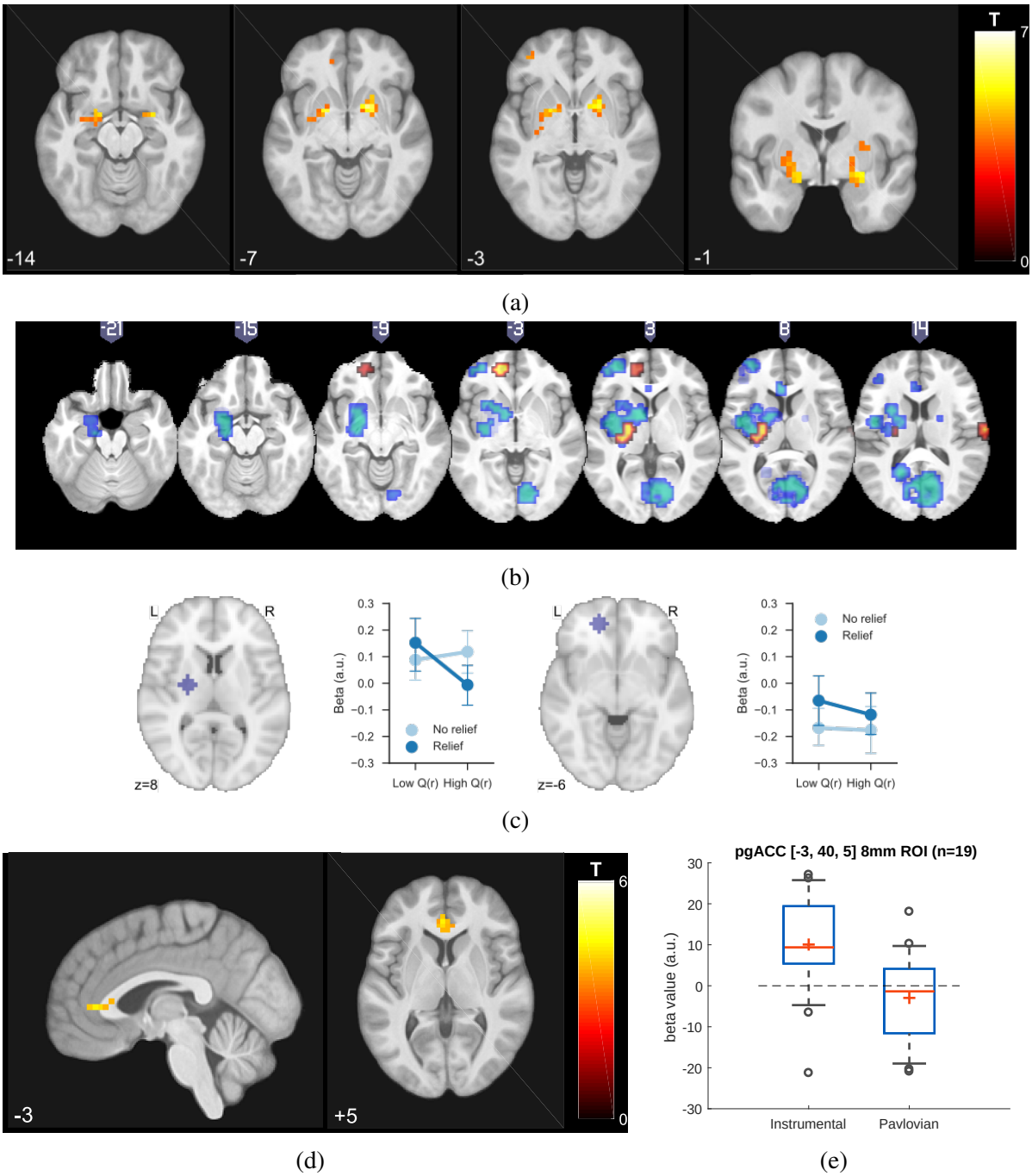


Fig. 4.3 Neuroimaging results, shown at p<0.001 uncorrected: (a) TD model prediction errors (PE) as parametric modulators at outcome onset time (duration=3s). (b) Model PE posterior probability maps (PPMs) from group-level Bayesian model selection (BMS) within PE cluster mask, warm colour: TD model PE, cool colour: hybrid model PE (shown at exceedance probability P>0.7), (c) Axiomatic analysis of hybrid model PEs in instrumental sessions, ROIs were 8mm spheres from BMS peaks favouring TD model PEs, in left putamen and VMPFC. (d) Associability uncertainty generated by hybrid model, as parametric modulators at choice time (duration=0), in instrumental sessions. (e) Comparing pgACC activations across instrumental/Pavlovian paradigms, ROI was 8mm sphere at [-3, 40, 5], peak from overlaying the pgACC clusters from Experiment 2 and 3 (Fig 5.6a).

pregenual anterior cingulate cortex (pgACC), posterior insula and ventrolateral prefrontal cortex (VLPFC) (Salomons et al., 2007; Wiech et al., 2006); and associability (amygdala) (Boll et al., 2013; Li et al., 2011; Zhang et al., 2016).

In instrumental sessions we found correlated responses only in the pgACC (Fig 4.3d, Table 4.4, MNI coordinates of peak: [-2, 37, 5]). No significant responses were observed in Pavlovian sessions. Fig 4.3e illustrates individual subjects' beta values extracted from an 8mm diameter spherical ROI mask built around peak coordinates [-3, 40, 5]. Instrumental sessions had higher response magnitude in the pgACC compared to Pavlovian sessions across subjects (Instrumental sessions: one-sample t-test against 0 $T(18)=3.746$, $p=0.0015$, Pavlovian sessions: one-sample t-test against 0 $T(18)=-1.230$, $p=0.235$, paired t-test for instrumental versus Pavlovian $T(18)=3.317$, $p=0.0038$).

4.4 Discussion

In summary, the results indicate that (i) relief action learning is well described by a RL (TD) learning process, with action prediction error signals observed in the dorsal putamen, (ii) that state-outcome learning proceeds in parallel to action-outcome learning, and can be described by an associability-dependent hybrid TD learning mechanism, and (iii) that this state associability modulates the level of ongoing tonic pain during instrumental learning, with associated responses in the pgACC.

Tonic pain characterises the altered behavioural homeostatic state after injury, in which motivation becomes orientated towards recovery. Whereas acute (phasic) pain functions as a teaching signal about harmful events, the physiological function of tonic pain is different: it provides a new affective baseline that endows relief with positive value as a behavioural goal, yet it also consumes other cognitive and behavioural functions to promote recuperation. Our results provide an explanation to this apparently paradoxical relationship: we identified a learning circuit that governs relief learning and decision-making, with its activities correlating with the information available during active learning, where higher uncertainty in the environment corresponded to lower subjective pain ratings. In so doing, it solves the problem of balancing tonic pain with the requirement to actively learn about behaviour that could lead to relief.

Endogenous modulation of tonic pain could in principle arise from any component of the learning system, including the associability signal. Associability has its theoretical underpinnings in classical theories of associative learning and attention (i.e. the Pearce-Hall theory, Pearce and Hall, 1980), and its mathematical implementation here is as an approximate uncertainty quantity derived from computing the running average of the magnitude of the prediction error (Le Pelley, 2004; Sutton, 1992). This uncertainty signal effectively captures

how predictable the environment is: when uncertainty is high (because of lots of recent large prediction errors), it increases the speed of acquisition through increasing the learning rate, and so accelerates convergence to stable predicted values. It is therefore an effective attention-like signal for mediating endogenous analgesia, because it selectively facilitates active relief seeking by suppressing pain only when it is necessary. This conception of the role of uncertainty in pain may explain why uncertainty has been shown to enhance phasic pain (Yoshida et al., 2013), where pain acts as the signal to drive learning, and suppresses tonic pain, where pain acts to reduce general cognition. In both instances, the role of uncertainty and attention is to facilitate learning.

The behavioural data modelling results suggest potentially dissociable learning systems. The instrumental choice data was best fitted by the TD model without an action-based associability term, while the SCR data was best fitted by the hybrid TD model with a state-based associability term. This distinction between the learning of prediction and control, captured as state and action learning respectively, has been observed in several previous studies (Boll et al., 2013; Gläscher et al., 2010; Li et al., 2011; G Morris et al., 2006; Zhang et al., 2016), and likely reflects the distinction between instrumental and Pavlovian learning systems. Indeed, Pavlovian and instrumental systems learn different responses – conditioned responses (e.g. autonomic responses) and instrumental actions respectively, each of which have independent biological functions (Dickinson and Balleine, 2002). While we did not find evidence that associability was used in determining actions (in keeping with previous reports), it was used for learning other forms of conditioned response. Although captured as independent models here, it is possible that different learning systems interact under appropriate circumstances (Holmes et al., 2010), however, that will require probing with a different task.

Compared to instrumental sessions, the results from Pavlovian relief learning sessions were less informative. While the SCR data model fitting were largely consistent with previous findings, pain ratings were found to increase with larger frustrated relief prediction errors, similar to nocebo effects. Unlike the associability signal that captures uncertainty of past events, the relief prediction error reflects frustration of having received the current outcome. Since pain ratings were always collected after such frustrated attempts in obtaining relief, it is therefore more difficult to conclude whether these ratings reflect pain or reactions to outcome. As for neuroimaging results, although not surviving multiple comparison, we identified state-based prediction errors in the left amygdala, consistent with a previous tonic pain relief study (Seymour et al., 2005). Previous evidence have also shown the amygdala representing multiple computationally distinct learning signals: aversive, reward, relief prediction errors, and associability (Belova et al., 2007; Johansen et al., 2010; Li et al., 2011; Paton et al., 2006;

Roesch et al., 2010, 2012). Based on these current findings, it is difficult to draw definitive conclusions regarding the effects of passive relief learning in Pavlovian sessions.

These results also led us to identify the major limitations of the current experimental design – the limited number and the post-outcome timing of pain ratings. This rating method was chosen as a compromise to reduce subjects' tonic pain exposure and increase the number of overall trials to boost learning. However, while the correlation analysis pointed towards an associability driven analgesic effect, the number of ratings may not be adequately powered for a comprehensive model comparison with other forms of uncertainty. Furthermore, collecting ratings after outcome occurrence made it difficult to distinguish pain and frustration. These limitations can be addressed by increasing the number of rated trials in a task design focusing on instrumental relief learning only, as demonstrated in the next experiment.

4.5 Tables

Table 4.1 Details of subjective ratings.

Experiment	Rating type	Rating timing	Avg # of ratings/subject
Active vs. Passive relief learning (Experiment 2)	Instrumental pain Instrumental relief Pavlovian pain Pavlovian relief	After 3s cue+choice window AND outcome (rating type depend on outcome)	8.2 7.7 8.1 7.7

Table 4.2 All learning models fitted (bold: winning model; AL - action-learning; SL - state-learning, F - variational Bayesian approximation to the model's marginal likelihood, used for model comparison, assoc - associability)

Instrumental sessions			
Choice	F (n=19, sum [SEM])	SCR	F (n=15, sum [SEM])
TD	-1330.920 [3.604]	RW - value	-1079.153 [8.024]
Hybrid (AL)	-1345.667 [3.664]	Hybrid (SL) - value	-1077.911 [8.059]
WSLS	-1486.723 [3.973]	**Hybrid (SL) - assoc**	-1077.699 [8.003]
Pavlovian sessions			
Choice (not available)		SCR	F (n=15, sum [SEM])
N/A		RW - value	-1101.079 [7.132]
		Hybrid (SL) - value	-1096.250 [7.195]
		Hybrid (SL) - assoc	-1095.135 [7.106]

Table 4.3 Learning model fitting results (std: standard deviation; μ_θ: Bayesian prior mean of evolution function parameter; σ_θ: prior variance of evolution function parameter; μ_ϕ: prior mean of observation function parameter, σ_ϕ: prior variance of observation function parameter).

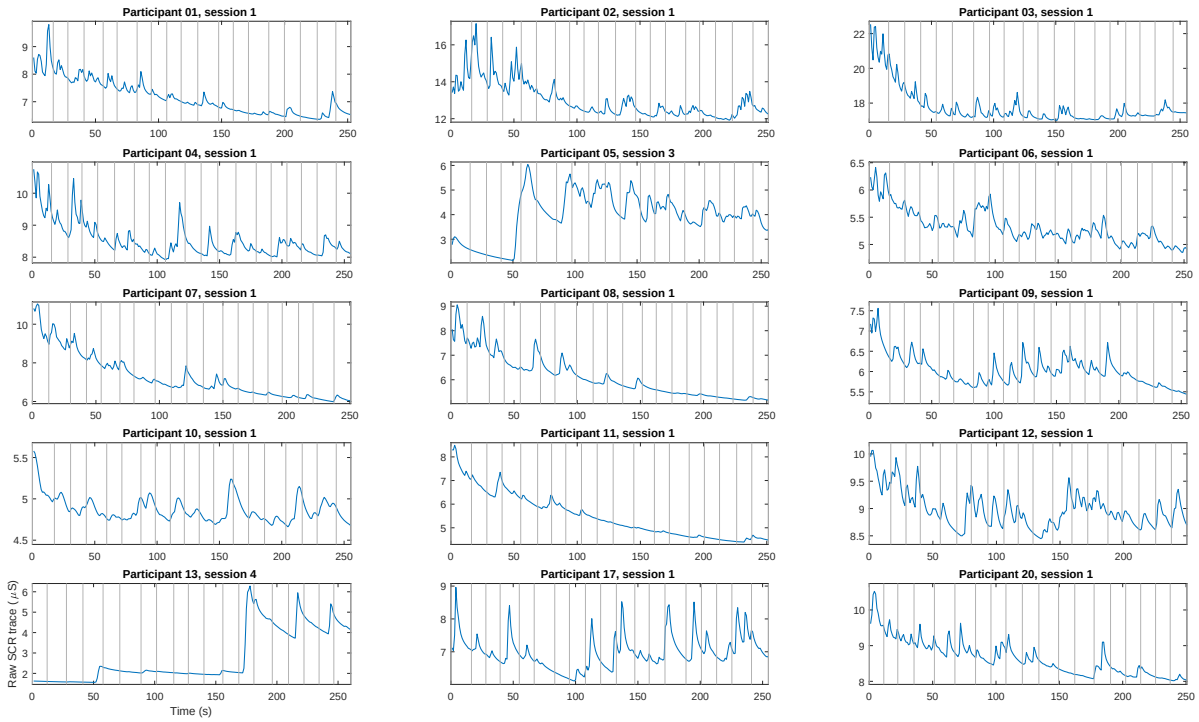
Model (Options)	Data fitted (sessions)	Parameters	Mean	Std	Initial states
TD (*)	choice (instrumental)	learning rate, α	0.401	0.087	Q_0=0
WSLS (*)	choice (instrumental)	pseudo Q (cue 1), p1	0.382	0.073	No
		pseudo Q (cue 2), p2	0.458	0.075	hidden states
Hybrid Action learning (*)	choice (instrumental)	free parameter κ	0.527	0.104	Q_0=0
		free parameter η	0.413	0.125	α_0=1
RW - V (†)	SCR (instrumental)	learning rate, α	0.492	0.013	V_0=0
RW - V (†)	SCR (Pavlovian)	learning rate, α	0.492	0.014	V_0=0
Hybrid - Assoc (†)	SCR (instrumental)	free parameter κ	0.497	0.004	V_0=0
		free parameter η	0.495	0.004	α_0=1
Hybrid - Assoc (†)	SCR (Pavlovian)	free parameter κ	0.498	0.003	V_0=0
		free parameter η	0.496	0.008	α_0=1
Hybrid - V (†)	SCR (instrumental)	free parameter κ	0.492	0.012	V_0=0
		free parameter η	0.499	0.003	α_0=1
Hybrid - V (†)	SCR (Pavlovian)	free parameter κ	0.494	0.005	V_0=0
		free parameter η	0.5	0.003	α_0=1

*Fitting options: $\mu_\theta = 0$, $\sigma_\theta = 1$, $\mu_\phi = 0$, $\sigma_\phi = 1$
†$\mu_\theta = 0$, $\sigma_\theta = 0.05$, $\mu_\phi = 0$, $\sigma_\phi = 1$

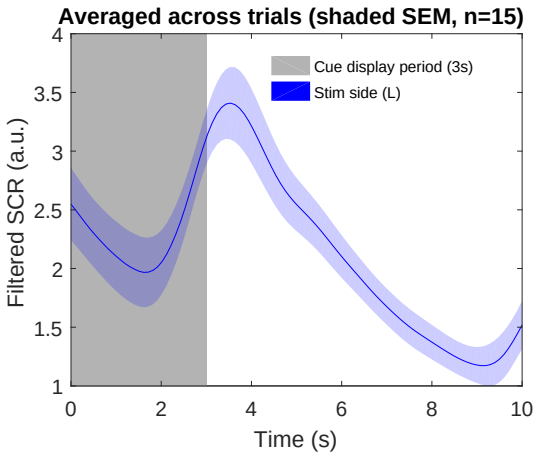
Table 4.4 Multiple correction (cluster-forming threshold of p <0.001 uncorrected, regions from Harvard-Oxford atlas. *FWE cluster-level corrected (showing p<0.05 only)

p*	k	t	Z	MNI coordinates (mm)			Region mask
				x	y	z	
TD model PE, instrumental sessions							
0.007	4	4.27	3.5	-21	-5	-14	Amygdala L
0.011	3	4.98	3.9	28	-1	-14	Amygdala R
0	28	5.31	4.07	-21	3	-7	Putamen L
		4.7	3.75	-28	-5	1	
0.003	14	5.73	4.27	20	7	-7	Putamen R
0.034	2	3.75	3.18	28	-1	8	
0.007	4	4.63	3.71	-17	3	-3	Pallidum L
0.003	9	5.2	4.01	17	7	-3	Pallidum R
Hybrid model PE, instrumental sessions							
0.005	5	4.3	3.52	-21	-5	-14	Amygdala L
0.014	2	4.53	3.65	28	-1	-14	Amygdala R
0.004	12	5.02	3.92	-21	3	-7	Putamen L
0.012	6	4.55	3.66	-28	3	8	
0.046	1	3.82	3.23	-28	11	-3	
0.001	23	5.03	3.92	20	7	-7	Putamen R
		4.92	3.87	20	7	1	
		4.39	3.57	24	-1	5	
0.006	5	4.04	3.36	-17	3	-3	Pallidum L
0.005	6	4.82	3.81	17	7	1	Pallidum R
Hybrid model PE, Pavlovian sessions							
None							
Hybrid model associability, instrumental sessions							
0.027	5	4.34	3.55	-2	37	5	Cingulate Anterior

4.6 Supplementary figures



(a)



(b)

Fig. 4.4 Skin conductance raw data. (a) Raw skin conductance traces, where vertical lines are beginning of each trial when cue display starts (n=15, excluded participants not shown, showing first non-excluded session from all participant). (b) Filtered skin conductance traces (band-pass at 0.0159-2Hz, 1st order Butterworth), averaged across all trials within participant (n=15, excluded participants not shown, shaded region represent SEM across all participants).

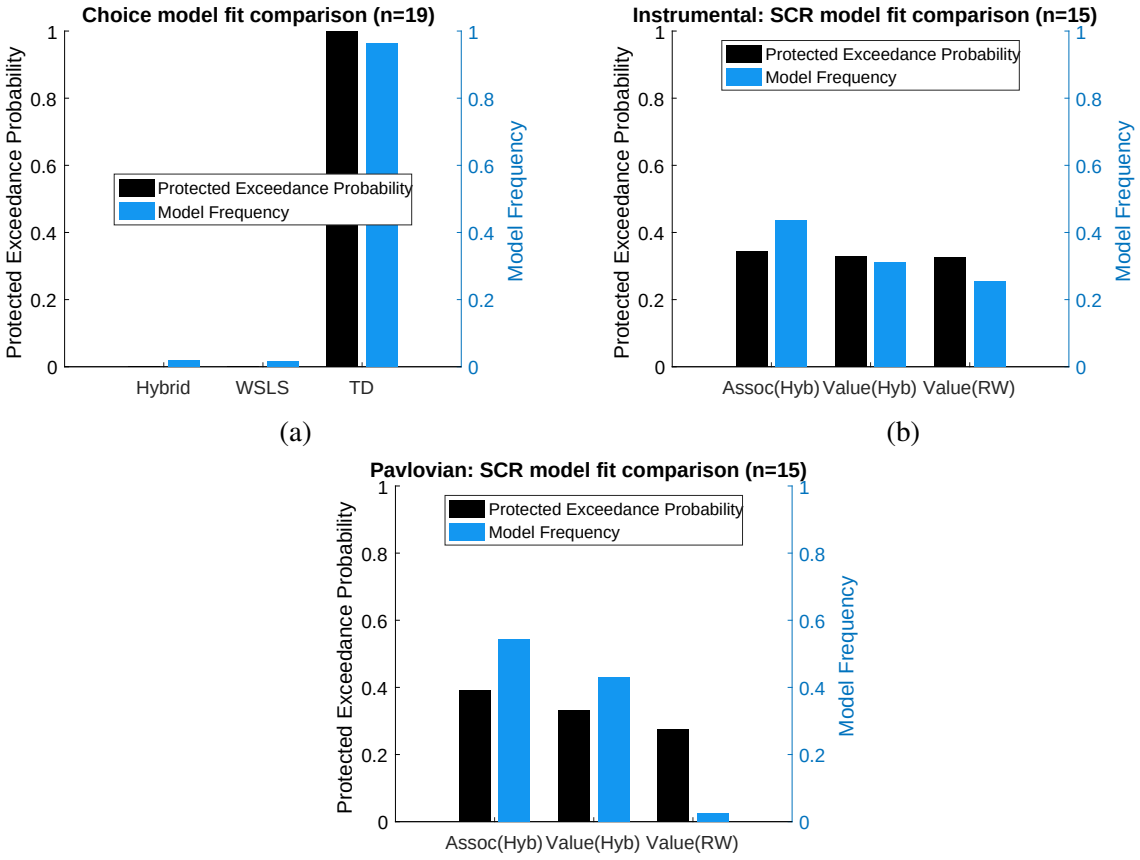


Fig. 4.5 Model protected exceedance probability. (a) Choice fitting remains similar to the original exceedance probability. (b-c) SCR fitting comparison becomes less clear regarding best fitting model. However, the best fitting models from comparison remained unchanged compared to the original comparison using exceedance probabilities in Experiment 3 (Chapter 5), where the number of trials was increased as fitting wasn't conducted separately for Instrumental / Pavlovian sessions, providing further validation of the results (Fig 5.5).

Chapter 5

Experiment 3: Active relief learning in a dynamic environment

Experiment 2 (Chapter 4) provided good evidence of a relief learning system capable of modulating tonic pain according to uncertainty during learning by tracking the associability signal. However, it remains unclear whether other forms of uncertainty signals, especially those arising from a more complicated learning environment, may be responsible for pain modulation. Using a modified tonic pain escape learning task with dynamic relief probabilities, we showed that active relief learning can still be modelled with a reinforcement learning process. With an increased number of pain ratings, we validated the results of associability-driven analgesia modulating tonic pain levels in a dynamic learning environment, as well as its correlation with BOLD activities in the pregenual anterior cingulate cortex. Together with the previous experiment, these results define a self-organising learning circuit that allows reduction of ongoing pain when learning about potential relief.

Adapted from Zhang et al. (2018a)

5.1 Introduction

In Experiment 2 (Chapter 4), the associability signal was found to reflect pain modulation during active relief learning, where tonic pain was suppressed according to the information available for learning. While associability as an internal uncertainty signal during learning is a potential candidate to bridge the gap to previous results on attention-driven endogenous analgesia (Bantick et al., 2002; Salomons et al., 2004; Valet et al., 2004; Wiech et al., 2006), this result also raised two important questions regarding the previous experimental design.

The first question was: can the associability signal be distinguished from other uncertainty signals that may arise in learning? Importantly, the use of fixed probabilities in the previous experiment means that associability tends to decline during sessions – as learning progresses, values of relief increase after a succession of relief trials, making subsequent relief outcomes less surprising, which leads to low associability/uncertainty. This raises the possibility that more complex models of uncertainty and attention might better explain the data. Model-based learning models, such as hidden Markov models (HMMs) or hierarchical Bayesian models, assume learners construct internal models of a complex, dynamic world (Behrens et al., 2007; Mathys et al., 2011; Prévost et al., 2011). These models incorporate changing beliefs of the environment during learning, which gives rise to different forms of uncertainty between the levels of environment, stimulus, actions, and outcomes (Daw et al., 2011; Mathys et al., 2011). In addition, predictors of model-based learning models have been found to correlate with activities in the prefrontal cortex, raising the possibility that previously observed activations in the pgACC might be tracking multiple learning signals (Daw et al., 2005; Gläscher et al., 2010). In order to tease apart the different uncertainty signals, a dynamic world with non-stationary relief probabilities is used in this study.

A second design question was: does the modulation of pain ratings occur throughout the trial? In the previous relief learning task, pain ratings were taken *after* choice action and trial outcome occurrence, which means only when relief was frustrated. This raises the possibility that the subjective rating reflects an outcome-driven response, as opposed to a learning-driven process modulating the ongoing pain. In addition, uncertainty was previously modelled before outcome onset, when it was assumed to combine with prediction errors to influence value update (Li et al., 2011). Therefore, we believe taking ratings *before* outcome is more likely to capture the modulatory effects on tonic pain with minimal impact from outcomes.

With these issues in mind, we designed a novel task to test if the active relief learning model could be generalised to a different paradigm with greater demands on flexible learning. We modified the relief learning task to include dynamic cue-relief contingencies, where participants learn to choose one cue out of three to obtain brief relief from tonic thermal pain, and need to adjust their choice throughout the task. We aimed to test for changes in learning strategies, as

well as other forms of uncertainty signals against associability for modulatory effects on tonic pain.

5.2 Methods

Subjects

23 healthy subjects participated in the neuroimaging experiment (five female, age 23.9±3.1 years). All subjects gave informed consent prior to participation, had normal or corrected to normal vision, and were free of pain conditions or pain medications. The experiment was approved by the Ethics and Safety committee of the Advanced Telecommunications Research Institute, Japan.

Experimental design

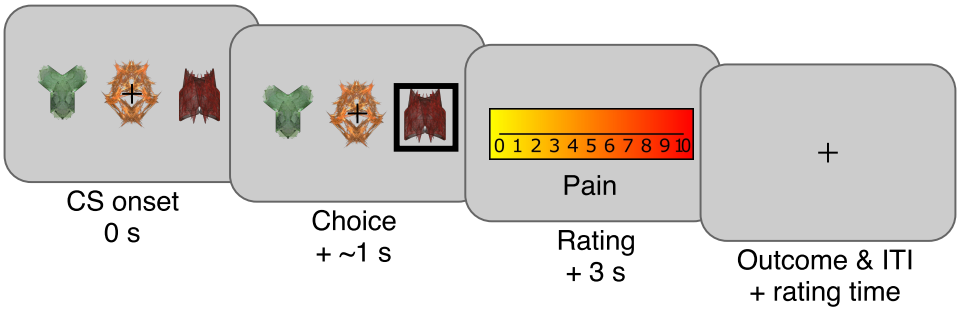
Subjects participated a purely instrumental relief conditioning task similar to that in Experiment 2 (Chapter 4). In this task, three visual cues were presented on screen simultaneously for 3s, during which the subject was asked to choose one (Fig 5.1a). Each one of these cues had varying relief probability, generated by a random walk program (probabilities changing at step size of 0.1, bound between 0.2-0.8, with random start). Relief outcomes were temporary reductions of 13°C from the tonic level, with the duration reduced to 3s, which was enough to produce a similar relief sensation with lower trial time. Subjects repeated the same task for eight sessions (24 trials each), with the same visual cues throughout. However, several subjects did not complete all sessions because of excess time in SCR experimental set-up which reduced the time available for the task; hence, the overall average was 7.08±1.44 sessions per subject.

In each session, subjective pain ratings were collected in 10 random trials out of the 24. The ratings were collected after the 3s choice period and before the outcome presentation, using the same 0-10 rating scale as Experiment 2 (red scale only, see Table 5.1).

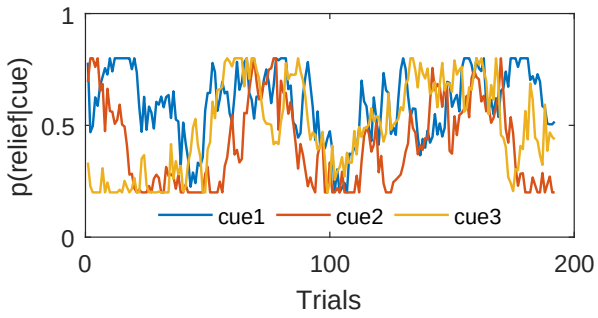
Stimulation

Painful tonic thermal stimuli were delivered with a contact heat-evoked potential stimulator (CHEPS, Medoc Pathway, Israel) to the subject's skin surface above the wrist on the left inner forearms.

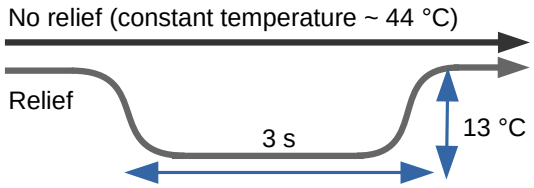
The temperature of painful tonic stimuli was thresholded slightly differently compared to the previous tonic pain relief experiment. 10 temperatures were presented in each series, both were randomly generated with 44.4±0.7°C. After the 8s stimulation, subjects were asked to



(a)



(b)



(c)

Fig. 5.1 Experiment paradigm. (a) Example trial, where subjects performed an instrumental paradigm only involving unstable relief probabilities. Three cues were presented alongside each other with subjects required to choose one of the three using a button press. The position of each cue varied from trial-to-trial, and the same three cues were presented throughout. Tonic pain rating being taken before the outcome was experienced, not after as in Experiment 2. (b) Example traces of dynamic relief probabilities for the three displayed cues throughout all trials in eight sessions, which required a constant trade-off of exploration and exploitation throughout the task. Dynamic relief probabilities also provide varying uncertainty throughout learning. (c) Relief and no relief outcomes, with constant temperature at around 44°C were used to elicit tonic pain similar to Experiment 2, a brief drop in temperature of 13°C was used as relief outcome for 3s, which was enough to produce relief sensation but with shorter trial time. Temperature did not change for the duration in no relief outcomes.

rate their pain on a linear VAS scale with 0-10 numerical markings, which were fitted with a sigmoid function. The temperature was chosen from the painful range of: 44, 44.2, 44.5, 44.8, 45°C, whichever closest and below the model fitted value of VAS=8.

Final temperature used was 44.5±0.4°C, similar to that of Experiment 2 (44.3±0.2°C). Relief temperature was at 13°C below threshold temperature for all subjects. Thermal stimulation settings were identical to Experiment 2, except relief duration was shorted to 3s.

Physiological measures

SCRs were recorded from both hands, in the same location as Experiment 2 on volar surfaces of distal phalanges of the second and fourth fingers on the left (pain side with thermode), and on the hypothenar eminences of the palm on the right (button press hand without thermode), with electrodes approximately 2cm apart. The signals were collected using MRI-compatible BrainAmp ExG MR System (Brain Products, Munich, Germany) and BrainVision software at 500Hz with no filter.

The offline processing and exclusion criteria were identical to Experiment 2. Sessions were excluded if more than 20% trials were labelled as not having enough viable event related SCRs (cue-evoked SCR amplitude below the threshold of 0.02). 19 subjects and 79 sessions remained for the left (thermal stimulation side), 20 subjects and 96 sessions remained for the right (no stimulation side). For model fitting, right side SCR reject criteria were used, since both channels' data were included as two data sources. Transformed SCRs on both sides were highly correlate (Fig 5.2b).

Computational learning models

Model-free reinforcement learning models were fitted to choice and SCR data, including the *TD model*, *hybrid model*, and *Rescorla-Wagner (RW) model* (see section 4.2 for model specifications). The softmax function (Equation 4.2) was used for mapping choices to model predictors, and a simple translation function was used for SCR mapping (Equation 5.8).

Hidden Markov Model (HMM)

For Experiment 3, where relief probability is unstable, model-based learning models were fitted to behavioural data. A Hidden Markov Model with dynamic expectation of change (Prevost et al., 2013; Schlagenhauf et al., 2014) was adapted to incorporate a hidden state variable S_t that represents the subject's estimation of an action-outcome pair (e.g. $S_t = (\text{cue}, \text{relief})$, three

cues × relief/no relief = 6 combinations). The state transition probabilities are calculated as:

$$P(S_t|S_{t-1}) = \begin{pmatrix} 1-\beta & \beta \\ \beta & 1-\beta \end{pmatrix} \tag{5.1}$$

where β is a free parameter ($0 \leq \beta \leq 1$). For each cue, the symmetry of the transition matrix encodes the reciprocal relationship between relief/no relief belief. Given the hidden state variable, the probability of actually observing this outcome is updated as:

$$P(O_t|S_t) = 0.5 \times \begin{pmatrix} 1+c & 1-c \\ 1-d & 1+d \end{pmatrix} \tag{5.2}$$

where the rows of the matrix represent relief/no relief outcomes, the columns represent the relief/no relief belief in S_t. c and d are free parameters ($0 \leq c \leq 1$, $0 \leq d \leq 1$) to incorporate potential discrimination between the two outcome types. The prior probability of S_t is calculated from the state transition probabilities and the posterior probability of S_{t-1} (Equation 5.3). The posterior probability of S_t is calculated from the prior $P(S_t)$ (from Equation 5.3) and the observed outcome O_t (Equation 5.4):

$$P(S_t) = \sum_{S_{t-1}} P(S_t|S_{t-1})P(S_{t-1}) \tag{5.3}$$

$$P(S_t) = \frac{P(O_t|S_t)P(S_t)}{\sum_{S_t} P(O_t|S_t)P(S_t)} \tag{5.4}$$

where Equation 5.3 is updated before observed outcome O_t, Equation 5.4 is updated after O_t.

S_t can be used to approximate state values by calculating the relative relief belief through a sigmoid function, with a free parameter m, and the preferred action to be inferred using the softmax function.

$$P(\text{r=1}|\text{cue}) = \frac{1}{1+\exp(-x)} \tag{5.5}$$

where $x = S_t(\text{r=1}) - S_t(\text{r=0}) + m$.

To represent uncertainty under i possible posterior relief probabilities, entropy H is calculated for the chosen cue as:

$$H(S_t) = -\sum_i P(S_t)\log P(S_t) \tag{5.6}$$

Hierarchical Bayesian model

The Hierarchical Bayesian model introduced by Mathys et al. (2011) incorporates different forms of uncertainty during learning on each level: irreducible uncertainty (resulting from

probabilistic relationship between prediction and outcome), estimation uncertainty (from imperfect knowledge of stimulus-outcome relationship), and volatility uncertainty (from potential environmental instability). This model has been shown to fit human acute stress response (de Berker et al., 2016). The model was adopted to our study with the basic structure unchanged. The second level estimated probabilities were used to approximate state values of different cues, and the preferred action calculated using the softmax function.

Modelling pain ratings

Our prior hypothesis suggests uncertainty is a likely modulator of tonic pain perception, hence model generated uncertainty signals (associability, entropy, and surprise) were used as the main pain rating predictors, given the increased number of ratings collected in Experiment 3. A generalised linear model includes the uncertainty predictor, and additional terms to control for potential temporal habituation/sensitization and between-session variation:

$$\text{Rating} = \beta_1 \cdot \text{Relief} + \beta_2 \cdot \log(\text{Trial}) + \beta_3 \cdot \text{Predictor} \tag{5.7}$$

where the 'Relief' term is the number of trials since the previous relief outcome, $log(\text{Trial})$ is the log of trial number within session (1-24), 'Predictor' is the model generated uncertainty value using group-averaged model parameters fitted with choice/SCR data. All trials were used for predictor calculation, but only rated trials were included in this regression. The $\log(\text{Trial})$ term is used because it can capture the nonlinear effects of habituation or sensitisation from prolonged thermal pain (Jepma et al., 2014). This GLM approach for pain rating modelling allows the use of Bayesian model comparison with resulting model fitting statistics, in line with other behavioural data modelling in Experiments 2 and 3. Repeated measure ANOVA results of the winning model predictor verified the results were not confounded by correlated predictors in the regression model.

Model fitting and comparison

Model fitting of choice and SCR data was performed with the Variational Bayesian Analysis (VBA) toolbox (https://mbb-team.github.io/VBA-toolbox/), as was described in section 4.2. Model parameters for individuals were optimised in a Bayesian equivalent of maximum likelihood (Table 5.2). The mean of best fit parameters from all subjects were summarised in Table 5.3. These parameters were also used to generate regressors for fMRI analysis.

For SCR fitting, as data were collected on both hands, left and right SCRs were fitted simultaneously as two data sources in the observation functions:

$$g(x_1) = \text{predictor} + b_1$$
$$g(x_2) = \text{predictor} + b_2 \tag{5.8}$$

where x represents trial-by-trial SCRs from two sources, with b_1 and b_2 as two free parameters to fit each side with the same predictor. Note that including two free parameters to bilateral SCR fitting might change the absolute fit statistics from those of Experiment 2, but it did not influence model comparisons since these were not conducted across experiments.

The TD, RW and hybrid models were assumed to have initial values of 0, and initial associability of 1. HMM and Bayesian models were assumed to have initial hidden states of relief belief as 0. All free parameters of the evolution function were assumed to have prior variance of 1, with the exception of SCR fitting where it was assumed to be 0.05 to reduce flexibility.

We also calculated the protected exceedance probabilities based on Rigoux et al. (2014) (Fig 5.5). In Experiment 3, the number of trials was increased as fitting was conducted with data from all sessions, not separately for instrumental/Pavlovian sessions as in Experiment 2. Best fitting models from protected comparison remained unchanged from the original comparison using exceedance probabilities. These results provided validation for Experiment 2 in the way similar to the neuroimaging analysis.

fMRI acquisition

A 3T Siemens Prisma scanner was used for neuroimaging data acquisition, with the Siemens standard 64 channel phased array head coil.

Scanning parameters were deliberately kept identical to the previous experiment. Functional images were collected with a single echo EPI sequence (repetition time TR=2500ms, echo time TE=30ms, flip angle=80, field of view=240mm), in 37 contiguous oblique-axial slices (voxel size 3.75 $\times$ 3.75 $\times$ 3.75 mm) parallel to the AC-PC line. Whole-brain high resolution T1-weighted structural images (dimension 208 $\times$ 256 $\times$ 256, voxel size 1 $\times$ 1 $\times$ 1 mm) were obtained with the standard MPRAGE sequence.

fMRI preprocessing

Functional images were slice time corrected using SPM12 (http://www.fil.ion.ucl.ac.uk/spm/software/spm12/). Image preprocessing was conducted using the fmriprep software (build

date 09/03/2017, freesurfer option turned off, https://github.com/poldracklab/fmriprep). The normalised images were smoothed using a Gaussian kernel of 8mm using SPM12. The 24-column matrix output by fmriprep were used in subsequent analysis to account for confounds.

fMRI GLM model

First level analysis for individual participant was conducted with SPM12. General linear models included best-fit learning model generated signals as parametric modulators, and regressors of no interest to account for habituation and motion effects. Second-level one-sample t-test of first-level GLM estimates were carried out using procedures identical to that of Experiment 2 (Section 4.2), and results with $p<0.05$ were summarised in Table 5.4 (FWE cluster-level corrected for multiple comparison, based on the Harvard-Oxford anatomical atlases).

fMRI model comparison and axiom analysis

Similar to Experiment 2, Bayesian model selection (BMS) was used to determine which brain regions were involved in state-based or action-based learning, within the conjunction activation mask of both prediction error signals.

In addition, axiom analysis was used to determine whether ROI activations were driven by outcomes or prediction errors (section 4.2). As we are primarily interested in the overall BOLD activity pattern, we did not include full statistics in this analysis. With increased number of trials, six regressors (2 outcomes $\times$ 3 value bins) were generated at outcome time (duration=3s) when PE was generated. Again, we focused on *a prior* selected ROIs related to relief learning.

5.3 Results

23 new subjects participated in a modified version of the instrumental escape learning task described in Experiment 2, with a number of important differences. First, subjects performed only instrumental sessions (8 sessions with 24 trials in each) given the absence of a global effect of instrumental versus Pavlovian pain in the previous relief experiment. Second, subjects were required to choose one out of *three* simultaneously presented visual cues to obtain relief, in which the position of each cue varied randomly from trial to trial. This was done to experimentally and theoretically better distinguish state-based and action-specific associability (Fig 5.1a). Third, the action-outcome contingencies were *non-stationary* (dynamic), such that the relief probability from selecting each cue varied slowly throughout the experiment duration, controlled by a random walk algorithm which varied between 20-80% (Fig 5.1b). This ensured that associability varied constantly through the task, encouraging continued relief exploration,

and allowed us to better resolve more complex models of uncertainty (see below). It also reduced the potential confounding correlation of associability and general habituation of SCRs. Fourth, we increased the frequency of tonic pain ratings (10 per session, 80 per subject in total, Table 5.1) to enhance power for identifying modulatory effects on pain. Fifth, the rating was taken after the action but before outcome, to provide an improved assessment of ongoing tonic pain modulation without interference by the outcome. Finally, we also collected SCRs bilaterally, to enhance the data quality given the importance of the SCR in inferences about associability.

Behavioural results

Choice In addition to the simple TD and hybrid action-learning TD models compared in Experiment 2, the modification in paradigm allowed us to test more sophisticated model-based learning models, including a hidden Markov model (HMM) (Prevost et al., 2013), and a hierarchical Bayesian model (Mathys et al., 2011). Both models incorporate a belief of environmental stability into learning, i.e. whether a cue previously predicting relief reliably has stopped being reliable during the course of the experiment. This is achieved by tracking the probability of state transition in the HMM, or environmental volatility in the hierarchical Bayesian model. Despite the greater demands of the non-stationary task compared to Experiment 2, the basic TD action learning model still best predicted choices following model comparison (model frequency=0.624, exceedance probability=0.989), followed by the HMM (model frequency=0.192, exceedance probability=0.006) and the hybrid action-learning model (model frequency=0.174, exceedance probability=0.004) (Fig 5.2a, see Methods for full details).

SCR SCRs were recorded from the side with thermal stimulation (left hand) and the side without stimulation (right hand). The left side had lower mean SCRs (Fig 5.2b, L/R paired t-test $T(19)=-2.67$, $p=0.015$, $n=20$, exclusion criteria followed from Experiment 2), however, trial-by-trial SCRs were highly correlated between both sides within individual subjects (mean Pearson correlation $\bar{\rho}=0.733$, 18 out of 20 participants with $p<0.001$). This suggests that although the overall SCR amplitude might be suppressed by the tonic heat stimulus, this did not affect event-related responses.

Using the same model-fitting procedure as in Experiment 2 (with the addition that the model now predicted SCR on both hands for each trial), we found that the associability from the state-outcome hybrid model again provided the best fit of trial-by-trial SCRs (Fig 5.2c, model frequency=0.667, exceedance probability=0.954). Indeed, the associability-SCR fit has a much higher model exceedance probability compared with that in Experiment 2 (calculated

by comparing a set of models within each experiment, not comparing fit across experiments), presumably from including the less attenuated SCRs from the non-stimulated right side.

Ratings Experiment 2 suggested that the associability acted was correlated with modulation of tonic pain ratings. Given the dynamic nature of the current design, we investigated whether uncertainty measures related to other aspects of learning might offer a better account. To do this, we fitted multiple regression models to trial-by-trial ratings for each participant as follows:

$$\text{Rating} = \beta_1 \cdot \text{Relief} + \beta_2 \cdot \log(\text{Trial}) + \beta_3 \cdot \text{Predictor} \tag{5.9}$$

where the 'Relief' term is the number of trials since the previous relief outcome, log(Trial) is the log of trial number within session (1-24), 'Predictor' is the model generated uncertainty value. The 'Relief' and log(Trial) terms were included to account for potential temporal and sessional effects of the tonic pain stimulus.

We built a regression model with different uncertainty signals as predictors for comparison: the state-based associability from hybrid model (as in Experiment 2), the entropy of state-action posterior probabilities (approximate of uncertainty over values) in an HMM, the absolute value of prediction error from previous trial in TD model (as a model of surprise), and a null model that did not include 'Predictor' term (Fig 5.2d). In this analysis, the state-learning hybrid associability again best fit the pain ratings (model frequency=0.698, exceedance probability=0.980; n=22, 1550 ratings, one participant was excluded for having >90% identical ratings). Regression coefficients with hybrid model associability as the uncertainty predictor were significant across subjects (Fig 5.2e, one-sample t-test for three sets of coefficients: 'Relief' term: T(21)=-4.004, p<0.001 (i.e. habituation, reduced pain over time after relief), log(trial) term: T(21)=1.017, p=0.321, associability term: T(21)=-2.643, p=0.015). Repeated measure Type-III sums-of-squares ANOVA showed all terms were statistically significant ('Relief' term: F(1,1545)=37.02, p<0.001, log(trial) term: F(1,1545)=14.13, p<0.001, associability term: F(1,1545)=4.35, p=0.037).

Neuroimaging results

Prediction errors We found that the TD model's action prediction errors were robustly correlated with BOLD responses in similar regions identified in Experiment 2, including left dorsal putamen, bilateral amygdala, and left DLPFC (Fig 5.3a, Table 5.4). Of these, BMS showed the TD model had higher posterior and exceedance probabilities in the dorsal putamen, as well as amygdala and DLPFC (Fig 5.3b warm colour clusters). The state-learning hybrid

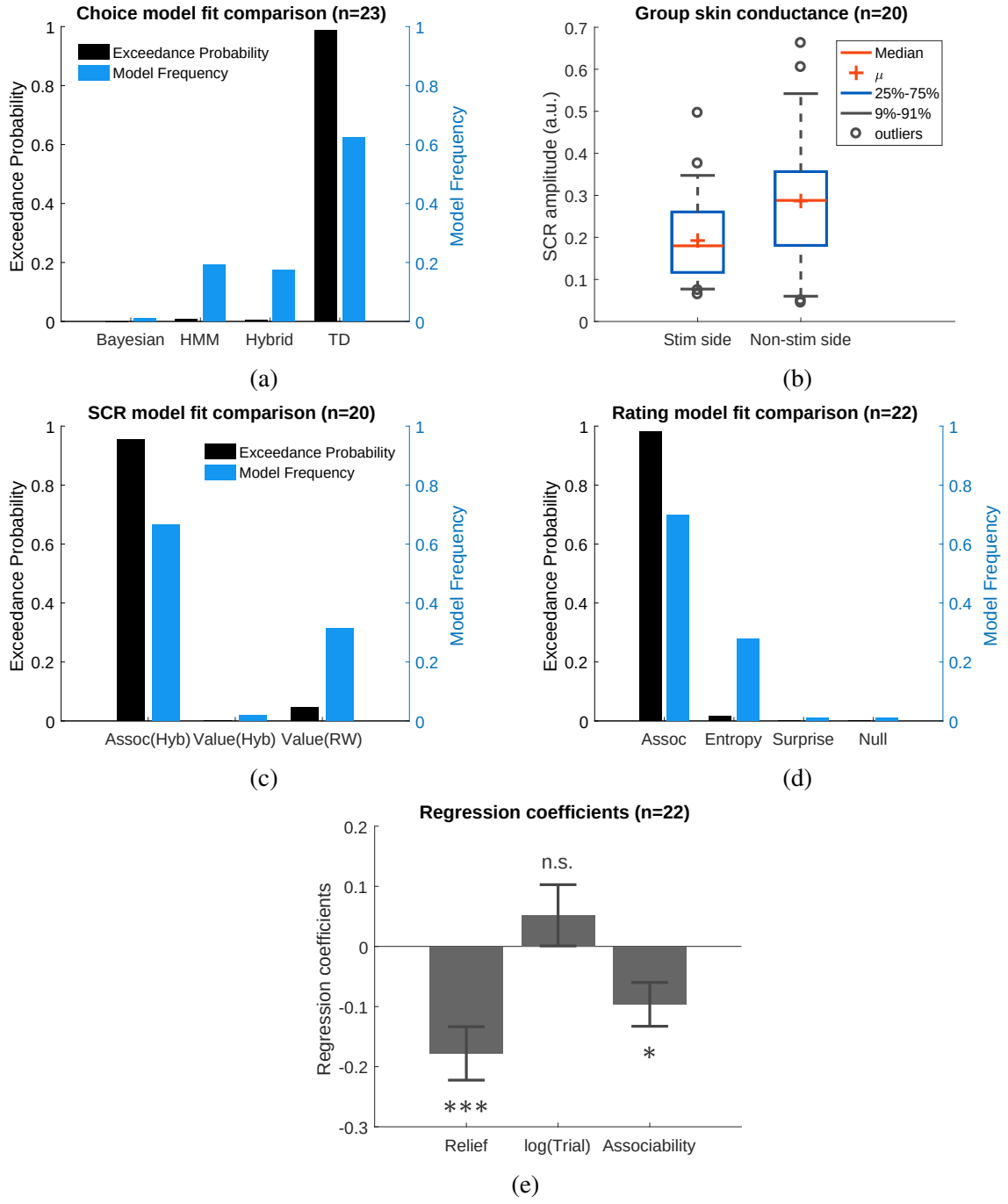


Fig. 5.2 Behavioural results. (a) Model comparison showed that TD model fitted choices best (Bayesian: hierarchical Bayesian model, HMM: hidden Markov model, Hybrid: action-learning model with associability as changing learning rate). (b) SCRs measured on the side with thermal stimulation ('Stim side', left hand) were lower than those on without stimulation ('Non-stim side', right hand), but both were highly correlated. (c) Associability from state-learning hybrid model fit SCRs best, similarly to Experiment 2. (d) Trial-by-trial associability from hybrid model fitted pain ratings best compared with other uncertain measures (entropy: HMM entropy, surprise: TD model prediction error magnitude from previous trial, null model: regression with no predictors). (e) Regression coefficients with associability as uncertainty predictor were significantly negative across subjects.

model explained prediction error responses in several areas, but outside our original *a priori* regions of interest (see Fig 5.3b cool colour clusters).

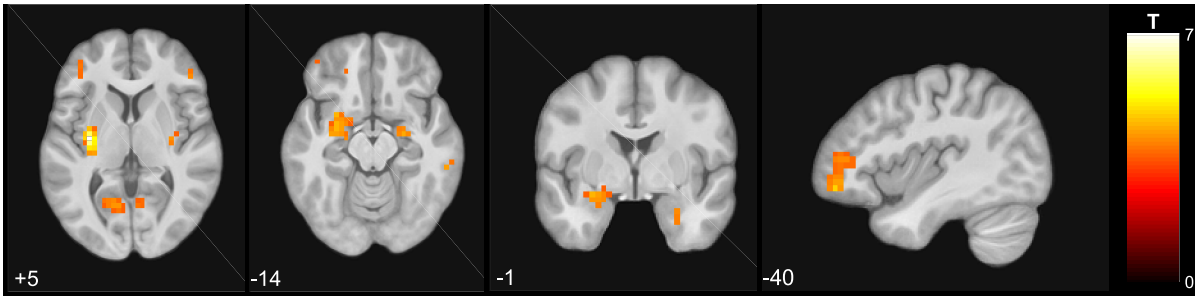
As previously, we further illustrated the pattern of outcome responses as a function of preceding cue value and relief/no-relief in an 'axiomatic' analysis. We split the trial values into 3 bins, allowing a better inspection of responses permitted by our larger number of trials. This revealed a clear prediction error-like pattern in the dorsal putamen, but it was somewhat less clear cut in the amygdala and DLPFC (Fig 5.3c). Therefore, across all analysis methods and the two experiments, the left dorsal putamen robustly exhibited a response profile consistent with an escape-based relief prediction error.

Associability Following the same analysis as in Experiment 2, we found again that pgACC BOLD responses correlated with trial-by-trial associability from the state-learning hybrid model (Fig 5.3d, Fig 5.3e, Table 5.4). The peak from this analysis was almost identical to that in Experiment 2 (Overlaid clusters can be found in Fig 5.6a). In addition, we used trial-by-trial pain ratings as a parametric modulator, but did not find significant pgACC responses, which suggested that it was unlikely to be solely driven by pain perception itself. Taken together, this indicates that the pgACC associability response is robust across experimental designs.

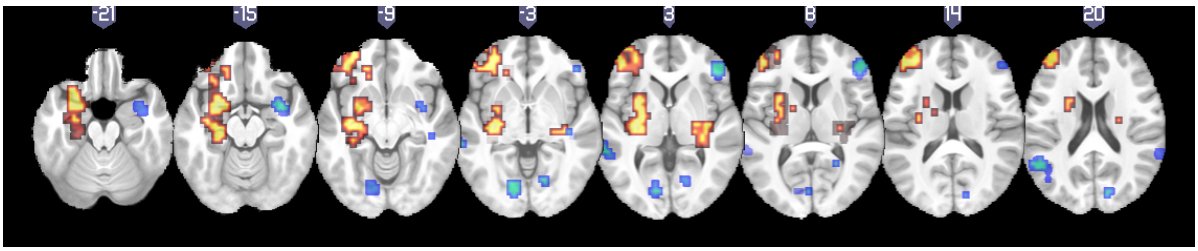
5.4 Discussion

In summary, this experiment reproduced the main results of active relief learning from Experiment 2 within a non-stationary relief environment. Firstly, dorsal putamen correlated with an action-relief prediction error from the RL model. And secondly, pgACC correlated with a state-based associability signal, that in turn was associated with reduced tonic pain. In particular, this modulation of pain was present after the cue was presented (and not just at the outcome as in Experiment 2), and was better explained by the associability signal when compared against alternative uncertainty measures.

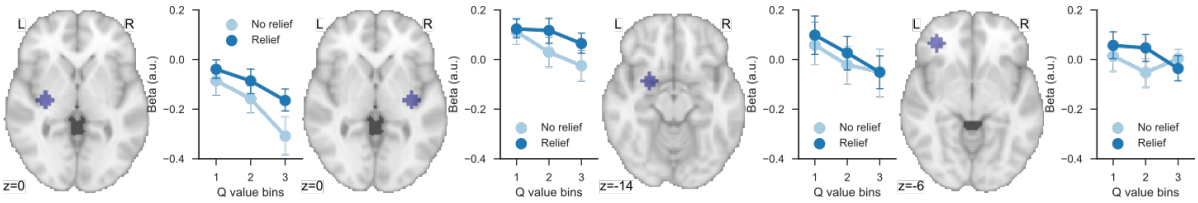
Across both relief learning experiments, the results provide convergent support for two key findings. First, we show that relief seeking from the state of tonic pain is supported by a reinforcement learning process, in which optimal escape actions are acquired using prediction error signals, which are observed as BOLD signals in the dorsal putamen. Second, we show that during learning, the level of ongoing pain is inversely correlated with the learned associability associated with state-based relief predictions. This signal thus corresponds to lower subjective pain when there is a greater capacity to learn new information, and is associated with BOLD responses in the pregenual anterior cingulate cortex. Together, these results identify a learning circuit that governs tonic pain escape learning whilst also suppressing pain according to the



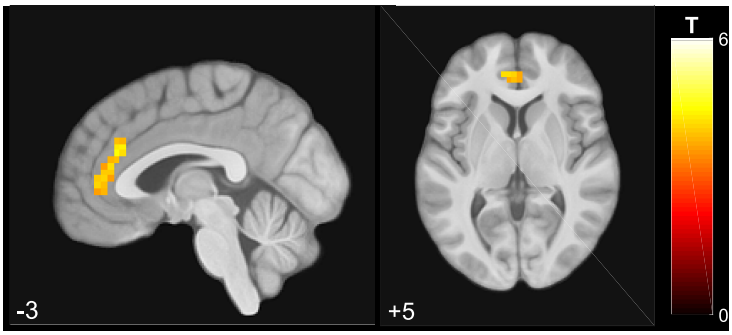
(a)

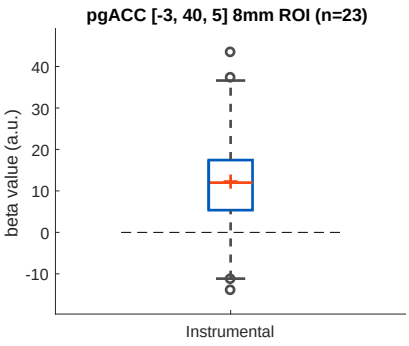


(b)



(c)





(d) (e)

Fig. 5.3 Neuroimaging results, shown at p<0.001 uncorrected: (a) TD model prediction errors (PE), at outcome onset time (duration=3s). (b) Model PE posterior probability maps (PPMs) from group-level Bayesian model selection, warm colour: TD model PE, cool colour: hybrid model PE (both shown at exceedance probability P>0.80). (c) Axiom analysis, separating trials according to outcomes and predicted relief values (bin 1-3 from low to high), BOLD activity pattern from striatum (putamen) satisfied those of relief PE. (d) Associability uncertainty generated by hybrid model correlating with pgACC activities, at choice time (duration=0). (e) pgACC activation beta values across all subjects, ROI was 8mm sphere at [-3, 40, 5], peak from overlaying the pgACC clusters from Experiment 2 and 3.

precise information available during learning. In so doing, it solves the problem of balancing tonic pain with the requirement to actively learn about behaviour that could lead to relief.

The findings highlight the dual function of a state-based relief associability signal during tonic pain escape. However, a caveat to this is that associability cannot distinguish *unreliable* cues – inherently poor predictors of outcomes, and so does not discriminate between reducible and irreducible uncertainty, bearing in mind there is little adaptive logic in suppressing pain for unreliable predictors. Over extended time-frames, it is possible that the learning system recognises this and reduces endogenous control. However in rodent studies of associative learning, associability is maintained even after several days of training (Holland et al., 2002), and it is possible that salient cues in aversive situations maintain the ability to command attention and learning longer than that would be predicted by 'optimal' Bayesian models.

The localisation of the associability signal to the pgACC is consistent with *a priori* predictions. The region is known to be involved in threat unpredictability (Nitschke et al., 2006; Rubio et al., 2015), computations of uncertainty during difficult approach-avoidance decision-making (Amemori and Graybiel, 2012), and in the perseverance of behaviour during foraging (Kolling et al., 2012; McGuire and Kable, 2015). It is distinct from a more anterior region in the ventromedial prefrontal cortex associated with action value (FitzGerald et al., 2012). More importantly, it has been specifically implicated in various forms of endogenous analgesia, including coping with uncontrollable pain (Salomons et al., 2007), distraction (Valet et al., 2004), and placebo analgesia (Bingel et al., 2006; Eippert et al., 2009). However, an open question remains about the role of conscious awareness in driving pgACC-related endogenous control – a factor that is often important in these other paradigms. Whether or not the role of associability is modulated by the metacognitive awareness of uncertainty or controllability would be an important question for future studies.

The pgACC has been suggested to be central to a 'medial pain system' and the descending control of pain, with its known anatomical and functional connectivity to key regions including the amygdala (Derbyshire et al., 1997; Salomons et al., 2014; Vogt et al., 2005) and PAG (Buchanan et al., 1994; Domesick, 1969; Stein et al., 2012; Vogt, 2005). Evidence of high level of μ-opioid receptors within pgACC (Vogt et al., 2005), where increased occupation has been found in both acute and chronic pain (Jones et al., 2004; Zubieta et al., 2005), further illustrates pgACC's potential role for cortical control of pain.

The results provide a formal computational framework that brings together theories of pain attention, controllability and endogenous analgesia. Previous demonstrations of reduced pain (albeit typically for phasic, not tonic pain) have been inconsistent (Becker et al., 2015; Mohr et al., 2012; Salomons et al., 2004, 2007; Wiech et al., 2014b, 2006). Our results offer

insight into why, by suggesting that endogenous analgesia is not a non-specific manifestation of control, but rather a specific process linked to the learnable information.

From the perspective of animal learning theory, the experiments here show how motivation during the persistent pain state can be understood as an escape learning problem, in which the state of relief is determined by the offset of a tonic aversive state (Mackintosh, 1983; Solomon and Corbit, 1974). This is theoretically distinct from the better-studied form of relief that results from *omission* of otherwise expected pain or punishment (Konorski, 1967), and which motivates avoidance behaviour (Mowrer, 1960). In our task, acquisition of dissociable behavioural responses (SCRs and choices) reveals the underlying theoretical architecture of the escape learning process, which involves both parallel state-outcome and action-outcome learning components. The action-outcome learning error signal localises to a region of the dorsolateral striatum (dorsal putamen). Striatal error signals are seen across a broad range of action learning tasks, although the region here appears more dorsolateral than previously noted in avoidance learning (Delgado et al., 2009; Kim et al., 2006; Seymour et al., 2012). It is not possible to definitively identify whether avoidance and escape use distinct errors, but it is well recognised that there are multiple error signals in dorsal and ventral striatum, for instance reflecting 'model-based' (cognitive), 'model-free' (including stimulus-response habits) and Pavlovian control (Schonberg et al., 2010; Tricomi et al., 2009; Yin et al., 2004). The reinforcement learning model we describe is a 'model-free' mechanism, since it learns action values but does not build an internal model of state-outcome identities and transition probabilities (Daw et al., 2005). However, it is likely that a model-based system coexists, and might be identifiable with appropriate task designs (Daw et al., 2011).

Developing a computational account of relief learning and endogenous control may also help us understand how the brain contributes to the pathogenesis and maintenance of chronic pain (Navratilova and Porreca, 2014). Adaptive learning processes are thought to be important in chronic pain: learning and controllability have been proposed to play a role in the pathogenesis and maintenance of chronic pain (Apkarian et al., 2004; Flor et al., 2002; Salomons et al., 2014; Vlaeyen, 2015), and brain regions such as the medial prefrontal cortex and striatum have been consistently implicated in clinical studies, e.g. in pain offset responses (Baliki et al., 2010) and resting functional connectivity in chronic back pain (Baliki et al., 2008, 2012; Fritz et al., 2016; R Yu et al., 2014). In addition to suggesting a possible computational mechanism that might underlie pain susceptibility in these patients, the results highlight the pgACC as a potential target for therapeutic intervention.

5.5 Tables

Table 5.1 Details of subjective ratings.

Experiment	Rating type	Rating timing	Avg # of ratings/subject
Active relief learning in dynamic environment (Experiment 3)	Instrumental pain	After 3s cue+choice window, BEFORE outcome	70.9

Table 5.2 All learning models fitted (bold: winning model; AL - action-learning; SL - state-learning, F - variational Bayesian approximation to the model's marginal likelihood, used for model comparison, assoc - associability)

Instrumental sessions only, Pavlovian not available			
Choice	F (n=23, sum [sem])	SCR	F (n=20, sum [sem])
TD	-3572.476 [8.736]	RW - value	-7867.834 [60.668]
Hybrid (AL)	-3626.478 [8.946]	Hybrid (SL) - value	-7857.341 [60.643]
HMM	-3571.020 [9.067]	**Hybrid (SL) - assoc**	-7841.864 [60.838]
Bayesian Hierarchical	-3784.372 [8.616]		

Table 5.3 Learning model fitting results (std: standard deviation; μ_θ: Bayesian prior mean of evolution function parameter; σ_θ: prior variance of evolution function parameter; μ_ϕ: prior mean of observation function parameter, σ_ϕ: prior variance of observation function parameter).

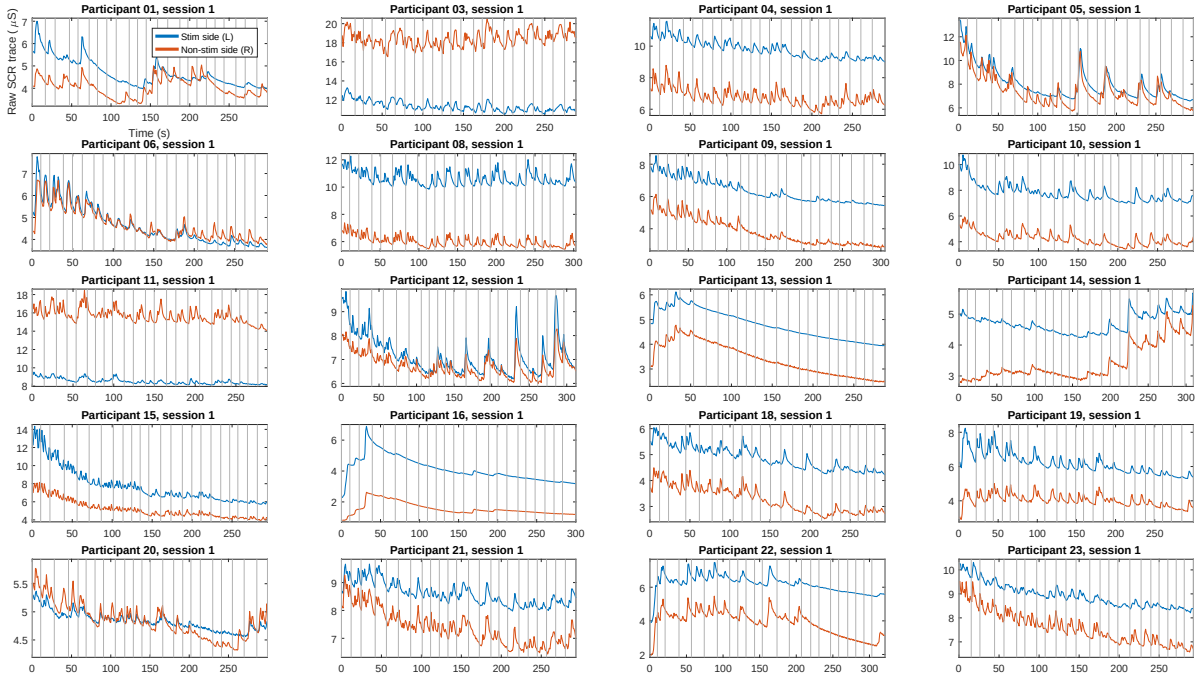
Model (Options)	Data fitted	Parameters	Mean	Std	Initial states
TD (*)	choice	learning rate, α	0.577	0.28	Q_0=0
Hybrid Action learning (*)	choice	free parameter κ	0.774	0.381	Q_0=0
		free parameter η	0.14	0.139	α_0=1
HMM (*)	choice	state transition probability β	0.275	0.213	Q_0=0.5
		relief outcome bias c	0.535	0.212	
		no relief outcome bias d	0.027	0.072	
Bayesian (‡)	choice	level 2 (outcome) κ	0.331	0.239	Q_0=0
		level 2 (outcome) ω	-0.423	1.396	
		level 3 (belief) θ	0.45	0.03	
RW - V (†)	SCR (bilateral)	learning rate, α	0.46	0.054	V_0=0
Hybrid - Assoc (†)	SCR (bilateral)	free parameter κ	0.49	0.01	V_0=0
		free parameter η	0.488	0.027	α_0=1
Hybrid - V (†)	SCR (bilateral)	free parameter κ	0.48	0.034	V_0=0
		free parameter η	0.496	0.013	α_0=1

*Fitting options: $\mu_\theta = 0$, $\sigma_\theta = 1$, $\mu_\phi = 0$, $\sigma_\phi = 1$
†$\mu_\theta = 0$, $\sigma_\theta = 0.05$, $\mu_\phi = 0$, $\sigma_\phi = 1$
‡$\mu_\theta = [0, -2, 0]$, $\sigma_\theta = 1$, $\mu_\phi = 0$, $\sigma_\phi = 1$

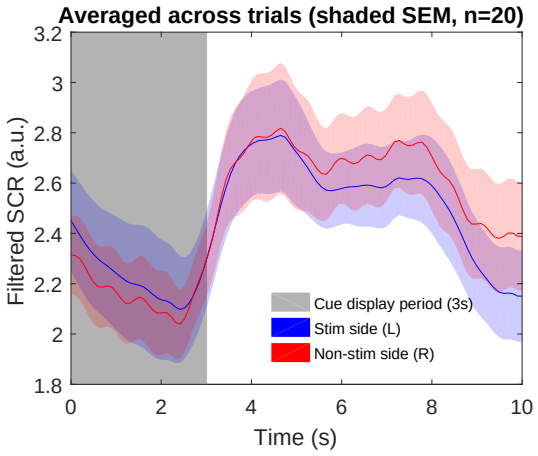
Table 5.4 Multiple correction (cluster-forming threshold of p <0.001 uncorrected, regions from Harvard-Oxford atlas. *FWE cluster-level corrected (showing p<0.05 only)

p*	k	t	Z	MNI coordinates (mm)			Region mask
				x	y	z	
TD model PE							
0.002	15	4.31	3.63	-25	-5	-22	Amygdala L
0.003	11	4.36	3.66	24	-8	-14	Amygdala R
0.018	1	3.97	3.41	28	-1	-26	
0.002	22	5.9	4.52	-32	-8	5	Putamen L
0.021	4	4.55	3.78	32	-16	1	Putamen R
Hybrid model PE							
0.001	16	4.36	3.66	-21	-12	-14	Amygdala L
		4.23	3.58	-21	-1	-18	
0.002	13	4.95	4.01	24	-8	-18	Amygdala R
		4.34	3.65	28	-1	-26	
0.003	17	5.49	4.31	-32	-8	5	Putamen L
Hybrid model associability							
0.001	29	4.5	3.75	-6	40	12	Cingulate Anterior
		4.44	3.71	-2	33	23	
		4.08	3.49	-2	44	5	
		3.93	3.38	2	40	1	

5.6 Supplementary figures



(a)



(b)

Fig. 5.4 Skin conductance raw data. (a) Raw skin conductance traces, where vertical lines are beginning of each trial when cue display starts (n=20, excluded participants not shown, showing first non-excluded session from all participant). (b) Filtered skin conductance traces (band-pass at 0.0159-2Hz, 1st order Butterworth), averaged across all trials within participant (n=20, excluded participants not shown, shaded region represent SEM across all participants). In the current experimental design, pain ratings were collected after cue display period, and participants could terminate rating whenever they finished. These variable time gaps between cue display and outcome account for the second peak in trial averaged SCR trace.

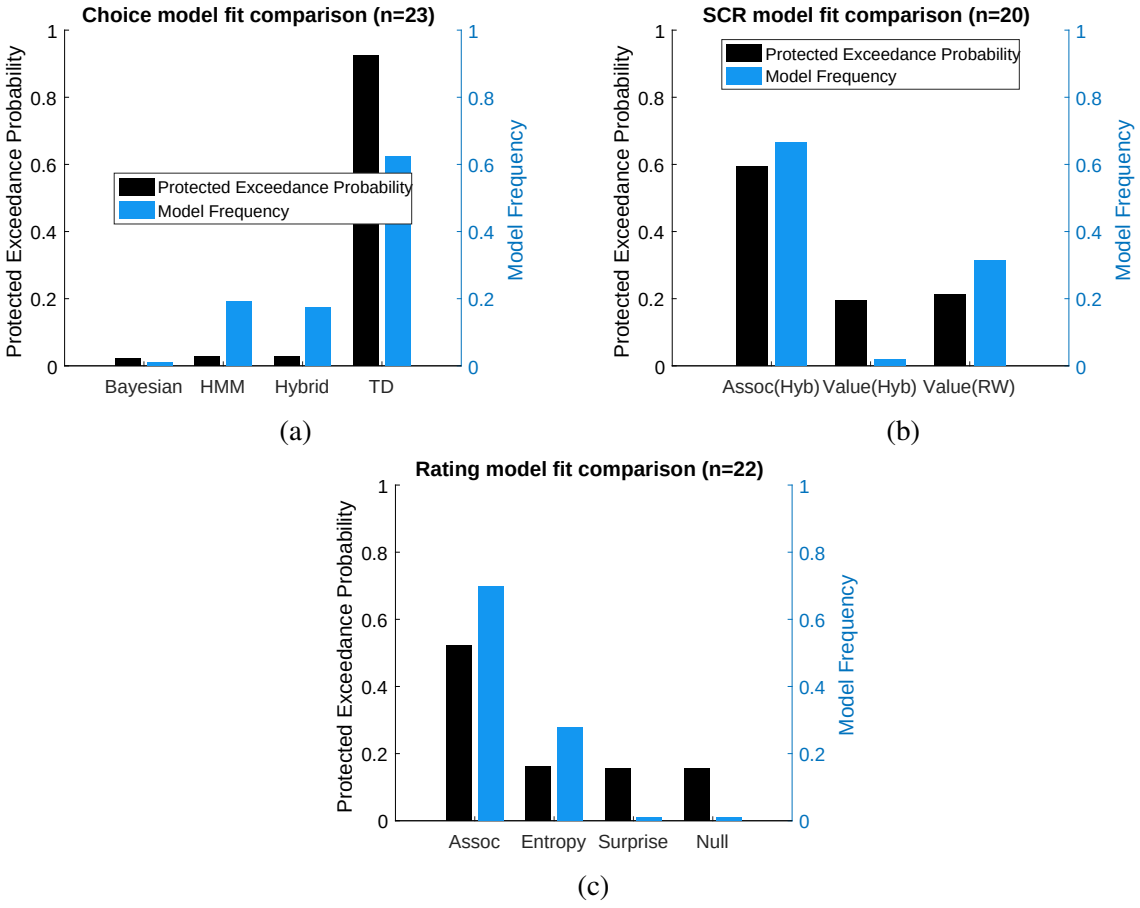
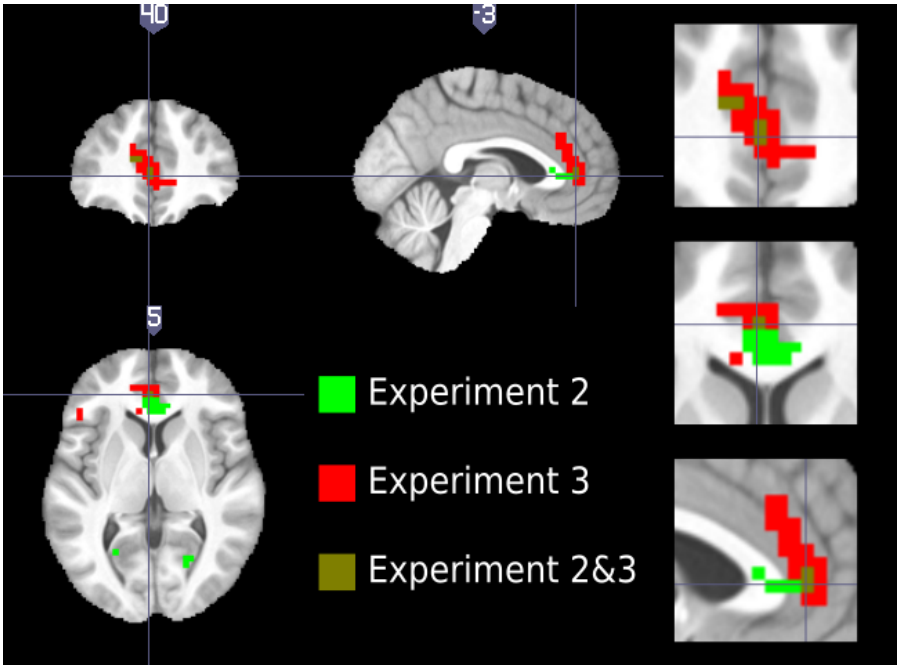
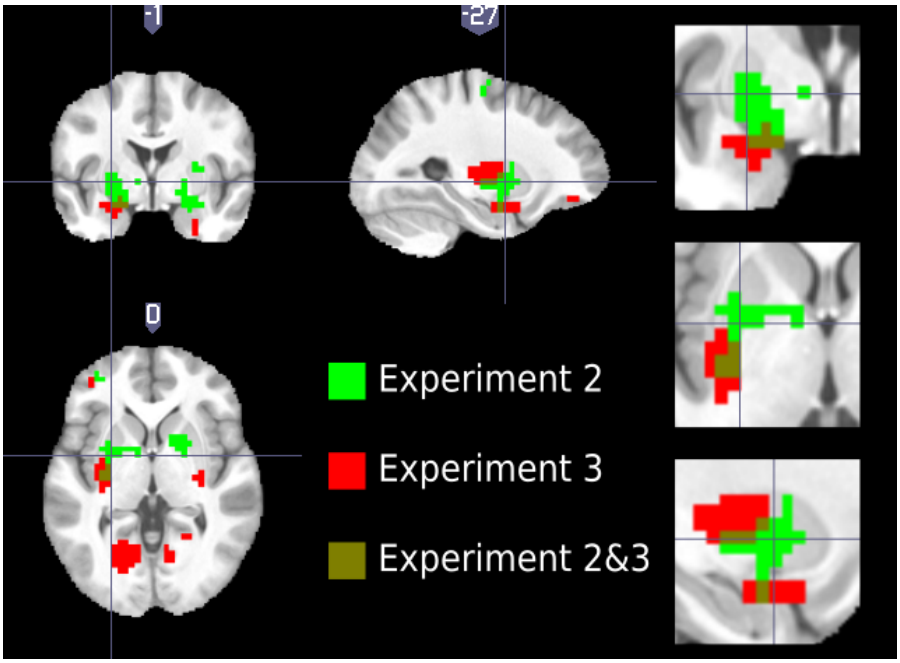


Fig. 5.5 Model protected exceedance probability. (a-c) Choice, SCR, rating fitting all remain similar to original exceedance probability figures.



(a)



(b)

Fig. 5.6 Overlaying clusters from both relief learning experiments to show similarities. (a) Overlaying associability associated pgACC responses from both relief learning experiments (displayed at p<0.001 unc., crosshair at [-3, 40, 5]). (b) Overlaying prediction error associated responses (displayed at p<0.001 unc., showing overlapping dorsal putamen and amygdala clusters).

Chapter 6

Experiment 4: Endogenous controllability of brain-machine interfaces for pain

A fundamental property of human pain is the ability to endogenously modulate perceived intensity as a function of cognitive control. It is especially important for the design of closed-loop brain-machine interfaces that aspire to use brain activity to control therapeutic devices, since it could act to substantially enhance or degrade performance. It remains unclear exactly what aspect of the neural representation of pain is under control. We studied brain activity whilst subjects received intermittent pain stimuli in an fMRI-based real-time neurofeedback task, in which multi-voxel pattern decoding of pain intensity was used to train a control algorithm to learn to deliver lower intensity stimuli (adaptive decoded neurofeedback). This created the incentive for subjects to enhance the neural discriminability of pain intensity, but we found that this was only achieved in a single brain region – the pregenual anterior cingulate cortex (pgACC). In contrast, discriminability was either reduced or unchanged in classic pain-processing regions, including insula, dorsolateral prefrontal, and somatosensory cortex. Furthermore, we also found that pain perception was modulated as subjects learned the success rate of the machine reducing high-intensity pain. The results indicate a primary role for the pgACC in the endogenous control of pain, and illustrate how regionally specific co-adaptive brain-machine learning can have a critical effect of the efficacy of closed-loop systems for pain.

Adapted from Zhang et al. (2018b)

6.1 Introduction

A striking feature of the pain system is its capacity for endogenous modulation, whereby the perception of pain can be reduced or increased by higher brain control processes (deCharms et al., 2005; Tracey and Mantyh, 2007; Wiech, 2016; Woo et al., 2017). It is thought to reflect a key mechanism by which animals cope with threat, and is mediated in part by control of descending pathways to the spinal dorsal horn neurons that transmit incoming nociceptive signals (Basbaum and Fields, 1984; Ossipov et al., 2010). Identifying the central cortical control site(s) is important, not least for the innovation of new pain treatments. But whereas a number of brain areas are clearly sensitive to subjective perception of modulated pain, it isn't clear whether they are functionally involved in control, or merely reflect the consequence of it. Importantly, it is not known whether *discriminability* between different intensities is under control – a possibility that would suggest a direct modulation of the informational representation of pain in the brain.

This issue is particularly important for the design of 'closed-loop' brain-machine interfaces, which aim to use brain-based activity patterns to dynamically control pain interventions (Zhang and Seymour, 2014). Recent studies have shown that multivariate pattern analysis (MVPA 'decoding') methods can be used to discriminate pain-related brain (BOLD) responses in humans with reasonable accuracy (Brodersen et al., 2012; Marquand et al., 2010; Schulz et al., 2012; Wager et al., 2013), and this can in principle be used to guide interventions in closed-loop settings where immediate intervention delivered is adjusted based on the real-time decoded pain state. If a person is aware that their brain is being decoded, however, the incentive is for them to learn to endogenously increase the discriminability of pain-related brain information, and therefore enhance the performance of any machine application that uses this information. But if pain discriminability patterns are not primarily under active control, then performance would be potentially susceptible to other adaptive changes associated with learning that might actually reduce discriminability – for instance modulation by expectancy effects in which pain predicts its subsequent relief (Atlas and Wager, 2012; Colloca et al., 2008). Consequently, understanding the nature of cognitive control of neural pain representations is of fundamental importance to predicting the performance and stability of closed-loop systems over time.

To identify whether pain discriminability could be enhanced through cognitive control, we designed a co-adaptive brain-machine interface system using real-time functional brain imaging (rtfMRI). Accordingly, a machine learned to reduce the intensity of experimental pain stimulation based on decoded BOLD representations of pain in the brain, creating the incentive for subjects to modulate their brain responses. We studied whether and where neural discriminability of pain increased when subjects were aware of the system, and the associated impact on the perception of pain itself.

6.2 Methods

Subjects

19 healthy participants enrolled in a two-day neuroimaging experiment (two female, age 23.5±4.0 years). All subjects gave informed consent prior to participation, had normal or corrected to normal vision, and were free of pain conditions or pain medications. The experiment was approved by the Ethics and Safety committee of the Advanced Telecommunications Research Institute, Japan.

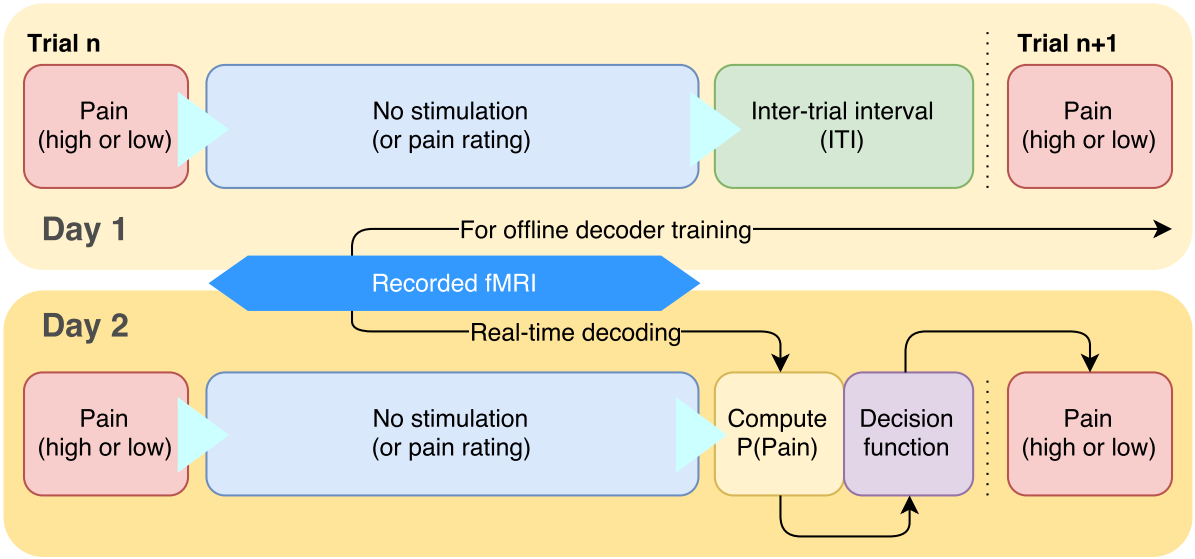
Experimental design

The experiment spanned two days. On each day, each participant completed 2 sessions of pain thresholding test outside the scanner and 6 sessions of task with high/low painful stimuli inside the scanner.

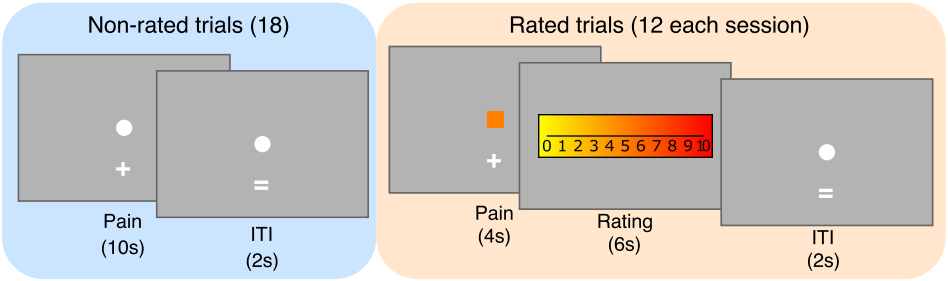
Day 1: Decoder construction

Participants received a sequence of high and low intensity painful electrical stimuli to the left hand, with no associated task demands other than intermittent pain ratings (Fig 6.1a). For decoding, we used BOLD responses in bilateral insula cortex, since this region is thought to have a primary role in the coding of pain (AD Craig, 2002; Geuter et al., 2017; Segerdahl et al., 2015; Woo et al., 2017). Individual participant's responses to high and low intensity stimuli were subsequently used to train a voxel-wise MVPA decoder that could classify the two stimulus levels (see 'Decoder construction').

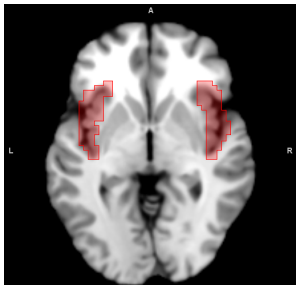
From the participant's perspective, a painful stimulus was delivered at the beginning of each trial when a '+' symbol appear on screen below the white bulls-eye fixation point. The '+' stayed on for 10s, then the '=' symbol replaced it for 2s, signalling a brief inter-trial interval (ITI). In 40% trials (12 randomly chosen out of 30 in each session), the '+' stayed on screen for 4s and the fixation point turned to an orange square (signalling upcoming rating), followed by a 0-10 visual analogue scale that stayed on for 6s, where participants were asked to rate how painful the stimulus was by pressing two buttons to move the slider on screen (Fig 6.1b). Brain images from 4-10s after pain delivery were used for both decoder training and real-time decoding. This allowed for BOLD delay and avoid movement contamination. The 30-trial session was repeated 6 times with a short break in between (180 trials, 72 ratings per subject in total). All participants were given the instruction to rest in the scanner and do nothing (see 'Participant instructions and survey questions').



(a)



(b)



(c)

Fig. 6.1 Experimental paradigm. (a) Participants took part in a two-day decoding neurofeedback experiment. On Day 1, their functional brain images were recorded while they were given high or low level of electrical pain at the beginning of each trial, which were used for offline MVPA decoder training. On Day 2, participant's probability of having received high pain (P(Pain)) in the current trial was computed in real-time from their brain activities by the decoder, which was then used by the pain delivery system to decide on the pain level of the next trial. The decision function of the system was based on a reinforcement learning algorithm that aimed to lower the participant's decoded P(Pain). (b) For 40% of all trials in a session, participants were asked for their pain ratings 4s after its delivery on a 0-10 VAS scale. The changed fixation point acted as a prompt for rating. The display were identical on both days. (c) Demonstration of bilateral insula ROI generated from AAL atlas (viewed at z=0).

Day 2: Neurofeedback adaptive control

High and low pain stimulation were embedded into a closed-loop adaptive control paradigm using real-time MVPA output as the feedback signal. Specifically, a machine controlling the two levels of pain stimulation tried to learn to deliver the level with the least pain based on the output of the MVPA classifier (the Day 2 scan was carefully realigned with Day 1 to allow generalisation across days). After delivering pain stimulus, the participants' probability of experiencing high pain from that stimulus (P(Pain)) was estimated by the decoder with insula BOLD activity from their brain images in that trial (realigned and resliced to the reference image from Day 1, following Shibata et al., 2011). The estimated probability was used as reward in the decision function (with lower probability of high pain equating to a more positive reward).

The reason for using an adaptive decision function was two-fold: to maximise the incentive for subjects to enhance discriminability of pain; and to show that in principle, pain-reducing interventions can be learned by an appropriate control algorithm. Specifically, stimulation was based on a reinforcement learning algorithm that learned the values for each of the two stimulus levels, initialised at zero at the beginning of each session (i.e. the machine doesn't 'know' which level results in less pain). The details of this algorithm are described below, but in brief, the stimulation state that elicited lower decoded pain signal in the participant was reinforced (see 'Neurofeedback adaptive control').

To allow meaningful comparisons to be made between the brain responses on the two days, bearing in mind their necessary sequential order, we yoked a previous participant's Day 2 pain sequence as the sequence for the current participant on Day 1. Day 1 and 2 had the same trial structure (Fig 6.1a), apart from the adaptive control process and subject instructions. The yoking procedure preserves the balance of high/low pain stimuli and their temporal sequences between days, which allows investigation of whether any brain-machine co-adaptation processes took place. 16 out of 19 participants used another participant's Day 2 trial sequences on Day 1 as yoked control. All participants were explicitly told that the pain level they received was controlled by the computer programme, and were aware that modulating their brain activity could therefore influence the computer (see 'Participant instructions' below).

Stimulus delivery

Painful electrical stimuli were delivered using two constant current stimulators (Digitimer model DS7A, Welwyn Garden City, Hertfordshire, UK), at two current levels for high/low pain determined using the participant's own threshold. On each day, the current levels were set according to the pain threshold test conducted before experiment started. The current levels

were fixed across sessions (except in 4 subjects, minor adjustments were made where pain ratings were either too high, or there were no difference between two levels). All stimuli were delivered as a train of 50 5ms square wave pulses at 10Hz, lasting 500ms (DS7 settings: x1 mA, 200μs).

The two stimulators were connected to a custom-made switch that allowed current delivery through the same custom-made, MRI-compatible ring electrode (10mm diameter). The electrode was taped to the back of the left hand of the participant, its location marked on Day 1 as reference for attachment on Day 2.

Pain thresholding (Day 1 and 2)

On each day, participants completed a thresholding procedure at the beginning of the experiment. In the first session, the staircase method was used to evaluate their highest pain limit. Stimuli current were linearly increased at 0.2-0.5mA interval, and participants were asked for verbal feedback of a 0-10 pain rating in person after each stimulation. This procedure was rerun a few times using different starting points and both stimulators. In the second session, 14 trials of randomised painful stimuli were given within the range of lowest perceivable to highest tolerable current level determined in session 1. Subjects rated each stimulus 1s after receiving it, on a 0-10 VAS scale on screen using a keyboard (as practice to the rating procedure used in the task). To determine the final current level to use, both a Weibull and a sigmoid function were fitted to session 2's stimuli and ratings, and current levels for VAS = 1 and 8 were used for low / high pain stimulus for the experiment respectively. The same procedure was repeated for Day 2, and the new fitted current levels were used.

fMRI data acquisition (Day 1 and 2)

Neuroimaging data was acquired with a 3T Siemens Prisma scanner with the standard 64 channel phased array head coil. Whole-brain functional images were collected with a single echo EPI sequence (repetition time TR=2000ms, echo time TE=26ms, flip angle=80, field of view=240mm), 33 contiguous oblique-axial slices (voxel size 3.2 $\times$ 3.2 $\times$ 4 mm) parallel to the AC-PC line were acquired. Whole-brain high resolution T1-weighted structural images (dimension 208 $\times$ 256 $\times$ 256, voxel size 1 $\times$ 1 $\times$ 1 mm) using standard MPRAGE sequence were also obtained.

Decoder construction (Day 1)

Preprocessing All preprocessing were conducted using SPM12 (http://www.fil.ion.ucl.ac.uk/spm/software/spm12/) in MATLAB (The MathWorks Inc., Natick, MA, USA).

All functional images were realigned and resliced to the reference functional volume (the first baseline TR after the first 3 dummy TRs obtained in the first session on Day 1). Structural T1 images were coregistered and segmented according to the canonical single subject T1 images. The resulting inverse transformation matrix was used to normalise the bilateral insula ROI obtained from the Automated Anatomical Labeling (AAL) atlas from MNI space to individual subject space. These warped ROI images were then coregistered to the reference functional TR.

Feature extraction Time series were extracted from all voxels within the individual's insula ROI. To account for BOLD delay and to minimise motion contamination, the times series from TR 3-5 (4-10s) were used from each trial, the first two TRs (0-4s) immediately following pain stimulus were omitted. For denoising, the 5 TRs following 3 dummy TRs at the beginning of each session were used as baseline, each trial ROI time series were normalised by subtracting session baseline mean and divided by baseline standard deviation, then the mean across the TR 3-5 from all trials were extracted for classifier training.

Decoder training Mean insula voxel activity as feature and high/low pain delivered as label were aggregated across all trials within participant for decoder training. Binary classification by Sparse Logistic Regression (SLR) with variational parameters approximation (Yamashita et al., 2008) was used. This results in a sparse matrix of weights for about 5 percent of all voxels within the given ROI. By multiplying weights with feature/voxel intensity signals, the decoder produces the probability of observing current label given trial features (referred as (P(pain) from here, P(pain)=1 means highly likely to have received high pain, P(pain)=0 means unlikely to have received high pain, or highly likely to have received low pain).

For decoder training, all trials were used for training. To estimate decoder accuracy, all trials were partitioned into 10 equal sets with 9 sets for training and 1 set for testing (10 fold cross-validation). The average testing accuracy of 10 iterations of cross-validation were used as estimated decoder accuracy (Table 6.1). The trained decoder was tested with another day's data using the experimental setting.

Neurofeedback adaptive control algorithm (Day 2)

To allow automated adaptive control of pain stimulus delivery, we used a simple reinforcement learning algorithm (Sutton and Barto, 1998) to update the value of high/low pain states trial-by-trial:

$$Q_{t+1}(a) = Q_t(a) + \alpha(-P(pain) - Q_t(a)) \tag{6.1}$$

where t represent trials, Q is the value of given state, a is the actions available for the algorithm (i.e. either giving high or low pain, collectively shown as action set A), α is learning rate fixed at 0.5.

P(pain) is the decoder-generated probability of current trial's stimulus being high pain. It's scaled between [-1,1] when used in the updating function. Higher P(pain) would decrease the value of the current pain state more and vice versa, while the value of the unchosen state remained unchanged.

The algorithm selects which pain level to deliver for the next trial using an ε-greedy action selection rule based on current values:

$$p_{t+1}(a|Q_t) = \begin{cases} \text{random action } a \in A, & \text{if } \xi > \varepsilon \\ \text{argmax}_{a \in A} Q_t(a), & \text{otherwise} \end{cases} \tag{6.2}$$

where ε is the explore ratio fixed at 0.4 (i.e. exploring by choosing a random action by either giving high or low pain 40% of the time, exploiting the other times), ξ is a uniform random number drawn within [0,1] at each trial. The random exploration allows a sufficient proportion of alternative pain level to be delivered, to ensure the next participant who uses current participant's Day 2 sequence to have enough trials of both high and low pain for decoder construction. We also set values to be 0 for both states at the beginning of each session.

Frequency learning model

The frequency learning model M assumes a participant estimates the posterior distribution of a given stimuli θ from a previously observed sequence of two possible stimuli $y_{1:t}$ (i.e. high or low pain) using Bayesian updating (Mars et al., 2008; Meyniel et al., 2016).

$$p(\theta|y_{1:t}, M) \propto p(y_{1:t}|\theta, M)p(\theta, M) \tag{6.3}$$

Given the experiment design, participants are assumed to have uninformative prior over the two stimuli at the beginning of each session, which can be represented by a Beta distribution with parameters [1,1]. Since the product of two Beta distributions results in a Beta distribution, the

posterior distribution depends only on the frequency of the high and low stimuli N_h, N_l, which has an analytical solution. The posterior mean of the predicted high pain distribution is:

$$p(h|N_h,N_l) = \frac{N_h+1}{N_h+N_l+2} \tag{6.4}$$

and $P(l|N_h,N_l) = 1 - p(h|N_h,N_l)$ given the reciprocal relationship between high/low pain stimuli.

The uncertainty/surprise of current stimulus h/l at trial t can be estimated as the entropy H of the posterior mean before updating from trial $t-1$:

$$H(P(h_t)) = -log_2(P(h_{t-1})) \tag{6.5}$$

This model does not require model fitting, as participants were assumed to cumulate stimulus counts over the entire session (30 trials), where we assumed perfect memory retention. It is possible to limit the number of trials for frequency memory, or introduce a forgetting 'leaky factor' to discount previously experienced trials. However, given that we had no other behavioural data for fitting apart from selective pain ratings, and relatively short sessions, we decided to use the simplest frequency model without fitted parameters.

To determine any learning effects on subjective ratings, we followed the method in Woo et al. (2017) to use subjective rating residuals for correlation analysis with learning model predictors. We regressed subjective ratings with a matrix of high/low pain stimulus identities (high=1, low=-1), and session numbers (1-6) for each individual to obtain rating residuals. The fluctuation of the resulting residuals can be interpreted as modulatory effects on pain beyond the level of nociceptive inputs.

Behavioural data

Behavioural data were analysed using Python 3.6, with pandas 0.19.2, scipy 0.18.1, afex 0.16-1.

fMRI data offline analyses

Preprocessing For offline analysis, functional images were preprocessed using the fmriprep software (build date 21/05/2017, pypi version 0.4.4, freesurfer option turned off, https://github.com/poldracklab/fmriprep), a pipeline that performs slicetime correction, motion correction, field unwarping, normalisation, field bias correction, and brain extraction using a various set of neuroimaging tools available. The confound files output by fmriprep include the following signals: mean global, mean white matter tissue class, three FSL-DVARS (stdDVARS, non-stdDVARS and voxel-wise stdDVARS), framewise displacement, six FSL-tCompCor, six

FSL-aCompCor, and six motion parameters (matrix size 24 × number of volumes). Resulting functional images were smoothed with an 8mm Gaussian kernel in SPM12, except for those in used searchlight analysis.

fMRI GLM model

All event-related fMRI data were analysed with GLM models constructed using SPM12, estimated for each participant in the first level. Stick functions at pain stimulation onset were convolved with a canonical hemodynamic response function (HRF). We also included rated trials (duration=10s, from beginning until ITI) as regressor of no interest, in addition to the 24 columns of confound matrix output by fmriprep. Day 1 and 2 data were included in the same GLM, but first-level contrasts were estimated separately for days.

Whole-brain comparison (Fig 6.3) 2 regressors: high/low pain onset (duration=0).

Switch trials differences (Fig 6.5) 4 regressors: trials stimulus different from or identical to that of the previous trial were labelled as switch or non-switch trials, separately for high/low pain (HH, LL, LH, HL), at pain onset (duration=0).

Frequency learning posterior probability and entropy (Figs 6.6b, 6.6c, 6.6d) 3 regressors at pain onset (duration=0) with parametric modulators: posterior probability of current stimulus (updated prediction), entropy of previous posterior probability of current stimulus (uncertainty of prediction before updating), actual identity of stimulus (high pain=1, low pain=-1). All parametric modulators mean centred within session, SPM orthogonalisation for these 3 regressors were turned off.

For multiple comparison, we used anatomical binary masks generated using the Harvard-Oxford Atlas (Desikan et al., 2006) (freely available with the FSL software, https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases), and periaqueductal grey probabilistic atlas (Ezra et al., 2015) for small volume correction. All probability maps were thresholded at 50%, and all masks were applied separately, not combined. We used the frontal medial cortex mask as approximation for the ventromedial prefrontal cortex (VMPFC). Bilateral masks for ventrolateral and lateral periaqueductal grey (vlPAG and lPAG) were combined respectively. We also used the pregenual anterior cingulate cortex (pgACC) peak identified in our previous study of active relief learning (Zhang et al., 2018a) for the 8mm spherical ROI mask (sphere peak used: [6,40,12]). We reported all results with $p<0.05$ (FWE cluster-level corrected), with the exception of searchlight analysis results (MFG/DLPFC SVC had p=0.06, see Table 6.2).

ROI analysis

Beta estimates were extracted from activation ROIs (see text for mask details). Beta values plotted were the average of all voxels within ROI masks, with statistics showing subject-level SEM. Post-hoc repeated measure ANOVA were conduced with the R package 'afex'. pgACC activations were overlaid on subject-averaged anatomical scans using MRIcroGL (https://www.nitrc.org/projects/mricrogl/).

Decoder comparison

Decoders were constructed using Day 2 data with the same procedure as Day 1 (Fig 6.4). This was done to determine whether the decoding performance of insula ROI remained the same, or whether any learning-induced changes might have changed the decoder properties.

Whole-brain searchlight analysis was conducted using the Decoding Toolbox (TDT, v3.98) in MATLAB (Hebart et al., 2015). The toolbox can conduct multivariate decoding analyses at combined trial types within fMRI runs, by extracting features from beta images of relevant regressors in the first level GLM analysis output by SPM. This could lead to higher classification accuracy and lower computation time, comparing to single trial decoding.

A searchlight analysis was carried out within a 10mm radius sphere for the whole brain, with high/low pain categories as unsmoothed beta images from each run for individual participant. TDT produced a decoding accuracy map for each voxel using a leave-one-run-out cross validation scheme, which can be interpreted as the local information content of each voxel (Kriegeskorte et al., 2006). The Day 1 and 2 accuracy maps from each individual were then smoothed with a Gaussian kernel of 4mm, and entered into a standard SPM second level paired t-test as in the GLM analysis above. The resulting T map indicates the changes in decodable information used for pain level decoding across days.

Participant instructions and survey questions

Day 1 (Decoder construction)

Please rest in the scanner. We are looking at your brain's response to different levels of pain. You don't have to do anything.

Day 2 (Adaptive control)

You don't need to do anything in this task. The computer is trying to work out if you feel pain or not, by looking at your brain activity. If it thinks you felt pain, it will try and change the pain

stimulation to stop you from having pain. If it thinks you did not feel much pain, it will try not to change anything. However, it cannot do this very reliably, as reading the brain activity is difficult, so it may often make mistakes.

During your first scan, we gave a random sequence of pain stimuli – some high, and some low. Using this data, we have trained a computer program to tell how much pain you were feeling during each shock, based on your brain activity. It is good, but not perfect – it gets it right about 80% of the time.

In today's scan, the computer program can influence the pain level you get. If it thinks you felt a lot of pain, it will influence the pain machine to give you less pain in the future. If it thinks you did not feel much pain, it will try to influence the pain machine to continue to give you little pain. In other words, it is trying to help you get less pain! This is a difficult job for the computer program, because it is not perfect at reading your brain activity as soon as it is active (i.e. within a few seconds).

It is up to you what you do in the task. You can do nothing, and hope that the system works well, and the computer learns to reduce the pain. Or you can try to influence the computer using your thoughts, in any way that you like.

Post-training survey (Day 2)

1. Do you think the machine was successful in reading your pain and trying to reduce it?
2. Did you try to influence the computer by doing or thinking anything?
3. If so, what did you do/think?
4. And if so, do you think you were successfully able to influence it?
5. Any other comments or feedback?

6.3 Results

Behavioural results

Within-subject decoder construction based on the insula ROI achieved reasonable classification accuracy (Day 1: 10-fold cross-validated test accuracy 65%, sensitivity 60%, specificity 67%, accuracy one-sample t-test vs 0.5 across subjects: $T(18)=8.967$, $p<1e\text{-}7$) (see also Table 6.1). When this classifier was used during neurofeedback when the subject returned on Day 2, decoding accuracy remained above chance (Day 2: accuracy 56%, sensitivity 51%, specificity 63%, accuracy t-test vs 0.5: $T(18)=4.053$, $p=0.0007$). Specifically, the real-time decoder output

following delivery of high pain (referred to as P(pain), Fig 6.2a) differed significantly for the high and low pain stimuli (repeated measure ANOVA of session and pain level effects, only pain level main effect significant: F(1, 18)=17.41, p=0.0006, post-hoc test Bonferroni corrected P(pain) for high pain 95%CI=[0.558, 0.676], low pain=[0.434, 0.551]).

Decoder performance was therefore sufficient for the reinforcement learning control system to learn differential values for high and low pain stimuli within a few trials in each session (Fig 6.2b, mean ± SEM, high pain=-0.264±0.0486, low pain=-0.0608±0.0479, paired t-test: T(18)=-3.651, p=0.0018). Accordingly, the control system delivered significantly fewer high compared to low pain stimuli (Fig 6.2c, number of high pain trials minus low pain trials: Day 1=-3.526±1.436, Day 2=-3.912±1.412, one-sample t-test vs 0: Day 1 T(18)=-2.455, p=0.0245, Day 2 T(18)=-2.771, p=0.0126).

Table 6.1 Decoder testing performance (high pain = positive, low pain = negative for sensitivity/specificity calculation; CV: 10-fold cross validation; D1: Day 1; D2: Day 2. All values are mean (SEM), n=19)

	Train D1, Test D1 (CV)	Train D1, Test D2	Train D2, Test D2 (CV)	Train D2, Test D1
Accuracy	0.6488 (0.0163)	0.5632 (0.0156)	0.5601 (0.0100)	0.4912 (0.0314)
Sensitivity	0.6017 (0.0258)	0.5057 (0.0159)	0.4982 (0.0313)	0.4379 (0.0258)
Specificity	0.6654 (0.0248)	0.6309 (0.0365)	0.5903 (0.0246)	0.5490 (0.0310)
# features (voxels)	24.053 (1.053)		28.737 (0.700)	

An important feature of our experimental design was to have matched sequences on Day 2 (neurofeedback) to Day 1 (decoder construction), to allow meaningful comparisons to be made. To achieve reasonable classification learning, therefore, we set the decision function (i.e. by which action values determine the actual machine choice of pain level) for the control system to be noisy, so that despite the relatively clear difference between action values, there were sufficiently large number of high stimuli delivered that would support decoder training performance when the sequence was used for a subsequent subject's Day 1 decoder construction. Note that the 3 initial subjects did not have yoked stimulus sequences, instead they used randomly generated 50/50 high/low pain sequences.

Participants completed pain threshold testing on both days before the experiment started, with aims to achieve VAS=1 and 8 for low/high pain respectively. Consequently, pain ratings were significantly different between the high and low levels of pain (Fig 6.2d, high pain=5.92±0.356, low pain=1.99±0.258, repeated measure ANOVA Pain levels: F(1,11)=86.00, p<1e-5), but did not show significant differences across sessions or days (day: F(1,11)=3.173, p=0.103, session: F(5,55)=0.470, p=0.797). There were no significant differences in the high/low pain stimulation current levels given between days (paired t-test p=0.12 and 0.27 respectively).

We conducted a questionnaire survey (see Methods) after Day 2's experiment concluded, in which participants showed evidence of awareness of and attempt in mentally influencing the system to reduce overall pain. 17 out of 19 (1 ambiguous) of all participants believed the machine was successful in reading their pain and trying to reduce it, 15 out of 19 (2 ambiguous) believed that they were successful in influencing the system to achieve that using some mental strategies. The strategies used most frequently included a combination of mental imagery of pain, distraction from pain, predicting/recalling stimulus sequence, and doing nothing.

Whole-brain comparison between days

Offline whole-brain analysis of fMRI data using a conventional general linear model showed evidence of a regional day × pain level interaction (Fig 6.3a). Specifically, within-subject comparison (Day 2 - Day 1) of the contrast (high pain - low pain) showed increased responses in PAG (statistics in figure legend, and correction for multiple comparisons detailed in Table 6.2), probably localising in dorsolateral or lateral PAG (Fig 6.3b), using PAG subdivision masks from (Ezra et al., 2015). The PAG is an important *a priori* ROI in this analysis, as it is a key part of the descending endogenous control system.

In contrast, we found decreased BOLD responses in the left amygdala and bilateral putamen (Fig 6.3a). Responses in the bilateral insula ROI, of which the decoder computed P(Pain) from, were not significantly different between days for either high or low pain, or overall (pain level main effect: $F(1,18)=63.911$, $p=2.475e-7$, session main effect: $F(5,90)=4.130$, $p=0.002$, none of the interactions significant. Note that images used in the GLM and subsequent analyses were fully preprocessed, as opposed to the limited (i.e. rapid) processing used in real-time neurofeedback).

Decoder comparison

To identify potential changes in the multi-voxel patterns in the insula, we compared MVPA decoder performance trained offline using the bilateral insula ROI on functional brain images from Day 1 and Day 2 (i.e. training the decoder separately on each day, and using cross-validation test sets to estimate performance). This analysis aimed to detect changes in the representation of pain, distinct from the mean BOLD signal across all ROI voxels, and despite the fact that there were no significant changes to pain stimuli or rating between days. Using a 10-fold cross-validation, we found that the decoding test accuracy decreased on day 2 compared to Day 1 (Fig 6.4a and Table 6.1, Day 1: 64.8%, Day 2: 56.0%, Wilcoxon signed-rank test $Z(18)=3.69$, $p<0.001$). Furthermore, the decoder trained with Day 2 data identified significantly more voxels as contributing to decoding performance compared to Day 1 (Fig 6.4b and Table

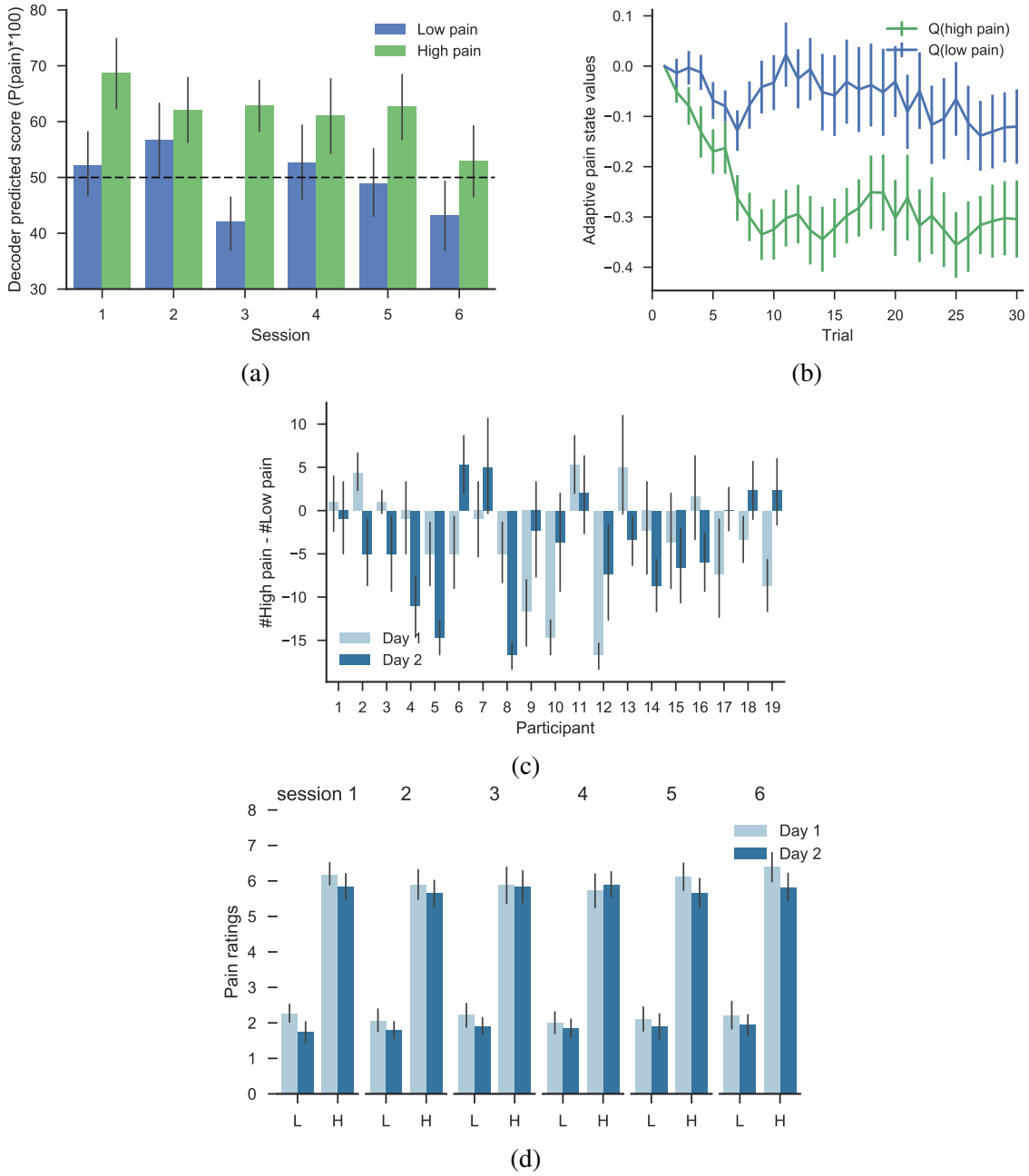
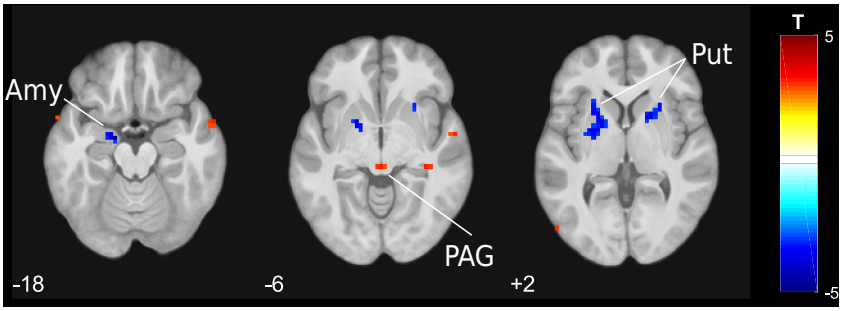
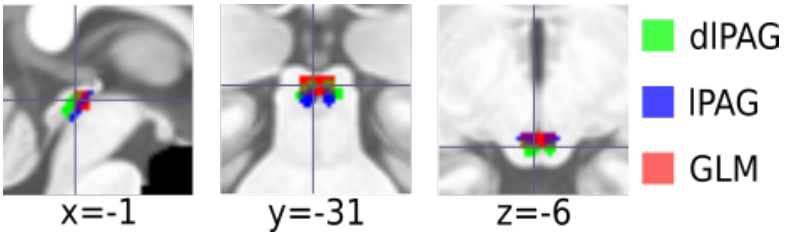


Fig. 6.2 Behavioural results (mean ± SEM across n=19 individuals). (a) Decoder predicted probabilities of having received high pain, P(pain), were able to distinguish high/low pain state (calculated on Day 2 only). (b) Within-session, the control system learned to value low pain states higher than high pain states (Q(low pain) > Q(high pain)) after several initial trials (Day 2). (c) Number of high pain trials minus low pain trials delivered to all participants on both days, where low pain levels were delivered more frequently (participant 1-3 used random stimulus sequences on Day 1 instead of yoked, error bars represent SEM calculated across sessions for a participant). (d) Raw pain ratings were significantly different for the two pain levels, but not across sessions or days (H=high pain, L=low pain).



(a)



(b)

Fig. 6.3 Whole brain comparison between days results, (a) Within-subject comparison (2nd level paired t-test, Day 2 - Day 1) of the high pain - low pain 1st level contrasts, interaction were observed in left amygdala (peak coordinates [-16, -7, -18], T=-4.38, k=11), bilateral putamen (left peak [-26, 12, 6], T=-4.09, k=47, right peak [26, 19, -2], T=-4.50, k=32) and the PAG (peak coordinates [0, -30, -6], T=3.27, k=3). All images shown at $p<0.005$, $k>0$. (b) Dorsal lateral PAG and lateral PAG masks (thresholded at 50%) from (Ezra et al., 2015) overlaying PAG activation from GLM (cluster formation at Z=2.9). SVC using bilateral dlPAG mask p(FWE-corr)=0.034, k=1, T=3.14, Z=2.76, peak in [-3,-30,-6], k=1, T=2.94, Z=2.62, [3,-30,-6], lPAG mask (FWE-corr)=0.036, other statistics the same as dlPAG mask. (note our voxel size is too big for very serious PAG subdivision). H: high pain, L: low pain.

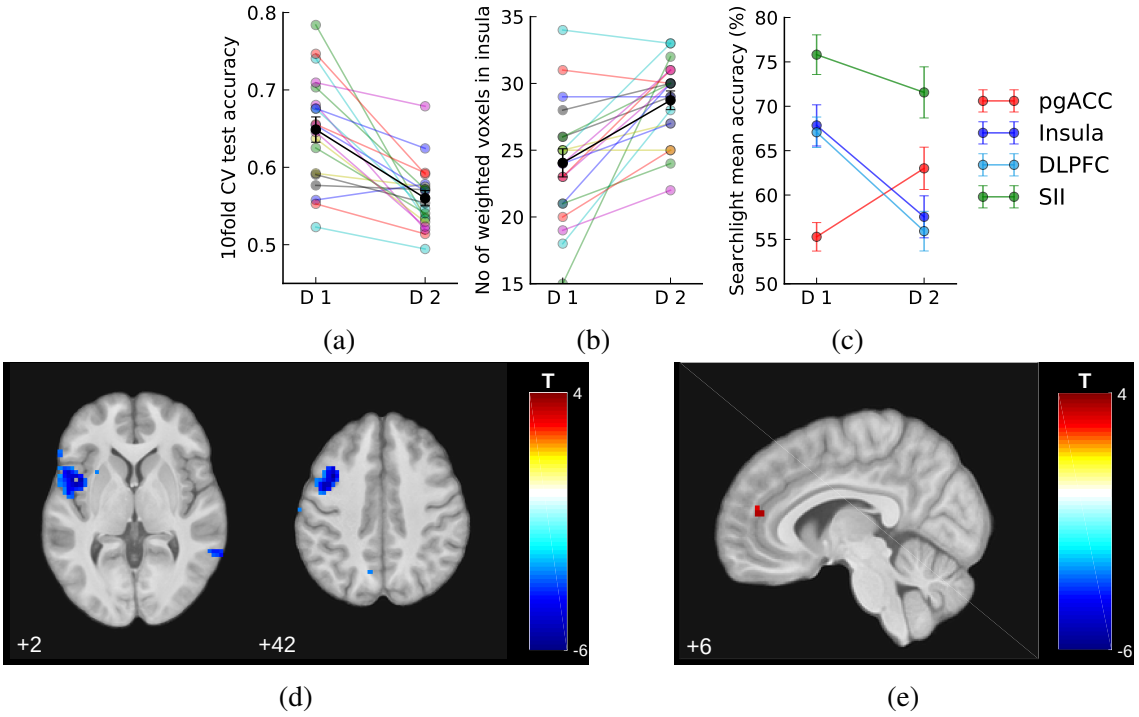


Fig. 6.4 Decoder comparison and searchlight analysis results. (a) 10 fold cross validation test accuracy of insula decoder decreased on Day 2 (mean ± SEM across N=19 individuals). (b) Number of weighted voxels increased on Day 2. (c) Comparison between the mean searchlight accuracy of pgACC and insula clusters (masks extracted from clusters shown in figures below, at p<0.005 unc.) and right SII (mask from accuracy>75% on Day 1, k=50 voxels, see text). (d-e) Whole-brain searchlight analysis showed that information content contributing to decoding accuracy decreased in left insula and DLPFC/MFG, and increased in pgACC, on Day 2 comparing to Day 1 (shown at p<0.005, k>0).

6.1, Day 1: 24.1±1.1 voxels, Day 2: 28.7±0.7 voxels, signed-rank test Z(18)=-3.21, p=0.001) (NB. the sparse logistic regression method prunes unnecessary features using regularisation that assumes the weight prior as a zero mean vector and a diagonal covariance matrix, Yamashita et al., 2008). There were no significant differences in the locations (i.e. average of x, y, or z coordinates) of these weighted voxels within individuals. These results suggest the insula as an ROI may have overall disrupted functional information content for pain level encoding on Day 2 (neurofeedback).

To identify any changes in pain intensity representations across the whole of the brain, we conducted a whole brain post-hoc searchlight analysis using data from Day 1 and 2. This identifies accuracy maps that reflect the local information content of each voxel (Hebart et al., 2015; Kriegeskorte et al., 2006), which can be used to search the brain for changes in pain level representation. A paired t-test of these maps (Day 2 - Day 1 in a second level paired t-test,

DF=18) revealed reduced pain level decoding accuracy localised to a region in the left insula (Fig 6.4d, [-45, 6, 2], T=-6.04, k=142, whole brain cluster level p(FWE-corr)=0.014, shown at p<0.005 uncorrected). This localisation presumably underlies the ROI-based reduction in accuracy in the preceding analysis. Extracting the exact values from accuracy maps from both days, the left insula showed decreased decoding accuracy on Day 2 (171 voxels, Day 1: 67.844±2.320, Day 2: 57.546±2.366, paired t-test T(18)=-5.335, p=4.525e-5). Outside of the insula, we also noted reduced accuracy (135 voxels, Day 1: 67.074±1.715, Day 2: 55.932±2.234, paired t-test T(18)=-4.996, p=9.359e-5) in the left medial frontal gyrus (i.e dorsolateral prefrontal cortex (DLPFC), [-38, 9, 42], T=-5.68, k=134, which survived correction for whole brain multiple comparisons, whole brain cluster level p(FWE-corr)=0.045).

In contrast, using the same searchlight procedure as above, we found *increased* information content in a small region of the medial prefrontal cortex consistent with the pregenual anterior cingulate cortex (pgACC) (Fig 6.4e, [6, 44, 14], T=3.50, k=5, small volume correction (SVC) using an 8-mm spherical mask based on our previous investigation, Zhang et al., 2018a). Extracting the exact values from the accuracy maps from both days, this pgACC ROI had significantly increased decoding accuracy across all participants (Fig 6.4c, Day 1 accuracy: 55.293±1.604, Day 2: 63.009±2.383, paired t-test T(18)=3.676, p=0.0017).

For comparison, we also looked at the average accuracy maps from all participants in the right secondary somatosensory cortex (SII), since this region is also known to encode intensity-related pain responses. We found reasonable accuracies across both days (Day 1 peak [45,-17,26], accuracy=77.414, Day 2 peak [55,-30,26], accuracy=74.510), however, SII decoding accuracy did not vary significantly across days (averaged within the cluster mask from Day 1, k=50, Day 1: 75.813±2.234, Day 2: 71.563±2.880, paired t-test T(18)=1.344, p=0.196). To look at regional differences formally, we computed a day × location interaction: we found that this was significant between pgACC and SII (F(1,18)=6.648, p=0.012), although not between insula and SII (F(1,18)=1.507, p=0.22).

Trial sequence comparison

The imaging results above indicate that there are differences in the way pain is processed and represented on Day 2. To examine this further, we next looked for behavioural and brain evidence that might reflect the subject's reported attempts to modulate their brain activity in the context of the closed-loop neurofeedback on Day 2. Specifically, we classified 'switch' and 'non-switch' trials, depending on whether the pain level delivered in the current trial differed from that of the previous trial, because these sequence information can guide participants as they adapt their mental strategies on Day 2.

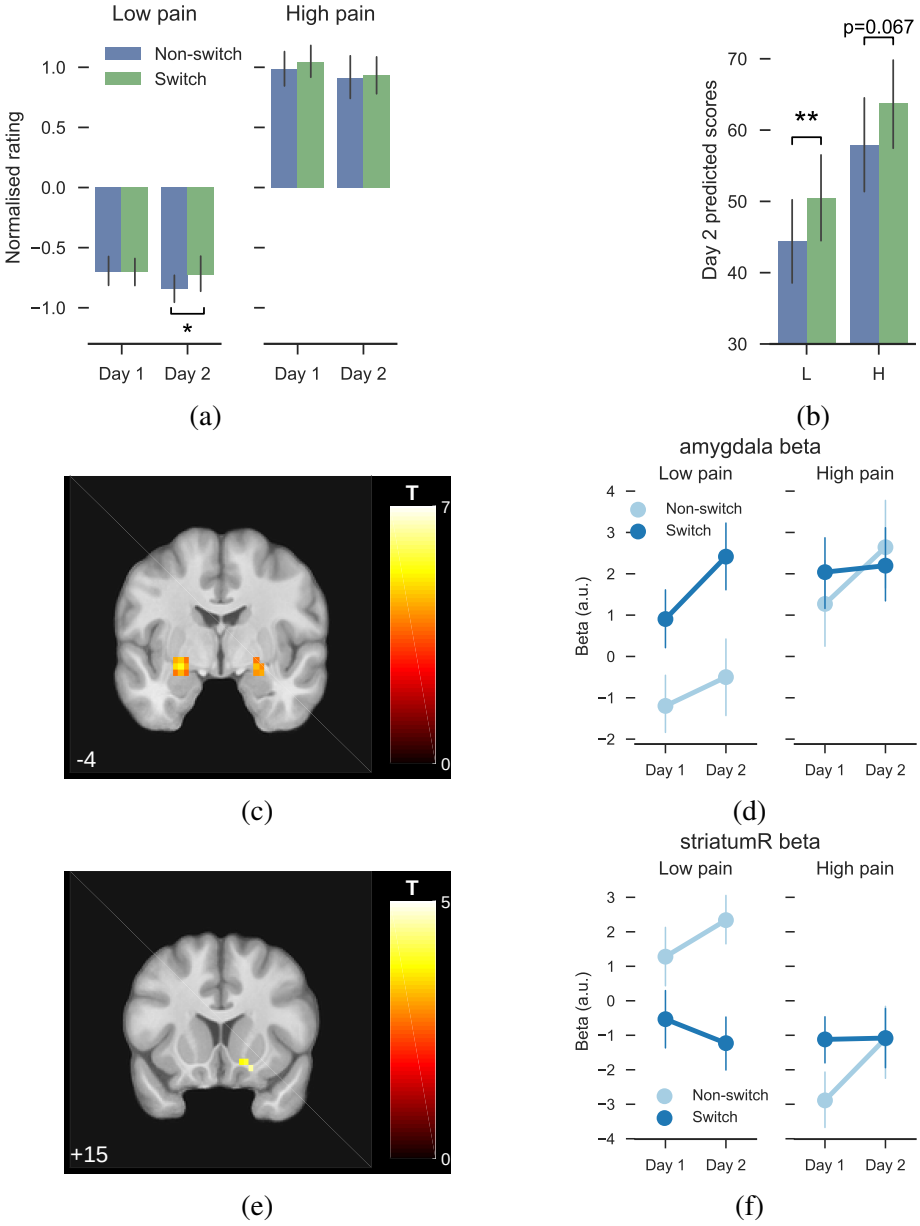


Fig. 6.5 Switch trials differences. All imaging results shown at p<0.001, k=0. (a) Within-subject normalised ratings, group by days, pain levels, and switch status, showing that Day 2 switch low pain trials were more painful than non-switch trials. (b) Day 2 decoder predicted scores (p(pain)*100) for switch/non-switch trials showed differences for high and low pain. (c) Day 2 HL>LL in bilateral amygdala. (d) Beta values extracted using bilateral activation cluster as ROI (at p<0.001 unc., k=30). (e) Day 2 LL>HL in right ventral striatum / OFC. (f) Beta values extracted from individuals using striatum activation cluster (k=9).

Using a simple t-contrast, we found that pain ratings for the low pain stimulus on switch trials was significantly higher than non-switch trials on Day 2, although there was no significant interaction across days (Fig. 6.5a, paired t-test T(18)=-2.466, p=0.0186, day × switch interaction for low pain: F(1,18)=2.477, p=0.133, interaction for high pain: F(1,18)=0.214, p=0.650). Similarly, we found that predicted insula P(Pain) score on switch trials was significantly higher for low pain (Fig 6.5b, paired t-test T(18)=2.990, p=0.0079), as well as being marginally so for high pain (T(18)=1.952, p=0.067). This provides evidence, from both ratings and insula patterns, that subjects were sensitive to the sequence of pain stimuli on Day 2.

In the imaging data, the contrast of switch vs non-switch low pain trials on Day 2 revealed HL>LL BOLD responses in bilateral amygdala (Fig 6.5c), and LL>HL in right striatum (Fig 6.5e). Exploring this result with an ROI analysis, Day 2 showed significant pain level × switch interaction in both ventral striatum (Fig 6.5f, repeated measure ANOVA F(1,18)=7.673, p=0.0126)), and amygdala (Fig 6.5d, ANOVA F(1,18)=11.991, p=0.0028)). This shows that learning-related brain regions track pain feedback.

Frequency learning evidence

The switch trial analyses provide behavioural and brain evidence that subjects are sensitive to the sequential identity of the stimuli on Day 2 (although this effect is not readily apparent on Day 1, the lack of an interaction by day means we can't necessarily conclude that subjects are *significantly* more sensitive to switches on Day 2). Switch trials are important because they contain more information than non-switch trials, an effect that can be formalised by a simple model in which people use the underlying frequency of high and low pain to infer how successful, overall, the machine is at reducing pain, i.e the overall probability of receiving low or high pain on any trial.

To capture a basic frequency learning process we applied a simple Bayesian learning model to quantify two key metrics: the ongoing probability of low/high pain, and the level of surprise on each trial (entropy). Previous studies have shown that such simple models provide a good account of behavioural and brain measures of surprise in comparable statistical environments (Mars et al., 2008; Meyniel et al., 2016).

We first looked at whether these metrics correlated with behaviour. Using a linear regression model of pain ratings (see methods), we found no correlation with *a posteriori* probability of low pain (z-transformed correlation coefficients Day 2 vs 0: T(18)=-0.582, p=0.568, Day 1 vs 0: T(18)=0.233, p=0.819, paired t-test between days: T(18)=0.601, p=0.555). However, we found a strong correlation with entropy, which was specific to Day 2, compared to Day 1 (Fig 6.6a, z-transformed correlation coefficients Day 2 vs 0: T(18)=4.648, p=1.99e-4, Day 1 vs 0:

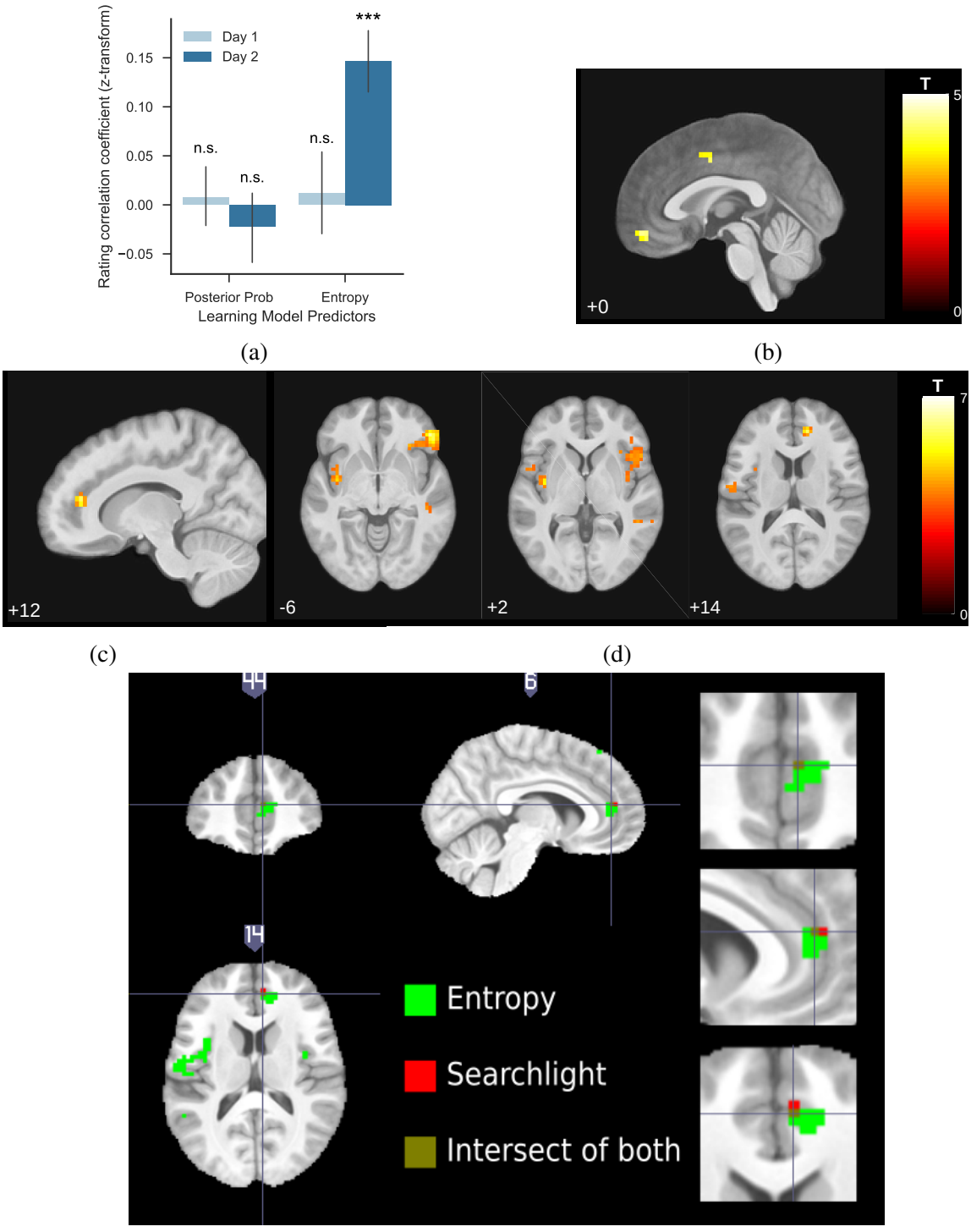


Fig. 6.6 Frequency learning evidence on Day 2. (a) Entropy (uncertainty regarding upcoming stimulus being high pain) from frequency learning model correlated with pain rating residuals from Day 2, but not Day 1 (using pain rating residuals with intensity and session numbers regressed out, see Methods). (b) Frequency learning model posterior probability of low pain correlated with VMPFC on Day 2 (peak coordinates [0, 51, -14], T=4.44, shown at p<0.001 unc.). (c-d) Frequency learning model entropy on Day 2 (i.e. entropy of posterior probability of current stimulus before updating) correlated with pgACC and bilateral insula (pgACC peak coordinates [13, 41, 14], T=5.91, sagittal and coronal views both at p<0.001 unc.). (e) Overlay of pgACC activations from both entropy (green) and searchlight (red) analysis (visualised at Z>3.2 for both).

T(18)=0.259, p=0.798, Paired t-test between days: T(18)=2.245, p=0.0376). That is, greater entropy was associated with greater subjective pain.

In the analysis of imaging data on Day 2, we found that the *a posteriori* probability of low pain was correlated with BOLD responses in VMPFC (Fig 6.6b), consistent with this regions strongly association with reward and relief value in previous studies of learning (Kim et al., 2006; Seymour et al., 2012). We found that entropy was correlated with BOLD responses in both bilateral insula and pgACC (Fig 6.6c and 6.6d). The insula response lay within the bilateral insula mask used for the decoder construction, and the pgACC response was part overlapping with the region associated with increased accuracy coding during adaptive neurofeedback (Fig 6.6e). When looking at the contrast of these responses across days, we found that although there was no significant effect of day in the insula response, the peak pgACC response was significantly greater on Day 2 (SVC corrected p(FWE-corr)=0.021, T=3.70, Z=3.15, peak coordinates [13,41,14]). That is, entropy – effectively an uncertainty metric during learning – correlated with both pain ratings and pgACC BOLD response specifically during neurofeedback.

6.4 Discussion

The results show that a real-time decoded brain response to an experimental pain stimulus can be used as feedback to teach a machine to reduce the intensity of future stimuli in an adaptive, closed-loop setting. Critically, subjects were not able to enhance discriminability in insula and somatosensory cortex, two regions commonly associated with intensity coding of pain, with the insula discriminability actually reduced. However, enhanced discriminability was observed in pgACC, reflecting the incentive to support the brain-machine system performance. Furthermore, the pgACC was also associated with modulation of subjective pain perception through uncertainty during frequency learning, a process that tracked the overall success of the closed-loop system. This indicates a central role for the pgACC in the dynamic control of neural representation and behavioural manifestation of pain. Across the whole brain, these results also reveal regionally-specific adaptive learning processes that can occur in a co-adaptive closed-loop brain-machine system for pain.

The enhancement of discriminability suggests that pain representation in the pgACC is under direct control. Along with other regions (including the insula and DLPFC), the pgACC has been implicated in endogenous control across a range of paradigms, including placebo/nocebo (Bingel et al., 2006; Eippert et al., 2009; Zubieta et al., 2005), attention/distraction (Bantick et al., 2002; Tracey et al., 2002; Valet et al., 2004), and controllability (Salomons et al., 2007; Zhang et al., 2018a). In at least some of these paradigms, it is sensitive to opioidergic

modulation – the neurotransmitter best associated with endogenous control – and is thought to mediate its effect through the descending pathway to the PAG and RVM (Valet et al., 2004; Zubieta et al., 2005). But it has remained unclear whether this modulation is a non-specific up- or down-regulation of pain responses, or a process in which the specific informational representation of pain is under control. Furthermore, the enhancement of discriminability cannot easily be explained by simple reinforcement (i.e. counter-conditioning of high pain to its subsequent relief, Koizumi et al., 2016), since it necessarily involves differential modulation of low and high intensity representations, e.g. reducing low intensity and enhancing high intensity representations.

The importance of the pgACC in endogenous pain control is further demonstrated by the fact that it was also associated with the behavioural modulation of pain (i.e. the subjective perception), through a learning process in which subjects appear to keep track of the underlying probability of low / high pain (which indicates the success of the overall system). In particular, the learned uncertainty (formally, the entropy) correlates with an increase in pain, in a manner at least partly independent to the modulation of the decoded intensity representation of pain. This increase in pain perception by learned uncertainty is consistent with the concept that pain is modulated according to its informational value to drive learning – when there is much to learn (i.e. high uncertainty), acute pain is increased to facilitate learning/attention (Yoshida et al., 2013) (with the opposite effect for tonic pain, whose role is suppress attention and cognition, Zhang et al., 2018a). Overall, therefore, the fact that pgACC activity independently reflects the control of both the neural and behavioural representations of pain suggests it has a sophisticated role in endogenous modulation. It is also distinct from the PAG, which although generally engaged during neurofeedback, did not show a parametric modulation according to discriminability or behavioural modulation.

The role of the pgACC contrasts with that of the decreased pain discriminability observed in the insula and DLPFC. The insula is strongly implicated in pain intensity coding and inference, it is shown to be involved in endogenous control through predictive coding, placebo analgesia and controllability (Bräscher et al., 2016; Geuter et al., 2017; Petrovic et al., 2002; Wager et al., 2004). The DLPFC is thought to have a modulatory role in endogenous control, where fMRI, structural and TMS evidence strongly associated the region with placebo analgesia (de Andrade et al., 2011; Krummenacher et al., 2010; Lui et al., 2010; Taylor et al., 2012; Wiech et al., 2006). Thus although pgACC, insula and DLPFC are co-implicated across a set of related behavioural contexts in which endogenous control is observed, our data shows that their function can be clearly dissociated. At least in the context of the current design, pgACC appears more closely tied to cognitive control and modulation of perceived pain.

The results demonstrate the feasibility of using online decoded pain responses to guide a closed-loop pain control system. The use of fMRI allows exploration of a range of target regions for decoding. Although multi-ROI classifiers can have much higher accuracy (Wager et al., 2013), here we used a single (bilateral) brain region as it is more realistic in terms of future applications that involve long-term implanted recordings, for example with electrocorticography (ECoG). However, this approach also has increased susceptibility to noise, warranting considerations for the trade-off between decoder simplicity and robustness. A key feature of our system is the incorporation of an RL decision function. This has a key advantage over fixed feedback decision policies because in principle RL algorithms can be used to search a much larger parameter space (as opposed to the binary levels of stimulation here) – something that has valuable applicability for many therapeutic interventions (e.g. spinal or brain stimulators). That is, when the optimal configuration of parameters for treatment under control is not known, the RL algorithm can search and find it over time. Combining the use of machine learning to generate an value approximation function with RL for optimal control provides an 'intelligent' system for pain control. However, larger parameter spaces require much longer time to learn, which moves it beyond the capability of fMRI as a measurement method. Although it is over such extended timescales that closed-loop systems are likely to have the highest clinical utility, it remains unclear how the brain may adapt to such systems in the long-term. Clinically, the combination of real-time signal decoding and RL control may speed up neurofeedback training, and making the procedure more intuitive for patients. While we showed that this co-adaptation engages the endogenous pain control system, its long-term impact and therapeutic value will need to be assessed in future studies.

The adaptive neurofeedback system revealed the presence of co-adaptive learning: whilst the machine is learning the value functions for different interventions, the brain is learning in parallel about the success of the machine in delivering the desired outcome. In our data, it appeared that some brain regions, i.e. the somatosensory cortex, were unaffected by learning, and together with a high decoding accuracy would make it a better candidate region than the insula – although this may change depending on the type of pain involved (thermal, as opposed to electrical for example). In the case of the insula, there are several mechanisms by which the representation may have been disrupted by learning, for instance it is well established that its pain responses change as a function of predictability (Geuter et al., 2017). In our study, identifying the mechanism of the difference in representation between Day 1 (when there was no task) and Day 2 (during neurofeedback) is complicated by an order effect – neurofeedback necessarily occurred after decoder construction. However, regardless of the mechanism, the lack of stability in decoder performance over time disrupts the machine's ability to learn the

value functions, creating a maladaptive feedback loop. This illustrates the importance of the specific choice of brain region for generating the feedback function.

6.5 Tables

Table 6.2 Multiple correction (cluster-forming threshold of p <0.001 uncorrected unless stated otherwise, regions from Harvard-Oxford, PAG probabilistic atlas, and previous study. *FWE cluster-level corrected. n=19. H: high pain, L: low pain.)

p*	k	t	Z	MNI coordinates (mm)			Region mask
				x	y	z	
Fig 6.3: Whole brain comparison (D2>D1, L>H, p<0.005)							
0.032	8	4.38	3.57	-16	-7	-18	Amygdala L
0.021	28	3.81	3.22	-22	9	6	Putamen L
		3.62	3.10	-19	6	2	
		3.62	3.09	-26	-1	2	
0.040	18	3.47	3	23	9	6	Putamen R
		3.29	2.88	26	15	-2	
Fig 6.3: Whole brain comparison (D2>D1, H>L, p<0.005)							
0.034	1	3.14	2.76	-3	-30	-6	dlPAG (Ezra et al., 2015)
0.034	1	2.94	2.62	3	-30	-6	
0.036	1	3.14	2.76	-3	-30	-6	lPAG (Ezra et al., 2015)
0.036	1	2.94	2.62	3	-30	-6	
Fig 6.4: Searchlight analysis - decreased information content (D2>D1)							
0.048	2	3.94	3.3	-42	3	-2	Insula L
0.061	2	4.41	3.59	-38	15	42	Middle Frontal Gyrus L
0.078	1	4.37	3.56	-38	35	30	
Fig 6.4: Searchlight analysis - increased information content (D2>D1, p<0.005)							
0.045	5	3.50	3.02	6	44	14	8mm pgACC sphere at [6,40,12] (Zhang et al., 2018a)
Fig 6.5: Whole brain comparison (D2, HL>LL)							
0.014	2	4.41	3.59	-26	-4	-14	Amygdala L
0.008	6	4.81	3.81	26	-7	-14	Amygdala R
Fig 6.5: Whole brain comparison (D2, LL >HL)							
striatum did not survive SVC							
Fig 6.6: Frequency learning model - posterior probability of low pain (D2)							
0.007	10	4.44	3.6	0	51	-14	Frontal Medial Cortex
Fig 6.6: Frequency learning model - entropy (D2)							
0.039	5	5.30	4.06	10	41	10	Cingulate Anterior
0.033	6	4.36	3.56	0	3	38	
0.002	14	5.91	4.35	13	41	14	8mm pgACC sphere at [6,40,12] (Zhang et al., 2018a)
0.002	31	5.24	4.03	-38	-7	2	Insular cortex (bilateral)
0.032	6	4.60	3.69	39	-4	6	

Chapter 7

Discussion

The aim of this thesis was to explore how pain and relief are represented in the human brain. Specifically, this posed three questions: (1) How do humans learn and adapt behaviourally and neurally in the presence of pain and relief, from a computational perspective? (2) How does learning influence the subjective experience of pain and relief, as well as their representations in the brain? (3) How does the understanding of pain/relief motivation inform the development of future technology-based pain therapeutics?

I investigated learning processes involved in pain and relief under three different sets of conditions: passive learning during Pavlovian conditioning with acute pain, active and passive relief learning during the escape of tonic pain, and human-machine co-adaptive learning in the presence of a closed-loop pain control system. Different types of data during these experiments were collected, including fMRI BOLD signals, physiological recordings, subjective ratings, and instrumental choices (Fig 2.1). Various formal learning models were then fitted to these data, allowing the use of model-generated predictors for computational characterisation of the underlying processes.

In this chapter, the contributions of these experimental studies to our current understanding of pain and relief is discussed, and their implications are summarised in section 7.1. Potential directions for future work are then given in section 7.2, including a discussion of the possibilities of extending current theoretical framework and improving pain therapeutics.

7.1 Contributions

Bridging the gap between conditioning paradigms and experimental data via the computational framework of learning, we demonstrated that pain and relief are not one-dimensional sensory experiences, but instead involve motivational aspects that transcend their various behavioural

and neurobiological measures. Learning to predict and control pain and relief in turn impacts on their perceptual experiences and information representations in the brain.

Uncertainty and pain/relief learning

We have demonstrated that uncertainty plays a crucial role in the control of learning during both acute and tonic pain. Computationally, the associability signal emerges as a key form of uncertainty used during pain/relief learning. It is calculated based on previously cached prediction error magnitudes arising during reinforcement learning. Associability keeps track of uncertainty by considering how surprising recent events were, i.e. how much actual outcomes differed from those that were predicted. It also acts to modulate learning rate for value updating, such that greater uncertainty leads to more rapid learning and hence faster convergence to stable predicted values (Pearce and Hall, 1980; AJ Yu and Dayan, 2005). Associability is also a biologically relevant signal: it can plausibly be implemented through activations of cerebellar climbing fibres (Schultz and Dickinson, 2000), and has previously been found to correlate with BOLD responses in the amygdala during aversive Pavlovian conditioning (Boll et al., 2013; Holland and Gallagher, 1993; Li et al., 2011).

The computational framework of associability allows us to identify the neural correlates of learning processes underlying different conditioned responses. Specifically, we found that trial-by-trial SCRs were best fitted by the state-based associability – both in Pavlovian learning of acute pain (Experiment 1), which has been well characterised in similar studies (Boll et al., 2013; Li et al., 2011); and in relief learning during tonic pain (Experiments 2 and 3), which has not been demonstrated previously. Active relief learning involves instrumental action learning, whereas Pavlovian conditioning involves only passive prediction learning. While action-based associability can be calculated, we did not find evidence of it being directly used in the learning of instrumental actions. This difference in the involvement of uncertainty for action and SCR learning is consistent with the fact that the instrumental learning process exists in parallel to the Pavlovian one. Our results were in accordance with existing evidence that draws a distinction between instrumental and Pavlovian learning systems (Gläscher et al., 2010; Li et al., 2011; G Morris et al., 2006), where the former learns actions, and the latter learns other conditioned (e.g. autonomic) responses, with the two systems interacting under appropriate circumstances (Holmes et al., 2010). Extending the active relief learning paradigm to a more complex environment, we demonstrated further that the parallel learning processes had a robust presence. The associability-SCR fitting results generalised in the relief learning case from a static learning environment (Experiment 2, where the probability of relief following an action does not change over time) to a dynamic one (Experiment 3, where a learning environment featuring more possible actions and changing relief probabilities over time). The latter kind of

context evokes higher variability in associability, ensuring the model fitting outcomes reflect learning effects instead of potentially confounding perceptual habituations. Together, these findings establish the relevance of associability beyond predictive learning of *acute* pain. As an uncertainty signal, it also has an important role in controlling learning of relief from *tonic* pain, where the restoration of homeostasis becomes the salient outcome.

We found that acute pain learning involves multiple dissociable neural processes. Using a lateralised experimental design, we were able to distinguish different types of behavioural conditioned responses ('preparatory' autonomic arousal, and 'consummatory' limb withdrawal) in human participants (Experiment 1). Importantly, this design could also distinguish the two underlying systems *computationally*: in the context of this experiment, the consummatory learning system is assumed to be sensitive to lateralised pain only, while the preparatory learning system attends to pain on both sides equally. This feature allowed us to model the two systems with two sets of independent predictors, which could be used to test for system-specific neural correlates in fMRI data. We showed that the cerebellum (lobules V and VI) tracked consummatory (laterality-specific) associability, while the amygdala tracked preparatory (non laterality-specific) associability during conditioning. These findings were in accordance with previous evidence regarding possible neural implementation of associability signals (Holland and Gallagher, 1993; Li et al., 2011; Schultz and Dickinson, 2000). Given this experiment prioritised neural process dissociation, we did not directly test for behavioural correlates of consummatory associability. However, this can be achieved with appropriate task designs that optimise consummatory behavioural recordings.

Our results provide the first computational account of the role of the cerebellum in pain, adding to existing evidence that aversive learning involves multiple processes outside of the amygdala (Balleine and Killcross, 2006; Cardinal et al., 2002; LeDoux, 2014). We linked the importance of the cerebellum in the acquisition of motor-related aversive conditioned responses (Daum et al., 1993; Lavond and Steinmetz, 1989) with its role in pain processing (Moulton et al., 2010; Seymour et al., 2004). Our task design satisfied the conditions for cerebellum recruitment in associative learning, including the elicitation of simple motor response, short CS-US interval (<3-4s), the use of aversive US, and the activation of the inferior olive by an intense US (Steinmetz, 2000). In addition, thermal pain stimuli have been shown to activate the cerebellum (Moulton et al., 2010; Wager et al., 2007). Together, these evidence suggest the identified cerebellar activations are specific to pain processing. In practical terms, these findings can serve as a reminder that observed pain behaviour should be used as an index of pain with caution. For example, evoked limb withdrawal in animal studies may be reflexive instead of reflecting 'affective' pain (Andersen et al., 2006). Indeed, as LeDoux (2017) recently pointed out, changing behavioural and physiological responses in animal research is not the

same as relieving fear or anxiety, and fixating on these behavioural indices could be detrimental to the development of translational treatments. Given that most preclinical pain studies rely heavily on non-verbal animal testing, interpretation of behavioural assessments should consider both evoked and spontaneous measures, as well as specific disease models, pharmacological, and genetic manipulations used (Burma et al., 2017).

In Experiments 2 and 3, we found that neural and behavioural responses underlying pain relief could be distinguished from those of reward learning. BOLD signals in the dorsal putamen was found to correlate with relief learning prediction errors from the action-learning TD model (Experiments 2 and 3), its activation patterns conformed to those required in an axiomatic test, which delineated these responses from simple reactions to relief outcomes (Roy et al., 2014). The dorsal putamen has previously been shown to be involved in controllability and habit learning (Maier and Seligman, 2016; Tricomi et al., 2009), and is anatomically distinct to the typical reward-related regions, notably the NAc, identified in previous pain relief studies (Andreatta et al., 2012; Leknes et al., 2011; Navratilova et al., 2012). Behaviourally, learned actions were better explained by a 'model-free' (or S-R habit) learning model, regardless of the learning environment being static or dynamic, as opposed to the 'model-based' (cognitive) learning strategies frequently identified in reward learning studies (Daw et al., 2005; Gläscher et al., 2010; Prévost et al., 2011). This difference is also evident from a motivational perspective – we approached restoring homeostasis as reinforcement, which is conceptually different from reward maximisation in typical reward reinforcement learning (learning 'what to do' can be independent from learning 'what not to do', Elfwing and Seymour, 2017). In addition, we have demonstrated that the associability signal correlated with both the changes in ongoing pain and the activity of the brain region heavily implicated in the endogenous pain control system, constituting a unique feature of pain relief learning (This will be discussed in more detail later in this section). In summary, these findings suggest that the processes underlying relief learning exhibit complex interactions with tonic pain, and warn against the simplistic view of equating pain relief with reward, or treating behavioural responses related to relief as indices of factually achieving pain relief.

The modulation of pain during learning

We have shown that uncertainty correlates with endogenously modulated pain, which suggests its potential capacity in maximising the impact of learning. In the case of tonic pain, rather than acting as a warning signal of imminent tissue damage, the prolonged presence of pain serves as a gauge of bodily homeostasis – it encourages recuperation by suppressing behaviour in order to conserve energy and limit further damage, and it resets pain relief as the new behavioural goal, where the immediate restoration of homeostasis is preferred (Crombez et al., 1997; Lorenz and

Bromm, 1997). This apparent paradox leads to an interesting question: if tonic pain generally suppresses behaviour and cognition, how does it balance the need to learn to achieve pain relief?

Our active relief learning studies showed through two independent fMRI experiments that the brain potentially has an ingenious solution: it selectively reduces pain when it detects learnable information about possible relief. We showed that the brain can achieve this by learning the uncertainty signal of associability, which measures the amount of learnable information about pain relief, and endogenously suppressing ongoing pain by a proportionate amount to aid relief-seeking. In this way, relief-seeking is facilitated by lowered levels of tonic pain when the environment is uncertain and learning is needed. Furthermore, the robustness of these findings was confirmed in different learning environments. This account therefore constitutes a generalised, mechanistic model of endogenous pain control during active relief learning, supported by behavioural, physiological and neurobiological data from healthy human participants.

We also showed evidence of perceived acute pain correlating with uncertainty (Experiment 4). In the absence of an explicit CS-US associative learning paradigm, the presence of a closed-loop system incentivised participants to learn, as altering mental strategies can influence the machine's decision on the pain intensity delivered in upcoming trials. In a within-subject yoked design, we compared responses to high/low pain sequences presented to participants with and without the closed-loop control, and showed behavioural and neural modulation of pain as a result of human learning. A Bayesian frequency learning model captured the learning process, such that the estimation of high/low pain probability and the level of uncertainty in the form of entropy were updated by accumulating stimuli appearances (i.e. tracking the overall success of the closed-loop system). We further identified a positive correlation between subjective pain ratings and uncertainty, which occurred only when closed-loop control was available. While entropy does not directly participate in the Bayesian updating of frequency learning, this subjective modulation of pain could dynamically enhance acute pain when uncertainty is high and learning is needed. Similar results has been shown in a cued acute pain study (Yoshida et al., 2013), suggesting the observed differential modulating responses to tonic and acute pain are unlikely to be driven solely by the absence of cued associations. Hence uncertainty can engage attention more effectively, leading to more efficient learning (AJ Yu and Dayan, 2005).

Based on these observations, we propose that uncertainty during learning can be a candidate computational construct to describe information-sensitive control of the endogenous analgesia system. This aspect of the endogenous system seeks to maximise the impact of learning by dynamically modulating pain according to current uncertainty, or the need for learning about the environment. Tracking the same uncertainty signal across contexts fits the parsimonious

assumption of biological systems, and in principle is able to accommodate different priorities of learning by prompting different responses downstream – suppressing tonic pain to facilitate relief-seeking, and enhancing acute pain to engage attention for effective learning of future pain predictions. Finally, this information-based pain control mechanism appears to be particularly relevant when active control (through either instrumental actions or mental strategies) is available, suggesting that the brain might selectively engage this endogenous control process only when effective learning is more likely to make an impact.

It should be noted that the proposed framework emerges from experiments designed to investigate pain/relief learning, however, learning is unlikely to be the sole contributing factor to the observed modulation of pain perception. For example, peripheral nociceptors alone have been shown to detect thermal pain and make magnitude judgements (Robinson et al., 1983), and pain relief induced by interupting a continuous noxious thermal stimulus with greater heat (offset analgesia) can activate endogenous pain control (Derbyshire and Osborn, 2009). In cases such as learned helplessness, learning may not have similar impacts on a maladapted brain. Hence the work in this thesis should be interpreted within the specific experimental context.

The role of pgACC in pain/relief learning

We have illustrated the central role of the pregenual anterior cingulate cortex (pgACC) in cortical pain control during pain and relief learning. In three different fMRI studies, we have localised uncertainty signals generated by learning models to the pgACC, in addition to showing these signals' correlation with subjective pain ratings. Specifically, these uncertainty signals were state-based associability from a hybrid TD model during active relief learning (Experiments 2 and 3), and entropy from a Bayesian frequency learning model of acute pain under closed-loop control (Experiment 4). It is worth noting that these uncertainty signals were not directly fitted to pain ratings – parameters of the learning model were fitted either using SCRs/choices or assumed to be fixed. In addition, we did not find pgACC responses using subjective pain ratings alone as a regressor in neuroimaging analyses. Therefore, BOLD responses identified in the pgACC with uncertainty signals were more likely to reflect learning processes than subjective pain experienced, despite the correlations between uncertainty and subjective pain.

The enhancement of pain discriminability suggests the pgACC may be a key node in the endogenous analgesia system responsible for active control (Experiment 4). The region has long been implicated in aspects of endogenous pain control related to attention and controllability (Bingel et al., 2006; Eippert et al., 2009; Salomons et al., 2007), especially opioid-dependent placebo/nocebo effects (Petrovic et al., 2002; Zubieta et al., 2005). However, it remains unclear

whether the pgACC is simply reflecting the consequences associated with pain responses, or actively representing pain information. Identified as the only region with increased discriminability of pain under closed-loop control, our results suggest that the pgACC could be reflecting information contents directly relating to endogenous control, since discriminability enhancement would require differential modulations of high/low pain intensity representations. In summary, our results indicate that the pgACC may be tracking uncertainty during learning, and it could be reflecting direct, active control of the endogenous analgesia system.

Clinically, the localisation of the pgACC might help elucidate the mechanism underlying previous ACC-targeted deep brain stimulation for intractable chronic pain treatment (Kringelbach et al., 2007; Mohseni et al., 2012), and potentially offer a more specific target for future stimulation-based therapeutics. However, we would caution against the simplistic view that this particular region might be specific to endogenous pain control, as previous evidence points to its relevance for computing uncertainty during general decision-making (Amemori and Graybiel, 2012), dealing with threat (Nitschke et al., 2006), processing emotions (Etkin et al., 2011), and encoding perseverance during foraging (Kolling et al., 2012; McGuire and Kable, 2015).

During our literature review, we noticed a discrepancy in terminology regarding the human pgACC. The ACC region similar to the one we have identified has been referred to as the rostral ACC (rACC) in many pain and placebo analgesia studies (Bingel et al., 2006; Eippert et al., 2009; Kong et al., 2006; Petrovic et al., 2002; Segerdahl et al., 2018; Zubieta et al., 2005), to distinguish it from the more caudal part of the ACC. However, many other studies, including ours, referred to the same region as the pregenual ACC (pgACC) (Atlas et al., 2010; Salomons et al., 2007; Shackman et al., 2011; Wager et al., 2007), following the anatomical definitions by Vogt (2009) to describe activations in the ACC subdivisions comprising Brodmann's areas 24 and 32. Some studies also used the term perigenual ACC (Bräscher et al., 2016; Valet et al., 2004). Anatomical evidence suggests that the pgACC can be distinguished from nearby regions such as the subgenual ACC (sgACC) and the anterior midcingulate cortex (aMCC) through differentiating distribution patterns of neurotransmitter receptors (Palomero-Gallagher et al., 2009), or tracing efferent/afferent projections (Vogt, 2009). Moreover, our results add support to the differentiation of the pgACC by providing further computational characterisation of the region. We believe future researchers would benefit from a clearer definition of terminology when pgACC-related results are being reported.

Implications for technology-based pain control systems

In our final experiment, we investigated the implications of learning for the design of future technology-based pain control systems. In a closed-loop system, subjective pain is decoded in real-time from neural activities, and immediate intervention is delivered according to the pain

experienced (Zhang et al., 2013; Zhang and Seymour, 2014). Developing a BMI for healthy participants, we explored how the brain might adapt to such a pain control system, where real-time decoded brain responses were used as feedback to teach a machine to reduce the intensity of future pain stimuli. The availability of closed-loop control can provide a context for learning – participants can either actively learn to increase the discriminability of pain in their brain to enhance machine performance, or passively associate pain with subsequent relief, reducing pain discriminability. If such learning occurs, as we have shown repeatedly in previous experiments, what does it entail for closed-loop therapeutic systems?

We showed that the availability of closed-loop control engaged learning and endogenous pain modulation, subsequently changing the neural representation of pain in the brain. The presence of the closed-loop system increased the decodable information contents of pain in the pgACC, while decreasing those in other pain-related brain regions including the insula and DLPFC. In addition, when closed-loop control was available, uncertainty measure (entropy from the Bayesian frequency learning model) appeared to modulate subjective pain, as well as being correlated with BOLD responses in the pgACC and insula. The identification of overlapping neural representation of these elements suggests that uncertainty-related pain modulation might play a role in changing the neural representation of pain. In practical terms, the engagement of endogenous control as a result of learning may contribute to the altered discriminability of pain in certain brain regions, reducing the decoding accuracy of pain decoders constructed for these brain regions (e.g. insula), and hence impacting the efficacy of the closed-loop system.

While learning might affect neural representation of pain, we showed that this can potentially be overcome by incorporating adaptive learning on the machine side of the system. We incorporated a basic reinforcement learning decision function which enabled the machine to adaptively learn the values of different pain states, instead of hardcoding fixed feedback decision policies. In our experiment, the machine was able to successfully adapt its pain state values to accommodate changing brain activities, subsequently delivering fewer high pain trials to participants. In theory, this type of human-machine co-adaptation could become the norm for future closed-loop control systems. However, given that fMRI is not suitable for recordings over extended timescales, the long-term consequences of these systems will require careful investigation with alternative BMI methods.

Methodological contributions

Methodologically, we illustrated the possibility of collecting data from different experiments – designed with different learning paradigms, conducted with different participants, and scanned in different scanners – to consolidate the same computational framework. The use

of open-source, version-controlled analysis software, with container software such as docker (https://www.docker.com/), a unified data structure protocol such as Brain Imaging Data Structure (BIDS, http://bids.neuroimaging.io/), and processing and analysis pipelines designed for the BIDS data structure (https://github.com/BIDS-Apps), ensures experimental data can be processed in a traceable, reproducible manner. This could be a potential solution to the current reproducibility problem within the neuroimaging community (Gorgolewski et al., 2017).

To our knowledge, the BMI system we presented in Experiment 4 represents the first attempt to develop rtfMRI-based, closed-loop control of pain in healthy human participants. This requires fewer training sessions compared to conventional decoded neurofeedback paradigms (Megumi et al., 2015; Shibata et al., 2011), and we have demonstrated its effectiveness in inducing human learning. Similar closed-loop systems could be adopted to test representation and discriminability changes in many other perceptual modalities for healthy human participants non-invasively, and to improve training efficiency of decoded neurofeedback in treatment settings.

Summary

We have demonstrated that pain and relief motivate learning, with uncertainty during learning playing an important role. In turn, learning to predict and control pain engages the endogenous pain control system, where uncertainty inversely correlates with pain perception, potentially to maximise the impact of learning. We have identified the pgACC to be a key brain region in the bidirectional relationship between perception and learning. Based on these observations, future pain therapeutics relying on decoded brain activities would need to take into account human learning and its engagement of endogenous pain control, potentially by employing adaptive control for the machine element. With the help of computational modelling, we can better understand the roles of the neural networks involved in the processing of pain/relief information during learning, beyond the conventional approach of stimulus-response brain mapping.

We have shown that the controllability of pain is inextricably linked to learning and endogenous modulation. From a learning perspective, controllability involves the detection and processing of learnable information regarding pain/relief occurrence in an environment, including the evaluation of uncertainty during learning. We propose that uncertainty concerning salient events, such as acute pain or relief from tonic pain, acts to direct attention to learnable associations between these events and the environment, and to engage the endogenous modulatory system to dynamically modulate pain perception in order to facilitate this learning. Depending on the circumstances, learning may have different priorities – obtaining relief when in tonic pain, or predicting the next occurrence of acute pain – which may prompt different responses of the endogenous pain control system. This may explain why previous studies have

inconsistently identified controllability-related pain suppression or enhancement (SM Miller, 1979; Salomons et al., 2007; Wiech et al., 2014b; Yoshida et al., 2013). In addition, the consistent localisation of uncertainty signals in the pgACC highlights the importance of this structure in the endogenous control of pain, an observation which is in accordance with numerous studies showing the prefrontal cortex's involvement in descending pain control (Bushnell et al., 2013; Wager and Atlas, 2015). We believe that a computational framework centred on uncertainty during learning can potentially unite pain attention, controllability, and endogenous analgesia, to explain the different modulatory effects attributed to the controllability of pain.

Our results also contribute to existing evidence of the pain/relief system's *degeneracy*, whereby structurally different brain regions and/or pathways yield the same behavioural output (Edelman and Gally, 2001; Tononi et al., 1999). Given their biological significance, it is unsurprising that pain and relief systems have evolved to engage multiple neural processes and behavioural responses. Unlike the redundancy built into many engineering systems, where the fail-safe mechanism is usually a standalone component, elements in a biological system not only coexist, but interact. Hence, with degeneracy comes adaptability, resilience, and complexity. These properties can manifest differently not only within, but also across individuals of a population (Noppeney et al., 2004; CJ Price and Friston, 2002). In-depth understanding of the various mechanisms underlying this complex system constitutes the first step towards understanding how they can go wrong, and the many consequences this may entail. By acknowledging the complex nature of pain and relief, we hope to provide realistic insights into pain-related disorders, and guidance for the future development of pain assessment, therapeutics, and research.

7.2 Future work

Extending theoretical framework

While the reinforcement learning framework is able to capture motivational states involved in pain and relief learning, there are other contributing components that have not been explored in the experimental designs in this thesis. Notably, working memory and effort valuation systems have both been shown to interact with reinforcement learning during decision-making (Collins and MJ Frank, 2012; Croxson et al., 2009). In real-life situations, the state-action space is much larger than that of laboratory settings, therefore learning is likely to involve a wider array of higher order cognitive functions. In the field of fear learning, a two-circuit view of threat processing has recently been proposed, whereby the subjective experience of fear emerges as a result of information integration from both cortical 'cognitive' (e.g. working memory)

and subcortical 'defensive survival' (e.g. behavioural) circuits (LeDoux, 2017; LeDoux and Brown, 2017). Experimental evidence from human electro-encephalography studies have also shown that working memory contributes to guiding reinforcement learning policy (Collins et al., 2017; Collins and M Frank, 2017). These recent developments have demonstrated that it is possible to expand existing experimental paradigms to test whether higher-order processes might be relevant for pain and relief motivation, and if this is the case, how the underlying computations might be organised in the brain. A possible approach is to adapt a pain or relief learning task with aspects from the decision-making task from Collins et al. (2017), which involves systematic manipulation of memory load and response delay. These manipulations could include stimuli with rotating changes in features such as shape and colour, and delayed trials to prevent a previously acquired action strategy being reused immediately. While this kind of design may require separate learning and testing phases in order to accommodate the acquisition of a larger state-action space, it would allow us to explore the influence of a working memory component that cannot be explored in our current designs. Interestingly, the DLPFC has been observed to play a role in working memory (Curtis and D'Esposito, 2003), opioid-dependent placebo effects (Zubieta et al., 2005), and encoding pain-related information (Experiment 4), suggesting a certain level of functional integration may be possible within this region.

Another extension to the current theoretical framework is to explore the potential role of effort valuation, specifically the costs of available actions, in pain and relief motivation. We have shown that action is an integral part of pain and relief learning, which appears to correlate with neural activities in distinct motor-related brain regions (cerebellum, dorsal putamen). Furthermore, previous studies have demonstrated a connection between pain relief and motor cortex stimulation (Hosomi et al., 2013; Yanagisawa et al., 2016; Zhang and Seymour, 2014). However, current pain/relief learning studies, especially fMRI experiments, are largely constrained in terms of available physical movements – an important part of natural pain/relief-related behaviour. A recent study has suggested a possible method for investigating the role of effort evaluation in pain-related learning. Using a 3 degrees-of-freedom, force-controlled robotic arm to elicit realistic arm movements, Meulders et al. (2016) paired increased effort with less frequent pain in order to study naturalistic pain avoidance learning. They found that individuals reported more pain-related fear for lower effort levels and exerted more effort to avoid pain, compared to a yoked control group where no effort-pain contingency was established. It might therefore be worth incorporating action costs or efforts as additional experimental variables for pain/relief learning. This can be achieved within an MRI environment by pairing varying levels of grip force with different pain/relief stimuli to the hand (Kawato et al., 2003), however, more naturalistic avoidance/escape actions would require the use of non-

MR modalities that can tolerate greater movement. It would also be interesting to investigate how thermal or electrical pain might differ from qualitatively different musculoskeletal pain in terms of eliciting naturalistic behavioural responses, with the use of EMG measurements. Notably, a previous study found evidence of action evaluation in the striatum and the ACC (Kurniawan et al., 2011), which are also important ROIs in pain and relief studies.

We also believe that our current studies would benefit from future replication and improvement. The theoretical framework of learning uncertainty and its engagement of endogenous pain control established here can be tested in other appropriately designed pain/relief learning paradigms, to address some of the issues we discussed in each experimental chapter – for example, the distinction between reducible and irreducible uncertainty and its potential effects on pain modulation (Experiments 2 and 3), and the role of conscious awareness, of either contingency or controllability, in uncertainty-related pgACC responses (Experiments 2-4). Our learning tasks could usefully be repeated under opioid manipulation (e.g. naloxone treatment), to investigate its effects on uncertainty-related pain modulation and the associated pgACC BOLD responses. In addition, the long-term consequences of uncertainty-related endogenous pain control need to be explored, possibly with non-fMRI modalities. This can be achieved by conducting pain/relief learning tasks over a period of days or weeks, in order to ascertain whether learning effects might extend to clinically relevant timescales. Experiment 2-4 used the yoked control design for across-paradigm comparisons. Yoking has been suggested to have inferrential limitations as random variations can cause systematic bias (Church, 1989), however, it offered direct comparison between response-contingent and nonresponse-contingent events in the same experiment, saving time and resources. It is possible to enhance the utility of yoking by including unreinforced stimuli in tasks, improving reinforcement scheduling, and interpreting from the directions of behavioural differences instead of magnitudes (Church, 1964; Jean-Richard-Dit-Bressel et al., 2018). Another possible improvement on the current designs is to adopt power analysis prior to determining sample size. Here, we based our sample size estimation on previous studies. However, if time and scheduling did not constitute major issues, prior power calculation based on pilot or archived data should become a standard practice in the planning of future fMRI experiments, with the help of open-source tools such as Neuropower (http://neuropowertools.org/).

Improving pain therapeutics

We have identified the pgACC as a potential therapeutic target for chronic pain treatment, however, the appropriate method for stimulation or modulation remains to be explored. Based on previous anatomical and functional evidence (Valet et al., 2004; Vogt et al., 2005), we attempted to increase the connectivity between pgACC and PAG using non-invasive neuro-

feedback training as a potential method to enhance endogenous pain control in an initial pilot study. We have recently modified the connectivity-based decoded neurofeedback training paradigm developed by Megumi et al. (2015), providing positive feedback (monetary incentives) when above-baseline pgACC-PAG connectivity was detected, and negative feedback for below-baseline connectivity, similar to a trial-by-trial reinforcement learning experiment. Following previous decoding neurofeedback protocols, no pain or relief stimulation were used during training.

Pilot data from two healthy human participants (not included in this thesis) showed inconsistent training effects on pgACC-PAG connectivity. There are two potential problems with this particular design. Firstly, the outcome of such training on endogenous analgesia is difficult to evaluate using our existing pain/relief learning paradigms. Secondly, pgACC-PAG connectivity appeared be modulated by participants controlling their breathing, a much simpler alternative to controlling brain activities. Although the latter problem can potentially be addressed by adding filters to reduce physiological noise in fMRI images online, there is no obvious solution to the first problem for healthy human volunteers. While the theoretical underpinnings behind this approach are sound, a more comprehensive design for outcome evaluation is needed to assess treatment efficacy and to establish placebo control, followed by a sufficiently powered blinded trial. A recent patient study narrowed the pathological alteration in pgACC-PAG connectivity down to the more specific region of ventrolateral PAG (Segerdahl et al., 2018), suggesting that the current decoded neurofeedback training may benefit from a more refined anatomical localisation.

Our current learning tasks could be adapted to characterise pain/relief motivation in patient populations, such as those with chronic pain and related comorbidities, and potentially supplement current procedures of diagnostic assessment in order to assist personalisation of treatment planning. Reinforcement learning tasks and choice data modelling have already been used to characterise deficits in patients with psychiatric disorders such as obsessive-compulsive disorder (Gillan et al., 2016), and evaluate treatment effects in Parkinson's disease (Seymour et al., 2016). While we have demonstrated the importance of pain/relief motivation in healthy human participants, we have yet to use these paradigms with pain patients. Establishing the correlation between modelled motivational deficits and the clinical manifestations of chronic pain, for example by combining learning tasks with pain questionnaires in both pain patients from different cohorts and healthy controls, will be valuable for establishing the standards needed for behavioural assessment tasks.

Finally, we have established the need for developing assessments of and adjustments for long-term effects in technology-based pain therapeutics. The definition of such therapeutics is not necessarily restricted to closed-loop pain control systems – a wide range of software

and devices capable of capturing and influencing pain-related behaviour as well as cognitive states should also be taken into account. With the technology industry starting to dedicate its infrastructure and capital to healthcare (Crow and A Gray, 2018), technology-based therapeutic systems, especially those based on BMI, will inevitably garner more attention (Palmer, 2018). It is likely that in the near future, more signal detection, processing, and storage options will be made available, expanding the existing pain therapeutic and management toolkit. However, as we have shown, the human brain learns from and adapts directly to a BMI pain control systems. Consequently, alongside the development of technology-based healthcare systems, we should be mindful of the potential consequences of using these tools in the long term.

To conclude, future research could extend the existing theoretical framework of pain/relief motivation to higher-order cognitive systems, and make use of these findings to improve existing assessment and treatment options for pain disorders. While all neurobiological models of pain and relief remain speculative, the work presented in this thesis shows the potential power of combining computational modelling with experimental evidence to improve our existing knowledge. Crucially, further understanding of 'everyday' pain and relief may shed light on how the human brain might be guided to prevent, or to better cope with chronic pain, by way of offering technology-based therapeutic alternatives.

References

1. Almeida TF, Roizenblatt S, and Tufik S. Afferent Pain Pathways: A Neuroanatomical Review. *Brain Research* **1000**.1 (2004), 40–56. DOI: 10.1016/j.brainres.2003.10.073.
2. Amat J, Baratta MV, Paul E, Bland ST, Watkins LR, and Maier SF. Medial Prefrontal Cortex Determines How Stressor Controllability Affects Behavior and Dorsal Raphe Nucleus. *Nature Neuroscience* **8**.3 (2005), 365–371. DOI: 10.1038/nn1399.
3. Amemori Ki and Graybiel AM. Localized Microstimulation of Primate Pregenual Cingulate Cortex Induces Negative Decision-Making. *Nature Neuroscience* **15**.5 (2012), 776–785. DOI: 10.1038/nn.3088.
4. Andersen OK, Mørch CD, and Arendt-Nielsen L. Modulation of Heat Evoked Nociceptive Withdrawal Reflexes by Painful Intramuscular Conditioning Stimulation. *Experimental Brain Research* **174**.4 (2006), 775–780. DOI: 10.1007/s00221-006-0674-5.
5. Andreatta M, Fendt M, Mühlberger A, Wieser MJ, Imobersteg S, Yarali A, Gerber B, and Pauli P. Onset and Offset of Aversive Events Establish Distinct Memories Requiring Fear and Reward Networks. *Learning & Memory* **19**.11 (2012), 518–526. DOI: 10.1101/lm.026864.112.
6. Andreatta M, Mühlberger A, Glotzbach-Schoon E, and Pauli P. Pain Predictability Reverses Valence Ratings of a Relief-Associated Stimulus. *Frontiers in Systems Neuroscience* **7** (2013). DOI: 10.3389/fnsys.2013.00053.
7. Andrew R, Derry S, Taylor RS, Straube S, and Phillips CJ. The Costs and Consequences of Adequately Managed Chronic Non-Cancer Pain and Chronic Neuropathic Pain. *Pain Practice* **14**.1 (2014), 79–94. DOI: 10.1111/papr.12050.
8. Apergis-Schoute AM, Schiller D, LeDoux JE, and Phelps EA. Extinction Resistant Changes in the Human Auditory Association Cortex Following Threat Learning. *Neurobiology of Learning and Memory*. Extinction **113** (2014), 109–114. DOI: 10.1016/j.nlm.2014.01.016.
9. Apkarian AV. The Brain in Chronic Pain: Clinical Implications. *Pain Management* **1**.6 (2011), 577–586. DOI: 10.2217/pmt.11.53.
10. Apkarian AV, Baliki MN, and Geha PY. Towards a Theory of Chronic Pain. *Progress in Neurobiology* **87**.2 (2009), 81–97. DOI: 10.1016/j.pneurobio.2008.09.018.
11. Apkarian AV, Bushnell MC, Treede RD, and Zubieta JK. Human Brain Mechanisms of Pain Perception and Regulation in Health and Disease. *European Journal of Pain* **9**.4 (2005), 463–463. DOI: 10.1016/j.ejpain.2004.11.001.
12. Apkarian AV, Sosa Y, Krauss BR, Thomas PS, Fredrickson BE, Levy RE, Harden RN, and Chialvo DR. Chronic Pain Patients Are Impaired on an Emotional Decision-Making Task. *Pain* **108** (2004), 129–136. DOI: 10.1016/j.pain.2003.12.015.

13. Arthurs OJ and Boniface S. How Well Do We Understand the Neural Origins of the fMRI BOLD Signal? *Trends in Neurosciences* **25**.1 (2002), 27–31. DOI: 10.1016/S0166-2236(00)01995-0.

14. Atlas LY, Bolger N, Lindquist MA, and Wager TD. Brain Mediators of Predictive Cue Effects on Perceived Pain. *Journal of Neuroscience* **30**.39 (2010), 12964–12977. DOI: 10.1523/JNEUROSCI.0057-10.2010.

15. Atlas LY and Wager TD. How Expectations Shape Pain. *Neuroscience Letters*. Neuroimaging of Pain: Insights into normal and pathological pain mechanisms **520**.2 (2012), 140–148. DOI: 10.1016/j.neulet.2012.03.039.

16. Bach DR, Daunizeau J, Friston KJ, and Dolan RJ. Dynamic Causal Modelling of Anticipatory Skin Conductance Responses. *Biological Psychology* **85**.1 (2010), 163–170. DOI: 10.1016/j.biopsycho.2010.06.007.

17. Bach DR and Dolan RJ. Knowing How Much You Don't Know: A Neural Organization of Uncertainty Estimates. *Nature Reviews Neuroscience* **13**.8 (2012), 572–586. DOI: 10.1038/nrn3289.

18. Bach DR, Flandin G, Friston KJ, and Dolan RJ. Time-Series Analysis for Rapid Event-Related Skin Conductance Responses. *Journal of Neuroscience Methods* **184**.2 (2009), 224–234. DOI: 10.1016/j.jneumeth.2009.08.005.

19. Baliki MN, Geha PY, Apkarian AV, and Chialvo DR. Beyond Feeling: Chronic Pain Hurts the Brain, Disrupting the Default-Mode Network Dynamics. *Journal of Neuroscience* **28**.6 (2008), 1398–1403. DOI: 10.1523/JNEUROSCI.4123-07.2008.

20. Baliki MN, Geha PY, Fields HL, and Apkarian AV. Predicting Value of Pain and Analgesia: Nucleus Accumbens Response to Noxious Stimuli Changes in the Presence of Chronic Pain. *Neuron* **66**.1 (2010), 149–160. DOI: 10.1016/j.neuron.2010.03.002.

21. Baliki MN, Petre B, Torbey S, Herrmann KM, Huang L, Schnitzer TJ, Fields HL, and Apkarian AV. Corticostriatal Functional Connectivity Predicts Transition to Chronic Back Pain. *Nature Neuroscience* **15**.8 (2012), 1117–1119. DOI: 10.1038/nn.3153.

22. Balleine BW, Delgado MR, and Hikosaka O. The Role of the Dorsal Striatum in Reward and Decision-Making. *Journal of Neuroscience* **27**.31 (2007), 8161–8165. DOI: 10.1523/JNEUROSCI.1554-07.2007.

23. Balleine BW and Dickinson A. Goal-Directed Instrumental Action: Contingency and Incentive Learning and Their Cortical Substrates. *Neuropharmacology* **37**.4 (1998), 407–419.

24. Balleine BW and Killcross S. Parallel Incentive Processing: An Integrated View of Amygdala Function. *Trends in Neurosciences* **29**.5 (2006), 272–279. DOI: 10.1016/j.tins.2006.03.002.

25. Balleine BW and O'Doherty JP. Human and Rodent Homologies in Action Control: Corticostriatal Determinants of Goal-Directed and Habitual Action. *Neuropsychopharmacology* **35**.1 (2010), 48. DOI: 10.1038/npp.2009.131.

26. Bantick SJ, Wise RG, Ploghaus A, Clare S, Smith SM, and Tracey I. Imaging How Attention Modulates Pain in Humans Using Functional MRI. *Brain* **125**.2 (2002), 310–319. DOI: 10.1093/brain/awf022.

27. Basbaum AI and Fields HL. Endogenous Pain Control Systems: Brainstem Spinal Pathways and Endorphin Circuitry. *Annual Review of Neuroscience* **7**.1 (1984), 309–338.
28. Becerra L, Navratilova E, Porreca F, and Borsook D. Analogous Responses in the Nucleus Accumbens and Cingulate Cortex to Pain Onset (Aversion) and Offset (Relief) in Rats and Humans. *Journal of Neurophysiology* **110**.5 (2013), 1221–1226. DOI: 10.1152/jn.00284.2013.
29. Becerra L and Borsook D. Signal Valence in the Nucleus Accumbens to Pain Onset and Offset. *European Journal of Pain* **12**.7 (2008), 866–869. DOI: 10.1016/j.ejpain.2007.12.007.
30. Becker S, Gandhi W, Kwan S, Ahmed AK, and Schweinhardt P. Doubling Your Payoff: Winning Pain Relief Engages Endogenous Pain Inhibition. *eNeuro* **2**.4 (2015). DOI: 10.1523/ENEURO.0029-15.2015.
31. Behrens TEJ, Woolrich MW, Walton ME, and Rushworth MFS. Learning the Value of Information in an Uncertain World. *Nature Neuroscience* **10**.9 (2007), 1214–1221. DOI: 10.1038/nn1954.
32. Behzadi Y, Restom K, Liau J, and Liu TT. A Component Based Noise Correction Method (CompCor) for BOLD and Perfusion Based fMRI. *NeuroImage* **37**.1 (2007), 90–101. DOI: 10.1016/j.neuroimage.2007.04.042.
33. Bellman R. A Markovian Decision Process. *Journal of Mathematics and Mechanics* **6**.5 (1957), 679–684. DOI: http://www.jstor.org/stable/24900506.
34. Belova MA, Paton JJ, Morrison SE, and Salzman CD. Expectation Modulates Neural Responses to Pleasant and Aversive Stimuli in Primate Amygdala. *Neuron* **55**.6 (2007), 970–984. DOI: 10.1016/j.neuron.2007.08.004.
35. Benedetti F, Mayberg HS, Wager TD, Stohler CS, and Zubieta JK. Neurobiological Mechanisms of the Placebo Effect. *Journal of Neuroscience* **25**.45 (2005), 10390–10402. DOI: 10.1523/JNEUROSCI.3458-05.2005.
36. Bermúdez JL. *Cognitive Science: An Introduction to the Science of the Mind*. Cambridge University Press, 2014. 555 pp.
37. Berridge KC. From Prediction Error to Incentive Salience: Mesolimbic Computation of Reward Motivation. *European Journal of Neuroscience* **35**.7 (2012), 1124–1143. DOI: 10.1111/j.1460-9568.2012.07990.x.
38. Betts SL, Brandon SE, and Wagner AR. Dissociation of the Blocking of Conditioned Eyeblink and Conditioned Fear Following a Shift in US Locus. *Animal Learning & Behavior* **24**.4 (1996), 459–470. DOI: 10.3758/BF03199017.
39. Bingel U, Gläscher J, Weiller C, and Büchel C. Somatotopic Representation of Nociceptive Information in the Putamen: An Event-Related fMRI Study. *Cerebral Cortex* **14**.12 (2004), 1340–1345. DOI: 10.1093/cercor/bhh094.
40. Bingel U, Lorenz J, Schoell E, Weiller C, and Büchel C. Mechanisms of Placebo Analgesia: rACC Recruitment of a Subcortical Antinociceptive Network. *Pain* **120** (2006), 8–15. DOI: 10.1016/j.pain.2005.08.027.
41. Blair HT, Huynh VK, Vaz VT, Van J, Patel RR, Hiteshi AK, Lee JE, and Tarpley JW. Unilateral Storage of Fear Memories by the Amygdala. *Journal of Neuroscience* **25**.16 (2005), 4198–4205. DOI: 10.1523/JNEUROSCI.0674-05.2005.

42. Blanchard DC and Blanchard RJ. Ethoexperimental Approaches to the Biology of Emotion. *Annual Review of Psychology* **39**.1 (1988), 43–68.

43. Boll S, Gamer M, Gluth S, Finsterbusch J, and Büchel C. Separate Amygdala Subregions Signal Surprise and Predictiveness during Associative Fear Learning in Humans. *European Journal of Neuroscience* **37**.5 (2013), 758–767. DOI: 10.1111/ejn.12094.

44. Bouton ME. *Learning and Behavior: A Contemporary Synthesis.* Sinauer Associates, 2007. 400 pp.

45. Bradley MM. Natural Selective Attention: Orienting and Emotion. *Psychophysiology* **46**.1 (2009), 1–11. DOI: 10.1111/j.1469-8986.2008.00702.x.

46. Bräscher AK, Becker S, Hoeppli ME, and Schweinhardt P. Different Brain Circuitries Mediating Controllable and Uncontrollable Pain. *Journal of Neuroscience* **36**.18 (2016), 5013–5025. DOI: 10.1523/JNEUROSCI.1954-15.2016.

47. Breivik H, Collett B, Ventafridda V, Cohen R, and Gallacher D. Survey of Chronic Pain in Europe: Prevalence, Impact on Daily Life, and Treatment. *European Journal of Pain* **10**.4 (2006), 287–287. DOI: 10.1016/j.ejpain.2005.06.009.

48. Brodersen KH, Wiech K, Lomakina EI, Lin Cs, Buhmann JM, Bingel U, Ploner M, Stephan KE, and Tracey I. Decoding the Perception of Pain from fMRI Using Multivariate Pattern Analysis. *NeuroImage* **63**.3 (2012), 1162–1170. DOI: 10.1016/j.neuroimage.2012.08.035.

49. Buchanan SL, Thompson RH, Maxwell BL, and Powell DA. Efferent Connections of the Medial Prefrontal Cortex in the Rabbit. *Experimental Brain Research* **79**.2 (1994), 469–483. DOI: 10.1007/BF00229186.

50. Burma NE, Leduc-Pessah H, Fan CY, and Trang T. Animal Models of Chronic Pain: Advances and Challenges for Clinical Translation. *Journal of Neuroscience Research* **95**.6 (2017), 1242–1256. DOI: 10.1002/jnr.23768.

51. Bushnell MC, Čeko M, and Low LA. Cognitive and Emotional Control of Pain and Its Disruption in Chronic Pain. *Nature Reviews Neuroscience* **14**.7 (2013), 502–511. DOI: 10.1038/nrn3516.

52. Buxton RB. The Physics of Functional Magnetic Resonance Imaging (fMRI). *Reports on Progress in Physics* **76**.9 (2013), 096601. DOI: 10.1088/0034-4885/76/9/096601.

53. Byrne JH and Dafny N. Neuroscience Online: An Electronic Textbook for the Neurosciences. *Department of Neurobiology and Anatomy, The University of Texas Medical School at Houston* (1997).

54. Cardinal RN, Parkinson JA, Hall J, and Everitt BJ. Emotion and Motivation: The Role of the Amygdala, Ventral Striatum, and Prefrontal Cortex. *Neuroscience & Biobehavioral Reviews* **26**.3 (2002), 321–352. DOI: 10.1016/S0149-7634(02)00007-6.

55. Carver CS and Scheier MF. Origins and Functions of Positive and Negative Affect: A Control-Process View. *Psychological Review* **97**.1 (1990), 19–35. DOI: 10.1037/0033-295X.97.1.19.

56. Chakour MC, Gibson SJ, Bradbeer M, and Helme RD. The Effect of Age on Aδ- and C-Fibre Thermal Pain Perception. *Pain* **64**.1 (1996), 143–152. DOI: 10.1016/0304-3959(95)00102-6.

57. Chen ACN, Niddam DM, and Arendt-Nielsen L. Contact Heat Evoked Potentials as a Valid Means to Study Nociceptive Pathways in Human Subjects. *Neuroscience Letters* **316**.2 (2001), 79–82. DOI: 10.1016/S0304-3940(01)02374-6.

58. Church RM. Systematic Effect of Random Error in the Yoked Control Design. *Psychological Bulletin* **62**.2 (1964), 122–131. DOI: 10.1037/h0042733.

59. Church RM. The Yoked Control Design. *Aversion, avoidance, and anxiety: Perspectives on aversively motivated behavior* (1989), 403–415.

60. Claeskens G and Hjort NL. *Model Selection and Model Averaging*. Cambridge University Press, 2008. 332 pp.

61. Cohen JD, Daw N, Engelhardt B, Hasson U, Li K, Niv Y, Norman KA, Pillow J, Ramadge PJ, Turk-Browne NB, and Willke TL. Computational Approaches to fMRI Analysis. *Nature Neuroscience* **20**.3 (2017), 304–313. DOI: 10.1038/nn.4499.

62. Collins AGE, Albrecht MA, Waltz JA, Gold JM, and Frank MJ. Interactions Among Working Memory, Reinforcement Learning, and Effort in Value-Based Choice: A New Paradigm and Selective Deficits in Schizophrenia. *Biological Psychiatry*. Computational Psychiatry **82**.6 (2017), 431–439. DOI: 10.1016/j.biopsych.2017.05.017.

63. Collins AGE and Frank M. Within and Across-Trial Dynamics of Human EEG Reveal Cooperative Interplay between Reinforcement Learning and Working Memory. *bioRxiv* (2017), 184812. DOI: 10.1101/184812.

64. Collins AGE and Frank MJ. How Much of Reinforcement Learning Is Working Memory, Not Reinforcement Learning? A Behavioral, Computational, and Neurogenetic Analysis. *European Journal of Neuroscience* **35**.7 (2012), 1024–1035. DOI: 10.1111/j.1460-9568.2011.07980.x.

65. Colloca L, Sigaudo M, and Benedetti F. The Role of Learning in Nocebo and Placebo Effects. *Pain* **136**.1-2 (2008), 211–218. DOI: 10.1016/j.pain.2008.02.006.

66. Corr PJ. Approach and Avoidance Behaviour: Multiple Systems and Their Interactions. *Emotion Review* **5**.3 (2013), 285–290. DOI: 10.1177/1754073913477507.

67. Craig AD. How Do You Feel — Now? The Anterior Insula and Human Awareness. *Nature Reviews Neuroscience* **10**.1 (2009), 59–70. DOI: 10.1038/nrn2555.

68. Craig AD. How Do You Feel? Interoception: The Sense of the Physiological Condition of the Body. *Nature Reviews Neuroscience* **3**.8 (2002), 655–666. DOI: 10.1038/nrn894.

69. Craig KD. Social Communication Model of Pain. *Pain* **156**.7 (2015), 1198–1199. DOI: 10.1097/j.pain.0000000000000185.

70. Criswell E. *Cram's Introduction to Surface Electromyography*. Jones & Bartlett Learning, 2011. 436 pp.

71. Crombez G, Eccleston C, Baeyens F, and Eelen P. Habituation and the Interference of Pain with Task Performance. *Pain* **70**.2 (1997), 149–154. DOI: 10.1016/S0304-3959(96)03304-0.

72. Crow D and Gray A. Amazon Makes Its Move in Healthcare. *Financial Times* (2018).

73. Croxson PL, Walton ME, O'Reilly JX, Behrens TEJ, and Rushworth MFS. Effort-Based Cost–Benefit Valuation and the Human Brain. *Journal of Neuroscience* **29**.14 (2009), 4531–4541. DOI: 10.1523/JNEUROSCI.4515-08.2009.

74. Curtis CE and D'Esposito M. Persistent Activity in the Prefrontal Cortex during Working Memory. *Trends in Cognitive Sciences* **7**.9 (2003), 415–423. DOI: 10.1016/S1364-6613(03)00197-9.

75. Dale AM. Optimal Experimental Design for Event-Related fMRI. *Human Brain Mapping* **8** (1999), 109–114.

76. Daum I, Schugens MM, Ackermann H, Lutzenberger W, Dichgans J, and Birbaumer N. Classical Conditioning after Cerebellar Lesions in Humans. *Behavioral Neuroscience* **107**.5 (1993), 748–756. DOI: 10.1037/0735-7044.107.5.748.

77. Daunizeau J, Adam V, and Rigoux L. VBA: A Probabilistic Treatment of Nonlinear Models for Neurobiological and Behavioural Data. *PLoS Computational Biology* **10**.1 (2014), e1003441. DOI: 10.1371/journal.pcbi.1003441.

78. Davis M. The Role of the Amygdala in Fear and Anxiety. *Annual review of neuroscience* **15**.1 (1992), 353–375.

79. Daw ND. Trial-by-Trial Data Analysis Using Computational Models. *Decision making, affect, and learning: Attention and performance XXIII* **23** (2011), 3–38. DOI: 10.1093/acprof:oso/9780199600434.003.0001.

80. Daw ND, Gershman SJ, Seymour B, Dayan P, and Dolan RJ. Model-Based Influences on Humans' Choices and Striatal Prediction Errors. *Neuron* **69**.6 (2011), 1204–1215. DOI: 10.1016/j.neuron.2011.02.027.

81. Daw ND, Niv Y, and Dayan P. Uncertainty-Based Competition between Prefrontal and Dorsolateral Striatal Systems for Behavioral Control. *Nature Neuroscience* **8**.12 (2005), 1704. DOI: 10.1038/nn1560.

82. Daw ND and Doya K. The Computational Neurobiology of Learning and Reward. *Current Opinion in Neurobiology*. Cognitive neuroscience **16**.2 (2006), 199–204. DOI: 10.1016/j.conb.2006.03.006.

83. Daw ND, Kakade S, and Dayan P. Opponent Interactions between Serotonin and Dopamine. *Neural Networks* **15**.4–6 (2002), 603–616. DOI: 10.1016/S0893-6080(02)00052-7.

84. Dawson ME, Schell AM, and Filion DL. The Electrodermal System. *Handbook of Psychophysiology*. Ed. by Cacioppo JT, Tassinary LG, and Berntson G. Cambridge: Cambridge University Press, 2007, 159–181.

85. Dayan P and Abbott LF. *Theoretical Neuroscience*. MIT Press, 2001. 480 pp.

86. Dayan P and Balleine BW. Reward, Motivation, and Reinforcement Learning. *Neuron* **36**.2 (2002), 285–298. DOI: 10.1016/S0896-6273(02)00963-7.

87. Dayan P, Kakade S, and Montague PR. Learning and Selective Attention. *Nature Neuroscience* **3** (2000), 1218–1223. DOI: 10.1038/81504.

88. De Andrade DC, Mhalla A, Adam F, Texeira MJ, and Bouhassira D. Neuropharmacological Basis of rTMS-Induced Analgesia: The Role of Endogenous Opioids. *Pain* **152**.2 (2011), 320–326. DOI: 10.1016/j.pain.2010.10.032.

89. De Berker AO, Rutledge RB, Mathys C, Marshall L, Cross GF, Dolan RJ, and Bestmann S. Computations of Uncertainty Mediate Acute Stress Responses in Humans. *Nature Communications* **7** (2016), 10996. DOI: 10.1038/ncomms10996.

90. deCharms RC. Applications of Real-Time fMRI. *Nature Reviews Neuroscience* **9**.9 (2008), 720. DOI: 10.1038/nrn2414.

91. deCharms RC, Christoff K, Glover GH, Pauly JM, Whitfield S, and Gabrieli JDE. Learned Regulation of Spatially Localized Brain Activation Using Real-Time fMRI. *NeuroImage* **21**.1 (2004), 436–443. DOI: 10.1016/j.neuroimage.2003.08.041.

92. deCharms RC, Maeda F, Glover GH, Ludlow D, Pauly JM, Soneji D, Gabrieli JDE, and Mackey SC. Control over Brain Activation and Pain Learned by Using Real-Time Functional MRI. *Proceedings of the National Academy of Sciences* **102**.51 (2005), 18626–18631. DOI: 10.1073/pnas.0505210102.

93. Deisseroth K. Optogenetics. *Nature Methods* **8**.1 (2011), 26–29. DOI: 10.1038/nmeth.f.324.

94. Delgado MR, Jou RL, LeDoux JE, and Phelps EA. Avoiding Negative Outcomes: Tracking the Mechanisms of Avoidance Learning in Humans During Fear Conditioning. *Frontiers in Behavioral Neuroscience* **3** (2009). DOI: 10.3389/neuro.08.033.2009.

95. Derbyshire SWG and Osborn J. Offset Analgesia Is Mediated by Activation in the Region of the Periaqueductal Grey and Rostral Ventromedial Medulla. *NeuroImage*. Brain Body Medicine **47**.3 (2009), 1002–1006. DOI: 10.1016/j.neuroimage.2009.04.032.

96. Derbyshire SWG, Jones AKP, Gyulai F, Clark S, Townsend D, and Firestone LL. Pain Processing during Three Levels of Noxious Stimulation Produces Differential Patterns of Central Activity. *Pain* **73**.3 (1997), 431–445. DOI: 10.1016/S0304-3959(97)00138-3.

97. Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, Albert MS, and Killiany RJ. An Automated Labeling System for Subdividing the Human Cerebral Cortex on MRI Scans into Gyral Based Regions of Interest. *NeuroImage* **31**.3 (2006), 968–980. DOI: 10.1016/j.neuroimage.2006.01.021.

98. Deutsch R, Smith KJM, Kordts-Freudinger R, and Reichardt R. How Absent Negativity Relates to Affect and Motivation: An Integrative Relief Model. *Emotion Science* **6:** (2015), 152. DOI: 10.3389/fpsyg.2015.00152.

99. Dickinson A, Smith J, and Mirenowicz J. Dissociation of Pavlovian and Instrumental Incentive Learning under Dopamine Antagonists. *Behavioral Neuroscience* **114**.3 (2000), 468–483.

100. Dickinson A and Balleine BW. The Role of Learning in the Operation of Motivational Systems. *Stevens' Handbook of Experimental Psychology*. John Wiley & Sons, Inc., 2002.

101. Diedrichsen J. A Spatially Unbiased Atlas Template of the Human Cerebellum. *NeuroImage* **33**.1 (2006), 127–138. DOI: 10.1016/j.neuroimage.2006.05.056.

102. Diedrichsen J, Balsters JH, Flavell J, Cussans E, and Ramnani N. A Probabilistic MR Atlas of the Human Cerebellum. *NeuroImage* **46**.1 (2009), 39–46. DOI: 10.1016/j.neuroimage.2009.01.045.

103. Dimitrova A, Kolb FP, Elles HG, Maschke M, Forsting M, Diener HC, and Timmann D. Cerebellar Responses Evoked by Nociceptive Leg Withdrawal Reflex as Revealed by Event-Related fMRI. *Journal of Neurophysiology* **90**.3 (2003), 1877–1886. DOI: 10.1152/jn.00053.2003.

104. Domesick VB. Projections from the Cingulate Cortex in the Rat. *Brain Research* **12**.2 (1969), 296–320. DOI: 10.1016/0006-8993(69)90002-X.

105. Edelman GM and Gally JA. Degeneracy and Complexity in Biological Systems. *Proceedings of the National Academy of Sciences* **98**.24 (2001), 13763–13768. DOI: 10.1073/pnas.231499798.

106. Eippert F, Bingel U, Schoell ED, Yacubian J, Klinger R, Lorenz J, and Büchel C. Activation of the Opioidergic Descending Pain Control System Underlies Placebo Analgesia. *Neuron* **63**.4 (2009), 533–543. DOI: 10.1016/j.neuron.2009.07.014.

107. Elfwing S and Seymour B. Parallel Reward and Punishment Control in Humans and Robots: Safe Reinforcement Learning Using the MaxPain Algorithm. IEEE International Conference on Development and Learning and on Epigenetic Robotics. 2017. DOI: 10.1109/DEVLRN.2017.8329799.

108. Emmert K, Kopel R, Sulzer J, Brühl AB, Berman BD, Linden DEJ, Horovitz SG, Breimhorst M, Caria A, Frank S, Johnston S, Long Z, Paret C, Robineau F, Veit R, Bartsch A, Beckmann CF, Van De Ville D, and Haller S. Meta-Analysis of Real-Time fMRI Neurofeedback Studies Using Individual Participant Data: How Is Brain Regulation Mediated? *NeuroImage* **124**.A (2016), 806–812. DOI: 10.1016/j.neuroimage.2015.09.042.

109. Esteban O, Blair R, Markiewicz CJ, Berleant SL, Moodie C, Ma F, Isik AI, Erramuzpe A, Kent JD, Goncalves M, Poldrack RA, and Gorgolewski KJ. *Poldracklab/Fmriprep: 1.0.0*. Zenodo, 2017. DOI: 10.5281/zenodo.1095198.

110. Etkin A, Egner T, and Kalisch R. Emotional Processing in Anterior Cingulate and Medial Prefrontal Cortex. *Trends in Cognitive Sciences* **15**.2 (2011), 85–93. DOI: 10.1016/j.tics.2010.11.004.

111. Everitt BJ and Robbins TW. Neural Systems of Reinforcement for Drug Addiction: From Actions to Habits to Compulsion. *Nature Neuroscience* **8**.11 (2005), 1481–1489. DOI: 10.1038/nn1579.

112. Ezra M, Faull OK, Jbabdi S, and Pattinson KT. Connectivity-Based Segmentation of the Periaqueductal Gray Matter in Human with Brainstem Optimized Diffusion MRI. *Human Brain Mapping* **36**.9 (2015), 3459–3471. DOI: 10.1002/hbm.22855.

113. Fernando ABP, Urcelay GP, Mar AC, Dickinson A, and Robbins TW. Comparison of the Conditioned Reinforcing Properties of a Safety Signal and Appetitive Stimulus: Effects of d-Amphetamine and Anxiolytics. *Psychopharmacology* **227**.2 (2013), 195–208. DOI: 10.1007/s00213-012-2952-1.

114. Fields HL. Pain Modulation: Expectation, Opioid Analgesia and Virtual Pain. *Progress in Brain Research* **122** (2000), 245–253.

115. Fields HL. State-Dependent Opioid Control of Pain. *Nature Reviews Neuroscience* **5**.7 (2004), 565–575. DOI: 10.1038/nrn1431.

116. FitzGerald TH, Friston KJ, and Dolan RJ. Action-Specific Value Signals in Reward-Related Regions of the Human Brain. *Journal of Neuroscience* **32**.46 (2012), 16417–1623a. DOI: 10.1523/JNEUROSCI.3254-12.2012.

117. Flor H, Knost B, and Birbaumer N. The Role of Operant Conditioning in Chronic Pain: An Experimental Investigation. *Pain* **95**.1 (2002), 111–118. DOI: 10.1016/S0304-3959(01)00385-2.

118. Friston KJ, Frith CD, Frackowiak RSJ, and Turner R. Characterizing Dynamic Brain Responses with fMRI: A Multivariate Approach. *NeuroImage* **2** (2, Part A 1995), 166–172. DOI: 10.1006/nimg.1995.1019.

119. Friston KJ, Jezzard P, and Turner R. Analysis of Functional MRI Time-Series. *Human Brain Mapping* **1**.2 (1994), 153–171. DOI: 10.1002/hbm.460010207.

120. Friston KJ, Stephan KE, Lund TE, Morcom A, and Kiebel S. Mixed-Effects and fMRI Studies. *NeuroImage* **24**.1 (2005), 244–252. DOI: 10.1016/j.neuroimage.2004.08.055.

121. Friston KJ. Functional and Effective Connectivity: A Review. *Brain Connectivity* **1**.1 (2011), 13–36. DOI: 10.1089/brain.2011.0008.

122. Friston KJ, Holmes AP, Worsley KJ, Poline JP, Frith CD, and Frackowiak RS. Statistical Parametric Maps in Functional Imaging: A General Linear Approach. *Human Brain Mapping* **2**.4 (1994), 189–210. DOI: 10.1002/hbm.460020402.

123. Fritz HC, McAuley JH, Wittfeld K, Hegenscheid K, Schmidt CO, Langner S, and Lotze M. Chronic Back Pain Is Associated With Decreased Prefrontal and Anterior Insular Gray Matter: Results From a Population-Based Cohort Study. *The Journal of Pain* **17**.1 (2016), 111–118. DOI: 10.1016/j.jpain.2015.10.003.

124. Ganesh G, Franklin DW, Gassert R, Imamizu H, and Kawato M. Accurate Real-Time Feedback of Surface EMG During fMRI. *Journal of Neurophysiology* **97**.1 (2007), 912–920. DOI: 10.1152/jn.00679.2006.

125. Garrison J, Erdeniz B, and Done J. Prediction Error in Reinforcement Learning: A Meta-Analysis of Neuroimaging Studies. *Neuroscience & Biobehavioral Reviews* **37**.7 (2013), 1297–1310. DOI: 10.1016/j.neubiorev.2013.03.023.

126. Geha P and Waxman SG. Pain Perception: Multiple Matrices or One? *JAMA Neurology* **73**.6 (2016), 628–630. DOI: 10.1001/jamaneurol.2016.0757.

127. Genud-Gabai R, Klavir O, and Paz R. Safety Signals in the Primate Amygdala. *Journal of Neuroscience* **33**.46 (2013), 17986–17994. DOI: 10.1523/JNEUROSCI.1539-13.2013.

128. Gerber B, Yarali A, Diegelmann S, Wotjak CT, Pauli P, and Fendt M. Pain-Relief Learning in Flies, Rats, and Man: Basic Research and Applied Perspectives. *Learning & Memory* **21**.4 (2014), 232–252. DOI: 10.1101/lm.032995.113.

129. Geuter S, Boll S, Eippert F, and Büchel C. Functional Dissociation of Stimulus Intensity Encoding and Predictive Coding of Pain in the Insula. *eLife* **6** (2017), e24770. DOI: 10.7554/eLife.24770.

130. Gillan CM, Kosinski M, Whelan R, Phelps EA, and Daw ND. Characterizing a Psychiatric Symptom Dimension Related to Deficits in Goal-Directed Control. *eLife* **5** (2016), e11305. DOI: 10.7554/eLife.11305.

131. Gläscher J, Daw N, Dayan P, and O'Doherty JP. States versus Rewards: Dissociable Neural Prediction Error Signals Underlying Model-Based and Model-Free Reinforcement Learning. *Neuron* **66**.4 (2010), 585–595. DOI: 10.1016/j.neuron.2010.04.016.

132. Goense J, Bohraus Y, and Logothetis NK. fMRI at High Spatial Resolution: Implications for BOLD-Models. *Frontiers in Computational Neuroscience* **10** (2016). DOI: 10.3389/fncom.2016.00066.

133. Gold MS and Gebhart GF. Nociceptor Sensitization in Pain Pathogenesis. *Nature Medicine* **16**.11 (2010), 1248–1257. DOI: 10.1038/nm.2235.

134. Gorgolewski KJ, Alfaro-Almagro F, Auer T, Bellec P, Capotă M, Chakravarty MM, Churchill NW, Cohen AL, Craddock RC, Devenyi GA, Eklund A, Esteban O, Flandin G, Ghosh SS, Guntupalli JS, Jenkinson M, Keshavan A, Kiar G, Liem F, Raamana PR, Raffelt D, Steele CJ, Quirion PO, Smith RE, Strother SC, Varoquaux G, Wang Y, Yarkoni T, and Poldrack RA. BIDS Apps: Improving Ease of Use, Accessibility, and Reproducibility of Neuroimaging Data Analysis Methods. *PLoS Computational Biology* **13**.3 (2017), e1005209. DOI: 10.1371/journal.pcbi.1005209.

135. Gray JA. *The Psychology of Fear and Stress*. CUP Archive, 1987.

136. Grondman I, Busoniu L, Lopes GA, and Babuska R. A Survey of Actor-Critic Reinforcement Learning: Standard and Natural Policy Gradients. *IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews)* **42**.6 (2012), 1291–1307. DOI: 10.1109/TSMCC.2012.2218595.

137. Grupe DW and Nitschke JB. Uncertainty and Anticipation in Anxiety: An Integrated Neurobiological and Psychological Perspective. *Nature Reviews Neuroscience* **14**.7 (2013), 488–501. DOI: 10.1038/nrn3524.

138. Hamm AO, Greenwald MK, Bradley MM, and Lang PJ. Emotional Learning, Hedonic Change, and the Startle Probe. *Journal of Abnormal Psychology* **102**.3 (1993), 453–465. DOI: 10.1037/0021-843X.102.3.453.

139. Hashmi JA, Baliki MN, Huang L, Baria AT, Torbey S, Hermann KM, Schnitzer TJ, and Apkarian AV. Shape Shifting Pain: Chronification of Back Pain Shifts Brain Representation from Nociceptive to Emotional Circuits. *Brain* **136**.9 (2013), 2751–2768. DOI: 10.1093/brain/awt211.

140. Haxby JV, Gobbini MI, Furey ML, Ishai A, Schouten JL, and Pietrini P. Distributed and Overlapping Representations of Faces and Objects in Ventral Temporal Cortex. *Science* **293**.5539 (2001), 2425–2430. DOI: 10.1126/science.1063736.

141. Hebart MN, Görgen K, and Haynes JD. The Decoding Toolbox (TDT): A Versatile Software Package for Multivariate Analyses of Functional Imaging Data. *Frontiers in Neuroinformatics* **8** (2015). DOI: 10.3389/fninf.2014.00088.

142. Higgins ET. Beyond Pleasure and Pain. *American Psychologist* **52**.12 (1997), 1280. DOI: 10.1037/0003-066X.52.12.1280.

143. Holden M. A Review of Geometric Transformations for Nonrigid Body Registration. *IEEE Transactions on Medical Imaging* **27**.1 (2008), 111–128. DOI: 10.1109/TMI.2007.904691.

144. Holland PC and Gallagher M. Amygdala Central Nucleus Lesions Disrupt Increments, but Not Decrements, in Conditioned Stimulus Processing. *Behavioral Neuroscience* **107**.2 (1993), 246–253. DOI: 10.1037/0735-7044.107.2.246.

145. Holland PC and Gallagher M. Different Roles for Amygdala Central Nucleus and Substantia Innominata in the Surprise-Induced Enhancement of Learning. *Journal of Neuroscience* **26**.14 (2006), 3791–3797. DOI: 10.1523/JNEUROSCI.0390-06.2006.

146. Holland PC and Schiffino FL. Mini-Review: Prediction Errors, Attention and Associative Learning. *Neurobiology of Learning and Memory* **131** (2016), 207–215. DOI: 10.1016/j.nlm.2016.02.014.

147. Holland PC, Bashaw M, and Quinn J. Amount of Training and Stimulus Salience Affect Associability Changes in Serial Conditioning. *Behavioural Processes* **59**.3 (2002), 169–183. DOI: 10.1016/S0376-6357(02)00092-X.

148. Holland PC and Gallagher M. Amygdala–Frontal Interactions and Reward Expectancy. *Current Opinion in Neurobiology* **14**.2 (2004), 148–155. DOI: 10.1016/j.conb.2004.03.007.

149. Holmes NM, Marchand AR, and Coutureau E. Pavlovian to Instrumental Transfer: A Neurobehavioural Perspective. *Neuroscience & Biobehavioral Reviews* **34**.8 (2010), 1277–1295. DOI: 10.1016/j.neubiorev.2010.03.007.

150. Hosomi K, Shimokawa T, Ikoma K, Nakamura Y, Sugiyama K, Ugawa Y, Uozumi T, Yamamoto T, and Saitoh Y. Daily Repetitive Transcranial Magnetic Stimulation of Primary Motor Cortex for Neuropathic Pain: A Randomized, Multicenter, Double-Blind, Crossover, Sham-Controlled Trial. *Pain* **154**.7 (2013), 1065–1072. DOI: 10.1016/j.pain.2013.03.016.

151. Huettel SA, Song AW, and McCarthy G. *Functional Magnetic Resonance Imaging*. Sinauer, 2014. 515 pp.

152. Huys QJM. Reinforcers and Control. Towards a Computational Ætiology of Depression. Gatsby Computational Neuroscience Unit, UCL, University of London, 2007.

153. Huys QJM, Cools R, Gölzer M, Friedel E, Heinz A, Dolan RJ, and Dayan P. Disentangling the Roles of Approach, Activation and Valence in Instrumental and Pavlovian Responding. *PLoS Computational Biology* **7**.4 (2011), e1002028. DOI: 10.1371/journal.pcbi.1002028.

154. IASP. Part III: Pain Terms, A Current List with Definitions and Notes on Usage. *Classification of Chronic Pain*. Ed. by Merskey H and Bogduk N. Second. Seattle: IASP (International Association for the Study of Pain) Press, 1994, 209–214.

155. Jean-Richard-Dit-Bressel P, Killcross S, and McNally GP. Behavioral and Neurobiological Mechanisms of Punishment: Implications for Psychiatric Disorders. *Neuropsychopharmacology* **43**.8 (2018), 1639–1650. DOI: 10.1038/s41386-018-0047-3.

156. Jensen MP, Karoly P, and Braver S. The Measurement of Clinical Pain Intensity: A Comparison of Six Methods. *Pain* **27**.1 (1986), 117–126. DOI: 10.1016/0304-3959(86)90228-9.

157. Jensen MP, Martin SA, and Cheung R. The Meaning of Pain Relief in a Clinical Trial. *The Journal of Pain* **6**.6 (2005), 400–406. DOI: 10.1016/j.jpain.2005.01.360.

158. Jepma M, Jones M, and Wager TD. The Dynamics of Pain: Evidence for Simultaneous Site-Specific Habituation and Site-Nonspecific Sensitization in Thermal Pain. *The Journal of Pain* **15**.7 (2014), 734–746. DOI: 10.1016/j.jpain.2014.02.010.

159. Johansen JP, Fields HL, and Manning BH. The Affective Component of Pain in Rodents: Direct Evidence for a Contribution of the Anterior Cingulate Cortex. *Proceedings of the National Academy of Sciences* **98**.14 (2001), 8077–8082. DOI: 10.1073/pnas.141218998.

160. Johansen JP, Tarpley JW, LeDoux JE, and Blair HT. Neural Substrates for Expectation-Modulated Fear Learning in the Amygdala and Periaqueductal Gray. *Nature Neuroscience* **13**.8 (2010), 979–986. DOI: 10.1038/nn.2594.

161. Jones AK, Watabe H, Cunningham VJ, and Jones T. Cerebral Decreases in Opioid Receptor Binding in Patients with Central Neuropathic Pain Measured by [11C]Diprenorphine Binding and PET. *European Journal of Pain* **8**.5 (2004), 479–485. DOI: 10.1016/j.ejpain.2003.11.017.

162. Kaulich T, Föhre W, Kutz DF, Gerwig M, Timmann D, and Kolb FP. Differences in Unconditioned and Conditioned Responses of the Human Withdrawal Reflex during Stance: Muscle Responses and Biomechanical Data. *Brain Research* **1326** (2010), 81–95. DOI: 10.1016/j.brainres.2010.02.051.

163. Kawato M, Kuroda T, Imamizu H, Nakano E, Miyauchi S, and Yoshioka T. Internal Forward Models in the Cerebellum: fMRI Study on Grip Force and Load Force Coupling. *Progress in Brain Research* **142** (2003), 171–188. DOI: 10.1016/S0079-6123(03)42013-X.

164. Kelly RM and Strick PL. Cerebellar Loops with Motor Cortex and Prefrontal Cortex of a Nonhuman Primate. *Journal of Neuroscience* **23**.23 (2003), 8432–8444.

165. Kim H, Shimojo S, and O'Doherty JP. Is Avoiding an Aversive Outcome Rewarding? Neural Substrates of Avoidance Learning in the Human Brain. *PLoS Biology* **4**.8 (2006), e233. DOI: 10.1371/journal.pbio.0040233.

166. Kobayashi K and Hsu M. Neural Mechanisms of Updating under Reducible and Irreducible Uncertainty. *Journal of Neuroscience* **37**.29 (2017), 6972–6982. DOI: 10.1523/JNEUROSCI.0535-17.2017.

167. Koizumi A, Amano K, Cortese A, Shibata K, Yoshida W, Seymour B, Kawato M, and Lau H. Fear Reduction without Fear through Reinforcement of Neural Activity That Bypasses Conscious Exposure. *Nature Human Behaviour* **1**.1 (2016), s41562-016-0006-016. DOI: 10.1038/s41562-016-0006.

168. Kolling N, Behrens TEJ, Mars RB, and Rushworth MFS. Neural Mechanisms of Foraging. *Science* **336**.6077 (2012), 95–98. DOI: 10.1126/science.1216930.

169. Kong J, Gollub RL, Rosman IS, Webb JM, Vangel MG, Kirsch I, and Kaptchuk TJ. Brain Activity Associated with Expectancy-Enhanced Placebo Analgesia as Measured by Functional Magnetic Resonance Imaging. *Journal of Neuroscience* **26**.2 (2006), 381–388. DOI: 10.1523/JNEUROSCI.3556-05.2006.

170. Konorski J. *Integrative Activity of the Brain: An Interdisciplinary Approach*. University of Chicago Press, 1967. 531 pp.

171. Kriegeskorte N, Goebel R, and Bandettini P. Information-Based Functional Brain Mapping. *Proceedings of the National Academy of Sciences* **103**.10 (2006), 3863–3868. DOI: 10.1073/pnas.0600244103.

172. Kringelbach ML. The Human Orbitofrontal Cortex: Linking Reward to Hedonic Experience. *Nature Reviews Neuroscience* **6**.9 (2005), 691–702. DOI: 10.1038/nrn1747.

173. Kringelbach ML, Jenkinson N, Owen SLF, and Aziz TZ. Translational Principles of Deep Brain Stimulation. *Nature Reviews Neuroscience* **8**.8 (2007), 623–635. DOI: 10.1038/nrn2196.

174. Krummenacher P, Candia V, Folkers G, Schedlowski M, and Schönbächler G. Prefrontal Cortex Modulates Placebo Analgesia. *Pain* **148**.3 (2010), 368–374. DOI: 10.1016/j.pain.2009.09.033.

175. Kurniawan IT, Guitart-Masip M, and Dolan RJ. Dopamine and Effort-Based Decision Making. *Frontiers in Neuroscience* **5** (2011). DOI: 10.3389/fnins.2011.00081.

176. Kwong KK, Belliveau JW, Chesler DA, Goldberg IE, Weisskoff RM, Poncelet BP, Kennedy DN, Hoppel BE, Cohen MS, and Turner R. Dynamic Magnetic Resonance Imaging of Human Brain Activity during Primary Sensory Stimulation. *Proceedings of the National Academy of Sciences* **89**.12 (1992), 5675–5679. DOI: 10.1073/pnas.89.12.5675.

177. Lang PJ, Bradley MM, and Cuthbert BN. Emotion, Attention, and the Startle Reflex. *Psychological Review* **97**.3 (1990), 377–395. DOI: 10.1037/0033-295X.97.3.377.

178. Lavond DG and Steinmetz JE. Acquisition of Classical Conditioning without Cerebellar Cortex. *Behavioural Brain Research* **33**.2 (1989), 113–164. DOI: 10.1016/S0166-4328(89)80047-6.

179. Lazarus RS. Cognition and Motivation in Emotion. *American Psychologist* **46**.4 (1991), 352–367. DOI: 10.1037/0003-066X.46.4.352.

180. Le Pelley ME. The Role of Associative History in Models of Associative Learning: A Selective Review and a Hybrid Model. *Quarterly Journal of Experimental Psychology Section B* **57**.3 (2004), 193–243. DOI: 10.1080/02724990344000141.

181. LeDoux JE. Coming to Terms with Fear. *Proceedings of the National Academy of Sciences* (2014), 201400335. DOI: 10.1073/pnas.1400335111.

182. LeDoux JE. Semantics, Surplus Meaning, and the Science of Fear. *Trends in Cognitive Sciences* **21**.5 (2017), 303–306. DOI: 10.1016/j.tics.2017.02.004.

183. LeDoux JE and Brown R. A Higher-Order Theory of Emotional Consciousness. *Proceedings of the National Academy of Sciences* **114**.10 (2017), E2016–E2025. DOI: 10.1073/pnas.1619316114.

184. Lee SW, Shimojo S, and O'Doherty JP. Neural Computations Underlying Arbitration between Model-Based and Model-Free Learning. *Neuron* **81**.3 (2014), 687–699. DOI: 10.1016/j.neuron.2013.11.028.

185. Leknes S, Berna C, Lee MC, Snyder GD, Biele G, and Tracey I. The Importance of Context: When Relative Relief Renders Pain Pleasant. *Pain* **154**.3 (2013), 402–410. DOI: 10.1016/j.pain.2012.11.018.

186. Leknes S, Brooks JCW, Wiech K, and Tracey I. Pain Relief as an Opponent Process: A Psychophysical Investigation. *European Journal of Neuroscience* **28**.4 (2008), 794–801. DOI: 10.1111/j.1460-9568.2008.06380.x.

187. Leknes S, Lee M, Berna C, Andersson J, and Tracey I. Relief as a Reward: Hedonic and Neural Responses to Safety from Pain. *PLoS ONE* **6**.4 (2011). DOI: 10.1371/journal.pone.0017870.

188. Leknes S and Tracey I. A Common Neurobiology for Pain and Pleasure. *Nature Reviews Neuroscience* **9**.4 (2008), 314–320. DOI: 10.1038/nrn2333.

189. Levine J, Gordon N, and Fields H. The Mechanism of Placebo Analgesia. *The Lancet* **312**.8091 (1978), 654–657. DOI: 10.1016/S0140-6736(78)92762-9.

190. Levita L, Hare TA, Voss HU, Glover G, Ballon DJ, and Casey BJ. The Bivalent Side of the Nucleus Accumbens. *NeuroImage* **44**.3 (2009), 1178–1187. DOI: 10.1016/j.neuroimage.2008.09.039.

191. Levita L, Hoskin R, and Champi S. Avoidance of Harm and Anxiety: A Role for the Nucleus Accumbens. *NeuroImage* **62**.1 (2012), 189–198. DOI: 10.1016/j.neuroimage.2012.04.059.

192. Li J, Schiller D, Schoenbaum G, Phelps EA, and Daw ND. Differential Roles of Human Striatum and Amygdala in Associative Learning. *Nature Neuroscience* **14**.10 (2011), 1250–1252. DOI: 10.1038/nn.2904.

193. Lieberman MD, Burns SM, Torre JB, and Eisenberger NI. Reply to Wager et Al.: Pain and the dACC: The Importance of Hit Rate-Adjusted Effects and Posterior Probabilities with Fair Priors. *Proceedings of the National Academy of Sciences* **113**.18 (2016), E2476–E2479. DOI: 10.1073/pnas.1603186113.

194. Lieberman MD and Eisenberger NI. The Dorsal Anterior Cingulate Cortex Is Selective for Pain: Results from Large-Scale Reverse Inference. *Proceedings of the National Academy of Sciences* **112**.49 (2015), 15250–15255. DOI: 10.1073/pnas.1515083112.

195. Linnman C, Moulton EA, Barmettler G, Becerra L, and Borsook D. Neuroimaging of the Periaqueductal Gray: State of the Field. *NeuroImage* **60**.1 (2012), 505–522. DOI: 10.1016/j.neuroimage.2011.11.095.

196. Littlewort GC, Bartlett MS, and Lee K. Automatic Coding of Facial Expressions Displayed during Posed and Genuine Pain. *Image and Vision Computing*. Visual and multimodal analysis of human spontaneous behaviour: **27**.12 (2009), 1797–1803. DOI: 10.1016/j.imavis.2008.12.010.

197. Locher C, Frey Nascimento A, Kirsch I, Kossowsky J, Meyer A, and Gaab J. Is the Rationale More Important than Deception? A Randomized Controlled Trial of Open-Label Placebo Analgesia. *Pain* **158**.12 (2017), 2320. DOI: 10.1097/j.pain.0000000000001012.

198. Logothetis NK, Pauls J, Augath M, Trinath T, and Oeltermann A. Neurophysiological Investigation of the Basis of the fMRI Signal. *Nature* **412**.6843 (2001), 150. DOI: 10.1038/35084005.

199. Lorenz J and Bromm B. Event-Related Potential Correlates of Interference between Cognitive Performance and Tonic Experimental Pain. *Psychophysiology* **34**.4 (1997), 436–445. DOI: 10.1111/j.1469-8986.1997.tb02387.x.

200. Lui F, Colloca L, Duzzi D, Anchisi D, Benedetti F, and Porro CA. Neural Bases of Conditioned Placebo Analgesia. *Pain* **151**.3 (2010), 816–824. DOI: 10.1016/j.pain.2010.09.021.

201. MacKay DJ. *Information Theory, Inference and Learning Algorithms*. Cambridge University Press, 2003. 640 pp.

202. Mackintosh NJ. *Conditioning and Associative Learning*. Clarendon Press Oxford, 1983. 322 pp.

203. Maia TV. Reinforcement Learning, Conditioning, and the Brain: Successes and Challenges. *Cognitive, Affective, & Behavioral Neuroscience* **9**.4 (2009), 343–364. DOI: 10.3758/CABN.9.4.343.

204. Maier SF and Seligman ME. Learned Helplessness: Theory and Evidence. *Journal of Experimental Psychology: General* **105**.1 (1976), 3. DOI: 10.1037/0096-3445.105.1.3.

205. Maier SF and Seligman ME. Learned Helplessness at Fifty: Insights from Neuroscience. *Psychological Review* **123**.4 (2016), 349. DOI: 10.1037/rev0000033.

206. Maier SF and Watkins LR. Stressor Controllability and Learned Helplessness: The Roles of the Dorsal Raphe Nucleus, Serotonin, and Corticotropin-Releasing Factor. *Neuroscience & Biobehavioral Reviews*. Animal Models of Depression and Antidepressant Activity **29** (2005), 829–841. DOI: 10.1016/j.neubiorev.2005.03.021.

207. Mallan KM and Lipp OV. Does Emotion Modulate the Blink Reflex in Human Conditioning? Startle Potentiation during Pleasant and Unpleasant Cues in the Picture–Picture Paradigm. *Psychophysiology* **44**.5 (2007), 737–748. DOI: 10.1111/j.1469-8986.2007.00541.x.

208. Marquand A, Howard M, Brammer M, Chu C, Coen S, and Mourão-Miranda J. Quantitative Prediction of Subjective Pain Intensity from Whole-Brain fMRI Data Using Gaussian Processes. *NeuroImage* **49**.3 (2010), 2178–2189. DOI: 10.1016/j.neuroimage.2009.10.072.

209. Mars RB, Debener S, Gladwin TE, Harrison LM, Haggard P, Rothwell JC, and Bestmann S. Trial-by-Trial Fluctuations in the Event-Related Electroencephalogram Reflect Dynamic Changes in the Degree of Surprise. *Journal of Neuroscience* **28**.47 (2008), 12539–12545. DOI: 10.1523/JNEUROSCI.2925-08.2008.

210. Mathys C, Daunizeau J, Friston KJ, and Stephan KE. A Bayesian Foundation for Individual Learning under Uncertainty. *Frontiers in Human Neuroscience* **5** (2011), 39. DOI: 10.3389/fnhum.2011.00039.

211. Matsumoto M and Hikosaka O. Lateral Habenula as a Source of Negative Reward Signals in Dopamine Neurons. *Nature* **447**.7148 (2007), 1111. DOI: 10.1038/nature05860.

212. McGuire JT and Kable JW. Medial Prefrontal Cortical Activity Reflects Dynamic Re-Evaluation during Voluntary Persistence. *Nature Neuroscience* **18**.5 (2015), 760–766. DOI: 10.1038/nn.3994.

213. McHugh SB, Barkus C, Huber A, Capitão L, Lima J, Lowry JP, and Bannerman DM. Aversive Prediction Error Signals in the Amygdala. *Journal of Neuroscience* **34**.27 (2014), 9024–9033. DOI: 10.1523/JNEUROSCI.4465-13.2014.

214. McNamee DC. Neural and Computational Representations of Decision Variables. California Institute of Technology, 2015.

215. Medoc. *PATHWAY Model CHEPS for Pain Assessment*. 2017.

216. Megumi F, Yamashita A, Kawato M, and Imamizu H. Functional MRI Neurofeedback Training on Connectivity between Two Regions Induces Long-Lasting Changes in Intrinsic Functional Network. *Frontiers in Human Neuroscience* **9** (2015). DOI: 10.3389/fnhum.2015.00160.

217. Meulders A, Franssen M, Fonteyne R, and Vlaeyen JWS. Acquisition and Extinction of Operant Pain-Related Avoidance Behavior Using a 3 Degrees-of-Freedom Robotic Arm. *Pain* (2016). DOI: 10.1097/j.pain.0000000000000483.

218. Meyniel F, Maheu M, and Dehaene S. Human Inferences about Sequences: A Minimal Transition Probability Model. *PLoS Computational Biology* **12**.12 (2016), e1005260. DOI: 10.1371/journal.pcbi.1005260.

219. Millan MJ. The Induction of Pain: An Integrative Review. *Progress in Neurobiology* **57**.1 (1999), 1–164. DOI: 10.1016/S0301-0082(98)00048-3.

220. Miller R and Escobar M. Laws and Models of Basic Conditioning. *Stevens' Handbook of Experimental Psychology*. John Wiley & Sons, Inc., 2004, 47.

221. Miller SM. Controllability and Human Stress: Method, Evidence and Theory. *Behaviour Research and Therapy* **17**.4 (1979), 287–304. DOI: 10.1016/0005-7967(79)90001-9.

222. Mohr C, Leyendecker S, Petersen D, and Helmchen C. Effects of Perceived and Exerted Pain Control on Neural Activity during Pain Relief in Experimental Heat Hyperalgesia: A fMRI Study. *European Journal of Pain* **16**.4 (2012), 496–508. DOI: 10.1016/j.ejpain.2011.07.010.

223. Mohseni HR, Smith PP, Parsons CE, Young KS, Hyam JA, Stein A, Stein JF, Green AL, Aziz TZ, and Kringelbach ML. MEG Can Map Short and Long-Term Changes in Brain Activity Following Deep Brain Stimulation for Chronic Pain. *PLoS ONE* **7**.6 (2012), e37993. DOI: 10.1371/journal.pone.0037993.

224. Montague PR and Berns GS. Neural Economics and the Biological Substrates of Valuation. *Neuron* **36**.2 (2002), 265–284. DOI: 10.1016/S0896-6273(02)00974-1.

225. Montgomery GH and Kirsch I. Classical Conditioning and the Placebo Effect. *Pain* **72**.1 (1997), 107–113. DOI: 10.1016/S0304-3959(97)00016-X.

226. Moran RJ, Kishida KT, Lohrenz T, Saez I, Laxton AW, Witcher MR, Tatter SB, Ellis TL, Phillips PE, Dayan P, and Montague PR. The Protective Action Encoding of Serotonin Transients in the Human Brain. *Neuropsychopharmacology* (2018). DOI: 10.1038/npp.2017.304.

227. Morris G, Nevet A, Arkadir D, Vaadia E, and Bergman H. Midbrain Dopamine Neurons Encode Decisions for Future Action. *Nature Neuroscience* **9**.8 (2006), 1057–1063. DOI: 10.1038/nn1743.

228. Morris JS and Dolan RJ. Dissociable Amygdala and Orbitofrontal Responses during Reversal Fear Conditioning. *NeuroImage* **22**.1 (2004), 372–380. DOI: 10.1016/j.neuroimage.2004.01.012.

229. Morville T, Friston K, Burdakov D, Siebner HR, and Hulme OJ. The Homeostatic Logic of Reward. *bioRxiv* (2018), 242974. DOI: 10.1101/242974.

230. Moulton EA, Schmahmann JD, Becerra L, and Borsook D. The Cerebellum and Pain: Passive Integrator or Active Participator? *Brain Research Reviews* **65**.1 (2010), 14–27. DOI: 10.1016/j.brainresrev.2010.05.005.

231. Mouraux A, Diukova A, Lee MC, Wise RG, and Iannetti GD. A Multisensory Investigation of the Functional Significance of the "Pain Matrix". *NeuroImage* **54**.3 (2011), 2237–2249. DOI: 10.1016/j.neuroimage.2010.09.084.

232. Mowrer OH. *Learning Theory and Behavior*. John Wiley & Sons, Inc., 1960. 555 pp.

233. Murray EA. The Amygdala, Reward and Emotion. *Trends in Cognitive Sciences* **11**.11 (2007), 489–497. DOI: 10.1016/j.tics.2007.08.013.

234. Myung IJ. Tutorial on Maximum Likelihood Estimation. *Journal of Mathematical Psychology* **47**.1 (2003), 90–100. DOI: 10.1016/S0022-2496(02)00028-7.

235. Navratilova E and Porreca F. Reward and Motivation in Pain and Pain Relief. *Nature Neuroscience* **17**.10 (2014), 1304–1312. DOI: 10.1038/nn.3811.

236. Navratilova E, Xie JY, Okun A, Qu C, Eyde N, Ci S, Ossipov MH, King T, Fields HL, and Porreca F. Pain Relief Produces Negative Reinforcement through Activation of Mesolimbic Reward–Valuation Circuitry. *Proceedings of the National Academy of Sciences* **109**.50 (2012), 20709–20713. DOI: 10.1073/pnas.1214605109.

237. Navratilova E, Xie JY, King T, and Porreca F. Evaluation of Reward from Pain Relief. *Annals of the New York Academy of Sciences* **1282**.1 (2013), 1–11. DOI: 10.1111/nyas.12095.

238. Navratilova E, Xie JY, Meske D, Qu C, Morimura K, Okun A, Arakawa N, Ossipov M, Fields HL, and Porreca F. Endogenous Opioid Activity in the Anterior Cingulate Cortex Is Required for Relief of Pain. *Journal of Neuroscience* **35**.18 (2015), 7264–7271. DOI: 10.1523/JNEUROSCI.3862-14.2015.

239. Nitschke JB, Sarinopoulos I, Mackiewicz KL, Schaefer HS, and Davidson RJ. Functional Neuroanatomy of Aversion and Its Anticipation. *NeuroImage* **29**.1 (2006), 106–116. DOI: 10.1016/j.neuroimage.2005.06.068.

240. Niv Y, Daw ND, and Dayan P. Choice Values. *Nature Neuroscience* **9**.8 (2006), 987–988. DOI: 10.1038/nn0806-987.

241. Niv Y and Schoenbaum G. Dialogues on Prediction Errors. *Trends in Cognitive Sciences* **12**.7 (2008), 265–272. DOI: 10.1016/j.tics.2008.03.006.

242. Noppeney U, Friston KJ, and Price CJ. Degenerate Neuronal Systems Sustaining Cognitive Functions. *Journal of Anatomy* **205**.6 (2004), 433–442. DOI: 10.1111/j.0021-8782.2004.00343.x.

243. O'Doherty JP, Dayan P, Schultz J, Deichmann R, Friston K, and Dolan RJ. Dissociable Roles of Ventral and Dorsal Striatum in Instrumental Conditioning. *Science* **304**.5669 (2004), 452–454. DOI: 10.1126/science.1094285.

244. O'Doherty JP, Hampton A, and Kim H. Model-Based fMRI and Its Application to Reward Learning and Decision Making. *Annals of the New York Academy of Sciences* **1104**.1 (2007), 35–53. DOI: 10.1196/annals.1390.022.

245. O'Reilly JX, Woolrich MW, Behrens TEJ, Smith SM, and Johansen-Berg H. Tools of the Trade: Psychophysiological Interactions and Functional Connectivity. *Social Cognitive and Affective Neuroscience* **7**.5 (2012), 604–609. DOI: 10.1093/scan/nss055.

246. O'Reilly RC and Munakata Y. *Computational Explorations in Cognitive Neuroscience: Understanding the Mind by Simulating the Brain*. MIT Press, 2000. 540 pp.

247. Ogawa S, Tank DW, Menon R, Ellermann JM, Kim SG, Merkle H, and Ugurbil K. Intrinsic Signal Changes Accompanying Sensory Stimulation: Functional Brain Mapping with Magnetic Resonance Imaging. *Proceedings of the National Academy of Sciences* **89**.13 (1992), 5951–5955. DOI: 10.1073/pnas.89.13.5951.

248. Ogawa S, Lee TM, Kay AR, and Tank DW. Brain Magnetic Resonance Imaging with Contrast Dependent on Blood Oxygenation. *Proceedings of the National Academy of Sciences* **87**.24 (1990), 9868–9872. DOI: 10.1073/pnas.87.24.9868.

249. Ossipov MH, Dussor GO, and Porreca F. Central Modulation of Pain. *The Journal of Clinical Investigation* **120**.11 (2010), 3779–3787. DOI: 10.1172/JCI43766.

250. Overmier JB and Seligman ME. Effects of Inescapable Shock upon Subsequent Escape and Avoidance Responding. *Journal of Comparative and Physiological Psychology* **63**.1 (1967), 28. DOI: 10.1037/h0024166.

251. Palmer A. Thought Experiments. *The Economist* (2018).

252. Palomero-Gallagher N, Vogt BA, Schleicher A, Mayberg HS, and Zilles K. Receptor Architecture of Human Cingulate Cortex: Evaluation of the Four-Region Neurobiological Model. *Human Brain Mapping* **30**.8 (2009), 2336–2355. DOI: 10.1002/hbm.20667.

253. Paton JJ, Belova MA, Morrison SE, and Salzman CD. The Primate Amygdala Represents the Positive and Negative Value of Visual Stimuli during Learning. *Nature* **439**.7078 (2006), 865–870. DOI: 10.1038/nature04490.

254. Pavlov IP. *Conditioned Reflexes*. Trans. by Anrep GV. Courier Corporation, 1927. 466 pp.

255. Pearce JM and Hall G. A Model for Pavlovian Learning: Variations in the Effectiveness of Conditioned but Not of Unconditioned Stimuli. *Psychological Review* **87**.6 (1980), 532. DOI: 10.1037/0033-295X.87.6.532.

256. Pearce JM, Montgomery A, and Dickinson A. Contralateral Transfer of Inhibitory and Excitatory Eyelid Conditioning in the Rabbit. *The Quarterly Journal of Experimental Psychology Section B* **33**.1 (1981), 45–61. DOI: 10.1080/14640748108400828.

257. Penny WD, Holmes AP, and Friston KJ. Hierarchical Models. *Human Brain Function (Second Edition)*. Burlington: Academic Press, 2004, 851–863. DOI: 10.1016/B978-012264841-0/50045-7.

258. Penny WD, Friston KJ, Ashburner JT, Kiebel SJ, and Nichols TE. *Statistical Parametric Mapping: The Analysis of Functional Brain Images*. Academic Press, 2011. 656 pp.

259. Peri T, Ben-Shakhar G, Orr SP, and Shalev AY. Psychophysiologic Assessment of Aversive Conditioning in Posttraumatic Stress Disorder. *Biological Psychiatry* **47**.6 (2000), 512–519. DOI: 10.1016/S0006-3223(99)00144-4.

260. Pernet CR. Misconceptions in the Use of the General Linear Model Applied to Functional MRI: A Tutorial for Junior Neuro-Imagers. *Frontiers in Neuroscience* **8** (2014). DOI: 10.3389/fnins.2014.00001.

261. Petrovic P, Kalso E, Petersson KM, and Ingvar M. Placebo and Opioid Analgesia–Imaging a Shared Neuronal Network. *Science* **295**.5560 (2002), 1737–1740. DOI: 10.1126/science.1067176.

262. Phelps EA and LeDoux JE. Contributions of the Amygdala to Emotion Processing: From Animal Models to Human Behavior. *Neuron* **48**.2 (2005), 175–187. DOI: 10.1016/j.neuron.2005.09.025.

263. Ploghaus A, Tracey I, Gati JS, Clare S, Menon RS, Matthews PM, and Rawlins JNP. Dissociating Pain from Its Anticipation in the Human Brain. *Science* **284**.5422 (1999), 1979–1981. DOI: 10.1126/science.284.5422.1979.

264. Ploner M, Lee MC, Wiech K, Bingel U, and Tracey I. Prestimulus Functional Connectivity Determines Pain Perception in Humans. *Proceedings of the National Academy of Sciences* **107**.1 (2010), 355–360. DOI: 10.1073/pnas.0906186106.

265. Poldrack RA, Mumford JA, and Nichols TE. *Handbook of Functional MRI Data Analysis*. Cambridge University Press, 2011. 238 pp.

266. Porreca F and Navratilova E. Reward, Motivation, and Emotion of Pain and Its Relief. *Pain* **158** (2017), S43–S49. DOI: 10.1097/j.pain.0000000000000798.

267. Prévost C, McCabe JA, Jessup RK, Bossaerts P, and O'Doherty JP. Differentiable Contributions of Human Amygdalar Subregions in the Computations Underlying Reward and Avoidance Learning. *European Journal of Neuroscience* **34**.1 (2011), 134–145. DOI: 10.1111/j.1460-9568.2011.07686.x.

268. Prevost C, McNamee D, Jessup RK, Bossaerts P, and O'Doherty JP. Evidence for Model-Based Computations in the Human Amygdala during Pavlovian Conditioning. *PLoS Computational Biology* **9**.2 (2013), e1002918. DOI: 10.1371/journal.pcbi.1002918.

269. Price CJ and Friston KJ. Degeneracy and Cognitive Anatomy. *Trends in Cognitive Sciences* **6**.10 (2002), 416–421. DOI: 10.1016/S1364-6613(02)01976-9.

270. Price DD, McGrath PA, Rafii A, and Buckingham B. The Validation of Visual Analogue Scales as Ratio Scale Measures for Chronic and Experimental Pain. *Pain* **17**.1 (1983), 45–56. DOI: 10.1016/0304-3959(83)90126-4.

271. Rescorla RA. Probability of Shock in the Presence and Absence of Cs in Fear Conditioning. *Journal of Comparative and Physiological Psychology* **66**.1 (1968), 1–5. DOI: 10.1037/h0025984.

272. Rescorla RA and Wagner AR. A Theory of Pavlovian Conditioning: Variations in the Effectiveness of Reinforcement and Nonreinforcement. *Classical conditioning II: Current research and theory* (1972), 64–99.

273. Rigoux L, Stephan KE, Friston KJ, and Daunizeau J. Bayesian Model Selection for Group Studies — Revisited. *NeuroImage* **84** (2014), 971–985. DOI: 10.1016/j.neuroimage.2013.08.065.

274. Robinson CJ, Erik Torebjörk H, and LaMotte RH. Psychophysical Detection and Pain Ratings of Incremental Thermal Stimuli: A Comparison with Nociceptor Responses in Humans. *Brain Research* **274**.1 (1983), 87–106. DOI: 10.1016/0006-8993(83)90523-1.

275. Roesch MR, Calu DJ, Esber GR, and Schoenbaum G. Neural Correlates of Variations in Event Processing during Learning in Basolateral Amygdala. *Journal of Neuroscience* **30**.7 (2010), 2464–2471. DOI: 10.1523/JNEUROSCI.5781-09.2010.

276. Roesch MR, Esber GR, Li J, Daw ND, and Schoenbaum G. Surprise! Neural Correlates of Pearce–Hall and Rescorla–Wagner Coexist within the Brain. *European Journal of Neuroscience* **35**.7 (2012), 1190–1200. DOI: 10.1111/j.1460-9568.2011.07986.x.

277. Roy M, Piché M, Chen JI, Peretz I, and Rainville P. Cerebral and Spinal Modulation of Pain by Emotions. *Proceedings of the National Academy of Sciences* **106**.49 (2009), 20900–20905. DOI: 10.1073/pnas.0904706106.

278. Roy M, Shohamy D, Daw N, Jepma M, Wimmer GE, and Wager TD. Representation of Aversive Prediction Errors in the Human Periaqueductal Gray. *Nature Neuroscience* **17**.11 (2014), 1607–1612. DOI: 10.1038/nn.3832.

279. Rubio A, Van Oudenhove L, Pellissier S, Ly HG, Dupont P, de Micheaux HL, Tack J, Dantzer C, Delon-Martin C, and Bonaz B. Uncertainty in Anticipation of Uncomfortable Rectal Distension Is Modulated by the Autonomic Nervous System — A fMRI Study in Healthy Volunteers. *NeuroImage* **107** (2015), 10–22. DOI: 10.1016/j.neuroimage.2014.11.043.

280. Rushworth MFS, Noonan MP, Boorman ED, Walton ME, and Behrens TE. Frontal Cortex and Reward-Guided Learning and Decision-Making. *Neuron* **70**.6 (2011), 1054–1069. DOI: 10.1016/j.neuron.2011.05.014.

281. Rutledge RB, Dean M, Caplin A, and Glimcher PW. Testing the Reward Prediction Error Hypothesis with an Axiomatic Model. *Journal of Neuroscience* **30**.40 (2010), 13525–13536. DOI: 10.1523/JNEUROSCI.1747-10.2010.

282. Saab CY and Willis WD. The Cerebellum: Organization, Functions and Its Role in Nociception. *Brain Research Reviews* **42**.1 (2003), 85–95. DOI: 10.1016/S0165-0173(03)00151-6.

283. Salamone JD. The Involvement of Nucleus Accumbens Dopamine in Appetitive and Aversive Motivation. *Behavioural Brain Research* **61**.2 (1994), 117–133. DOI: 10.1016/0166-4328(94)90153-8.

284. Salomons TV, Iannetti GD, Liang M, and Wood JN. The "Pain Matrix" in Pain-Free Individuals. *JAMA Neurology* **73**.6 (2016), 755–756. DOI: 10.1001/jamaneurol.2016.0653.

285. Salomons TV, Johnstone T, Backonja MM, and Davidson RJ. Perceived Controllability Modulates the Neural Response to Pain. *Journal of Neuroscience* **24**.32 (2004), 7199–7203. DOI: 10.1523/JNEUROSCI.1315-04.2004.

286. Salomons TV, Johnstone T, Backonja MM, Shackman AJ, and Davidson RJ. Individual Differences in the Effects of Perceived Controllability on Pain Perception: Critical Role of the Prefrontal Cortex. *Journal of Cognitive Neuroscience* **19**.6 (2007), 993–1003. DOI: 10.1162/jocn.2007.19.6.993.

287. Salomons TV, Nusslock R, Detloff A, Johnstone T, and Davidson RJ. Neural Emotion Regulation Circuitry Underlying Anxiolytic Effects of Perceived Control over Pain. *Journal of Cognitive Neuroscience* **27**.2 (2014), 222–233. DOI: 10.1162/jocn_a_00702.

288. Sangha S, Chadick JZ, and Janak PH. Safety Encoding in the Basal Amygdala. *Journal of Neuroscience* **33**.9 (2013), 3744–3751. DOI: 10.1523/JNEUROSCI.3302-12.2013.

289. Schaible HG and Grubb BD. Afferent and Spinal Mechanisms of Joint Pain. *Pain* **55**.1 (1993), 5–54. DOI: 10.1016/0304-3959(93)90183-P.

290. Schiller D, Levy I, Niv Y, LeDoux JE, and Phelps EA. From Fear to Safety and Back: Reversal of Fear in the Human Brain. *Journal of Neuroscience* **28**.45 (2008), 11517–11525. DOI: 10.1523/JNEUROSCI.2265-08.2008.

291. Schlagenhauf F, Huys QJM, Deserno L, Rapp MA, Beck A, Heinze HJ, Dolan R, and Heinz A. Striatal Dysfunction during Reversal Learning in Unmedicated Schizophrenia Patients. *NeuroImage* **89** (2014), 171–180. DOI: 10.1016/j.neuroimage.2013.11.034.

292. Schlund MW, Siegle GJ, Ladouceur CD, Silk JS, Cataldo MF, Forbes EE, Dahl RE, and Ryan ND. Nothing to Fear? Neural Systems Supporting Avoidance Behavior in Healthy Youths. *NeuroImage* **52**.2 (2010), 710–719. DOI: 10.1016/j.neuroimage.2010.04.244.

293. Schonberg T, O'Doherty JP, Joel D, Inzelberg R, Segev Y, and Daw ND. Selective Impairment of Prediction Error Signaling in Human Dorsolateral but Not Ventral Striatum in Parkinson's Disease Patients: Evidence from a Model-Based fMRI Study. *NeuroImage* **49**.1 (2010), 772–781. DOI: 10.1016/j.neuroimage.2009.08.011.

294. Schultz W. Behavioral Theories and the Neurophysiology of Reward. *Annual Review of Psychology* **57** (2006), 87–115. DOI: 10.1146/annurev.psych.56.091103.070229.

295. Schultz W. Predictive Reward Signal of Dopamine Neurons. *Journal of Neurophysiology* **80**.1 (1998), 1–27. DOI: 10.1152/jn.1998.80.1.1.

296. Schultz W. Updating Dopamine Reward Signals. *Current Opinion in Neurobiology*. Macrocircuits **23**.2 (2013), 229–238. DOI: 10.1016/j.conb.2012.11.012.

297. Schultz W, Apicella P, and Ljungberg T. Responses of Monkey Dopamine Neurons to Reward and Conditioned Stimuli during Successive Steps of Learning a Delayed Response Task. *Journal of Neuroscience* **13**.3 (1993), 900–913.

298. Schultz W and Dickinson A. Neuronal Coding of Prediction Errors. *Annual Review of Neuroscience* **23**.1 (2000), 473–500. DOI: 10.1146/annurev.neuro.23.1.473.

299. Schulz E, Zherdin A, Tiemann L, Plant C, and Ploner M. Decoding an Individual's Sensitivity to Pain from the Multivariate Analysis of EEG Data. *Cerebral Cortex* **22**.5 (2012), 1118–1123. DOI: 10.1093/cercor/bhr186.

300. Schwarz G. Estimating the Dimension of a Model. *Annals of Statistics* **6**.2 (1978), 461–464. DOI: 10.1214/aos/1176344136.

301. Segerdahl AR, Mezue M, Okell TW, Farrar JT, and Tracey I. The Dorsal Posterior Insula Subserves a Fundamental Role in Human Pain. *Nature Neuroscience* **18**.4 (2015), 499–500. DOI: 10.1038/nn.3969.

302. Segerdahl AR, Themistocleous AC, Fido D, Bennett DL, and Tracey I. A Brain-Based Pain Facilitation Mechanism Contributes to Painful Diabetic Polyneuropathy. *Brain* **141**.2 (2018), 357–364. DOI: 10.1093/brain/awx337.

303. Seymour B. Aversive Reinforcement Learning. UCL (University College London), 2010.

304. Seymour B, Barbe M, Dayan P, Shiner T, Dolan R, and Fink GR. Deep Brain Stimulation of the Subthalamic Nucleus Modulates Sensitivity to Decision Outcome Value in Parkinson's Disease. *Scientific Reports* **6** (2016). DOI: 10.1038/srep32509.

305. Seymour B, Daw ND, Roiser JP, Dayan P, and Dolan R. Serotonin Selectively Modulates Reward Value in Human Decision-Making. *Journal of Neuroscience* **32**.17 (2012), 5833–5842. DOI: 10.1523/JNEUROSCI.0053-12.2012.

306. Seymour B, O'Doherty JP, Dayan P, Koltzenburg M, Jones AK, Dolan RJ, Friston KJ, and Frackowiak RS. Temporal Difference Models Describe Higher-Order Learning in Humans. *Nature* **429**.6992 (2004), 664–667. DOI: 10.1038/nature02581.

307. Seymour B, O'Doherty JP, Koltzenburg M, Wiech K, Frackowiak R, Friston K, and Dolan R. Opponent Appetitive-Aversive Neural Processes Underlie Predictive Learning of Pain Relief. *Nature Neuroscience* **8**.9 (2005), 1234–1240. DOI: 10.1038/nn1527.

308. Shackman AJ, Salomons TV, Slagter HA, Fox AS, Winter JJ, and Davidson RJ. The Integration of Negative Affect, Pain and Cognitive Control in the Cingulate Cortex. *Nature Reviews Neuroscience* **12**.3 (2011), 154–167. DOI: 10.1038/nrn2994.

309. Shibata K, Watanabe T, Sasaki Y, and Kawato M. Perceptual Learning Incepted by Decoded fMRI Neurofeedback Without Stimulus Presentation. *Science* **334**.6061 (2011), 1413–1415. DOI: 10.1126/science.1212003.

310. Singer T, Seymour B, O'Doherty J, Kaube H, Dolan RJ, and Frith CD. Empathy for Pain Involves the Affective but Not Sensory Components of Pain. *Science* **303**.5661 (2004), 1157–1162. DOI: 10.1126/science.1093535.

311. Sitaram R, Ros T, Stoeckel L, Haller S, Scharnowski F, Lewis-Peacock J, Weiskopf N, Blefari ML, Rana M, Oblak E, Birbaumer N, and Sulzer J. Closed-Loop Brain Training: The Science of Neurofeedback. *Nature Reviews Neuroscience* **18**.2 (2017), 86–100. DOI: 10.1038/nrn.2016.164.

312. Skinner BF. *The Behavior of Organisms: An Experimental Analysis*. B. F. Skinner Foundation, 1938. 488 pp.

313. Solomon RL. The Opponent-Process Theory of Acquired Motivation: The Costs of Pleasure and the Benefits of Pain. *American Psychologist* **35**.8 (1980), 691. DOI: 10.1037/0003-066X.35.8.691.

314. Solomon RL and Corbit JD. An Opponent-Process Theory of Motivation: I. Temporal Dynamics of Affect. *Psychological Review* **81**.2 (1974), 119. DOI: 10.1037/h0036128.

315. Staib M, Castegnetti G, and Bach DR. Optimising a Model-Based Approach to Inferring Fear Learning from Skin Conductance Responses. *Journal of Neuroscience Methods* **255** (Supplement C 2015), 131–138. DOI: 10.1016/j.jneumeth.2015.08.009.

316. Stein N, Sprenger C, Scholz J, Wiech K, and Bingel U. White Matter Integrity of the Descending Pain Modulatory System Is Associated with Interindividual Differences in Placebo Analgesia. *Pain* **153**.11 (2012), 2210–2217. DOI: 10.1016/j.pain.2012.07.010.

317. Steinmetz JE. Brain Substrates of Classical Eyeblink Conditioning: A Highly Localized but Also Distributed System. *Behavioural Brain Research* **110**.1 (2000), 13–24. DOI: 10.1016/S0166-4328(99)00181-3.

318. Stephan KE, Penny WD, Daunizeau J, Moran RJ, and Friston KJ. Bayesian Model Selection for Group Studies. *NeuroImage* **46**.4 (2009), 1004–1017. DOI: 10.1016/j.neuroimage.2009.03.025.

319. Stoodley CJ and Schmahmann JD. Functional Topography in the Human Cerebellum: A Meta-Analysis of Neuroimaging Studies. *NeuroImage* **44**.2 (2009), 489–501. DOI: 10.1016/j.neuroimage.2008.08.039.

320. Sutton RS. Adapting Bias by Gradient Descent: An Incremental Version of Delta-Bar-Delta. *AAAI*. 1992, 171–176.

321. Sutton RS and Barto AG. *Introduction to Reinforcement Learning*. Cambridge, MA, USA: MIT Press, 1998. 344 pp.

322. Taylor JJ, Borckardt JJ, and George MS. Endogenous Opioids Mediate Left Dorsolateral Prefrontal Cortex rTMS-Induced Analgesia. *Pain* **153**.6 (2012), 1219–1225. DOI: 10.1016/j.pain.2012.02.030.

323. Thieme A, Thürling M, Galuba J, Burciu RG, Göricke S, Beck A, Aurich V, Wondzinski E, Siebler M, Gerwig M, Bracha V, and Timmann D. Storage of a Naturally Acquired Conditioned Response Is Impaired in Patients with Cerebellar Degeneration. *Brain* (2013), awt107. DOI: 10.1093/brain/awt107.

324. Thorndike EL. *Animal Intelligence: Experimental Studies*. Routledge, 1911. 335 pp.

325. Thorndike EL. *The Fundamentals of Learning*. Vol. xvii. New York, NY, US: Teachers College Bureau of Publications, 1932. 638 pp.

326. Timmann D, Baier PC, Diener HC, and Kolb FP. Classically Conditioned Withdrawal Reflex in Cerebellar Patients. 1. Impaired Conditioned Responses. *Experimental Brain Research* **130**.4 (2000), 453–470. DOI: 10.1007/s002219900225.

327. Tononi G, Sporns O, and Edelman GM. Measures of Degeneracy and Redundancy in Biological Networks. *Proceedings of the National Academy of Sciences* **96**.6 (1999), 3257–3262. DOI: 10.1073/pnas.96.6.3257.

328. Torrecillos F, Albouy P, Brochier T, and Malfait N. Does the Processing of Sensory and Reward-Prediction Errors Involve Common Neural Resources? Evidence from a Frontocentral Negative Potential Modulated by Movement Execution Errors. *Journal of Neuroscience* **34**.14 (2014), 4845–4856. DOI: 10.1523/JNEUROSCI.4390-13.2014.

329. Tracey I and Mantyh PW. The Cerebral Signature for Pain Perception and Its Modulation. *Neuron* **55**.3 (2007), 377–391. DOI: 10.1016/j.neuron.2007.07.012.

330. Tracey I, Ploghaus A, Gati JS, Clare S, Smith S, Menon RS, and Matthews PM. Imaging Attentional Modulation of Pain in the Periaqueductal Gray in Humans. *Journal of Neuroscience* **22**.7 (2002), 2748–2752.

331. Tricomi E, Balleine BW, and O'Doherty JP. A Specific Role for Posterior Dorsolateral Striatum in Human Habit Learning. *European Journal of Neuroscience* **29**.11 (2009), 2225–2232. DOI: 10.1111/j.1460-9568.2009.06796.x.

332. Turk DC and Melzack R. *Handbook of Pain Assessment*. Guilford Press, 2011. 562 pp.

333. Valet M, Sprenger T, Boecker H, Willoch F, Rummeny E, Conrad B, Erhard P, and Tolle TR. Distraction Modulates Connectivity of the Cingulo-Frontal Cortex and the Midbrain during Pain—an fMRI Analysis. *Pain* **109**.3 (2004), 399–408. DOI: 10.1016/j.pain.2004.02.033.

334. Vlaeyen JW. Learning to Predict and Control Harmful Events: Chronic Pain and Conditioning. *Pain* **156** (2015), S86–S93. DOI: 10.1097/j.pain.0000000000000107.

335. Vogt BA. *Cingulate Neurobiology and Disease*. Oxford University Press, 2009. 865 pp.

336. Vogt BA. Pain and Emotion Interactions in Subregions of the Cingulate Gyrus. *Nature Reviews Neuroscience* **6**.7 (2005), 533–544. DOI: 10.1038/nrn1704.

337. Vogt BA, Vogt L, Farber NB, and Bush G. Architecture and Neurocytology of Monkey Cingulate Gyrus. *The Journal of Comparative Neurology* **485**.3 (2005), 218–239. DOI: 10.1002/cne.20512.

338. Wager TD and Atlas LY. The Neuroscience of Placebo Effects: Connecting Context, Learning and Health. *Nature Reviews Neuroscience* **16**.7 (2015), 403–418. DOI: 10.1038/nrn3976.

339. Wager TD, Atlas LY, Botvinick MM, Chang LJ, Coghill RC, Davis KD, Iannetti GD, Poldrack RA, Shackman AJ, and Yarkoni T. Pain in the ACC? *Proceedings of the National Academy of Sciences* **113**.18 (2016), E2474–E2475. DOI: 10.1073/pnas.1600282113.

340. Wager TD, Atlas LY, Lindquist MA, Roy M, Woo CW, and Kross E. An fMRI-Based Neurologic Signature of Physical Pain. *New England Journal of Medicine* **368**.15 (2013), 1388–1397. DOI: 10.1056/NEJMoa1204471.

341. Wager TD and Barrett LF. From Affect to Control: Functional Specialization of the Insula in Motivation and Regulation. *bioRxiv* (2017), 102368. DOI: 10.1101/102368.

342. Wager TD, Rilling JK, Smith EE, Sokolik A, Casey KL, Davidson RJ, Kosslyn SM, Rose RM, and Cohen JD. Placebo-Induced Changes in fMRI in the Anticipation and Experience of Pain. *Science* **303**.5661 (2004), 1162–1167. DOI: 10.1126/science.1093065.

343. Wager TD, Scott DJ, and Zubieta JK. Placebo Effects on Human μ-Opioid Activity during Pain. *Proceedings of the National Academy of Sciences* **104**.26 (2007), 11056–11061. DOI: 10.1073/pnas.0702413104.

344. Watkins CJCH. Learning from Delayed Rewards. PhD Thesis. King's College, Cambridge, 1989.

345. Wiech K, Jbabdi S, Lin C, Andersson J, and Tracey I. Differential Structural and Resting State Connectivity between Insular Subdivisions and Other Pain-Related Brain Regions. *Pain* **155**.10 (2014), 2047–2055. DOI: 10.1016/j.pain.2014.07.009.

346. Wiech K. Deconstructing the Sensation of Pain: The Influence of Cognitive Processes on Pain Perception. *Science* **354**.6312 (2016), 584–587. DOI: 10.1126/science.aaf8934.

347. Wiech K, Edwards R, Moseley GL, Berna C, Ploner M, and Tracey I. Dissociable Neural Mechanisms Underlying the Modulation of Pain and Anxiety? An fMRI Pilot Study. *PLoS ONE* **9**.12 (2014), e110654. DOI: 10.1371/journal.pone.0110654.

348. Wiech K, Kalisch R, Weiskopf N, Pleger B, Stephan KE, and Dolan RJ. Anterolateral Prefrontal Cortex Mediates the Analgesic Effect of Expected and Perceived Control over Pain. *Journal of Neuroscience* **26**.44 (2006), 11501–11509. DOI: 10.1523/JNEUROSCI.2568-06.2006.

349. Wiech K, Ploner M, and Tracey I. Neurocognitive Aspects of Pain Perception. *Trends in Cognitive Sciences* **12**.8 (2008), 306–313. DOI: 10.1016/j.tics.2008.05.005.

350. Wiech K and Tracey I. Pain, Decisions, and Actions: A Motivational Perspective. *Decision Neuroscience* **7** (2013), 46. DOI: 10.3389/fnins.2013.00046.

351. Wiech K, Vandekerckhove J, Zaman J, Tuerlinckx F, Vlaeyen JWS, and Tracey I. Influence of Prior Information on Pain Involves Biased Perceptual Decision-Making. *Current Biology* **24**.15 (2014), R679–R681. DOI: 10.1016/j.cub.2014.06.022.

352. Williams ACdC and Craig KD. Updating the Definition of Pain. *Pain* **157**.11 (2016), 2420. DOI: 10.1097/j.pain.0000000000000613.

353. Woo CW, Schmidt L, Krishnan A, Jepma M, Roy M, Lindquist MA, Atlas LY, and Wager TD. Quantifying Cerebral Contributions to Pain beyond Nociception. *Nature Communications* **8** (2017). DOI: 10.1038/ncomms14211.

354. Worsley KJ, Marrett S, Neelin P, Vandal AC, Friston KJ, and Evans AC. A Unified Statistical Approach for Determining Significant Signals in Images of Cerebral Activation. *Human Brain Mapping* **4**.1 (1996), 58–73. DOI: 10.1002/(SICI)1097-0193(1996)4:1<58::AID-HBM4>3.0.CO;2-O.

355. Worsley KJ, Taylor JE, Tomaiuolo F, and Lerch J. Unified Univariate and Multivariate Random Field Theory. *NeuroImage*. Mathematics in Brain Imaging **23** (Supplement 1 2004), S189–S195. DOI: 10.1016/j.neuroimage.2004.07.026.

356. Yamashita O, Sato Ma, Yoshioka T, Tong F, and Kamitani Y. Sparse Estimation Automatically Selects Voxels Relevant for the Decoding of fMRI Activity Patterns. *NeuroImage* **42**.4 (2008), 1414–1429. DOI: 10.1016/j.neuroimage.2008.05.050.

357. Yanagisawa T, Fukuma R, Seymour B, Hosomi K, Kishima H, Shimizu T, Yokoi H, Hirata M, Yoshimine T, Kamitani Y, and Saitoh Y. Induced Sensorimotor Brain Plasticity Controls Pain in Phantom Limb Patients. *Nature Communications* **7** (2016), 13209. DOI: 10.1038/ncomms13209.

358. Yin HH, Knowlton BJ, and Balleine BW. Lesions of Dorsolateral Striatum Preserve Outcome Expectancy but Disrupt Habit Formation in Instrumental Learning. *European Journal of Neuroscience* **19**.1 (2004), 181–189. DOI: 10.1111/j.1460-9568.2004.03095.x.

359. Yoshida W, Seymour B, Koltzenburg M, and Dolan RJ. Uncertainty Increases Pain: Evidence for a Novel Mechanism of Pain Modulation Involving the Periaqueductal Gray. *Journal of Neuroscience* **33**.13 (2013), 5638–5646. DOI: 10.1523/JNEUROSCI.4984-12.2013.

360. Yu AJ and Dayan P. Uncertainty, Neuromodulation, and Attention. *Neuron* **46**.4 (2005), 681–692. DOI: 10.1016/j.neuron.2005.04.026.

361. Yu R, Gollub RL, Spaeth R, Napadow V, Wasan A, and Kong J. Disrupted Functional Connectivity of the Periaqueductal Gray in Chronic Low Back Pain. *NeuroImage: Clinical* **6** (2014), 100–108. DOI: 10.1016/j.nicl.2014.08.019.

362. Zhang S, Green A, and Smith PP. An Automatic Classifier of Pain Scores in Chronic Pain Patients from Local Field Potentials Recordings. *2013 6th International IEEE/EMBS Conference on Neural Engineering (NER)*. 2013, 1194–1197. DOI: 10.1109/NER.2013.6696153.

363. Zhang S, Mano H, Ganesh G, Robbins T, and Seymour B. Dissociable Learning Processes Underlie Human Pain Conditioning. *Current Biology* **26**.1 (2016), 52–58. DOI: 10.1016/j.cub.2015.10.066.

364. Zhang S, Mano H, Lee M, Yoshida W, Kawato M, Robbins TW, and Seymour B. The Control of Tonic Pain by Active Relief Learning. *eLife* **7** (2018), e31949. DOI: 10.7554/eLife.31949.

365. Zhang S and Seymour B. Technology for Chronic Pain. *Current Biology* **24**.18 (2014), R930–R935. DOI: 10.1016/j.cub.2014.07.010.

366. Zhang S, Yoshida W, Mano H, Yanagisawa T, Shibata K, Kawato M, and Seymour B. Endogenous Controllability of Closed-Loop Brain-Machine Interfaces for Pain. *bioRxiv* (2018), 369736. DOI: 10.1101/369736.

367. Zubieta JK, Bueller JA, Jackson LR, Scott DJ, Xu Y, Koeppe RA, Nichols TE, and Stohler CS. Placebo Effects Mediated by Endogenous Opioid Activity on μ-Opioid Receptors. *Journal of Neuroscience* **25**.34 (2005), 7754–7762. DOI: 10.1523/JNEUROSCI.0439-05.2005.