Dissociable Systems for Body-oriented and World-oriented Pain Learning

Yijia Yan

New College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hilary 2025

Abstract

Pain is a crucial cognitive construct that guides flexible decision-making in dangerous environments, evaluating both the external world and our bodily integrity, and linking cognitive evaluation on exteroception with perceptive bodily integrity of interoception. A challenge is to understand its neural representations in the human brain. To address this, we aimed to acquire new knowledge on the conceptual basis of pain underlying two fundamental distinctions: i) cognitive problems located within the external world (as world-model pain) and induced by the bodily actions (as body-model pain) from nociception for pain avoidance learning, and ii) nociceptive problems from the external world (as exteroception) and in the internal body (as interoception) for cognitive modulation on nociception. For the cognitive problem, we suggested a reinforcement learning (RL) model with multidimensional bodily pain input. The model arbitrated between model-free and model-based learning for efficient pain avoidance behaviour. Simulation results showed its effective predictive error-driven avoidance learning over time, akin to human behaviour in different pain conditions, as well as flexible learning in varying pain conditions. We also showed the necessity of input pain perception as a multisensory vector embedded in a body map for effective pain avoidance, as well as partially dysfunctionalised pain avoidance learning following the extent of perceptive bodily impairment. With a desktop-based Virtual Reality (dVR) navigation task conducted in magnetic resonance imaging (MRI), we further provided neuroimaging support to our learning model fitted to the behavioural data. The corresponding pain learning strategy could best describe the data within the respective environmental pain conditions. Neuroimaging results further implied a temporal-parietal difference in favour of different pain learning pathways, a dorsolateral-medial prefrontal difference in predictive pain learning, sensory-motor and insula activations with various pain-predictive signals, and an anterior cingulate circuit correlating with reliability-based arbitration. These suggested a single arbitrative learning

system for human cognitive pain representation learning both external threatening environments and internal bodily integrity with neuroimaging support. For the nociceptive problem, we designed the experiment to discover the interaction of persistent nociception and sensation modulated by discrepant exteroceptive contexts with electroencephalography (EEG) and an immersive Virtual Reality (iVR) game allowing free-operant moving in the open arena. The iVR technique enabled an immersive pain learning environment for human participants to conduct natural bodily movements for realistic pain learning. We showed discrepant representations of body-model and world-model exteroceptive pain in EEG as posterior-anterior potentials anchored to both stimuli and action events, and as body-model beta and world-model frontal theta power oscillations. We further discovered the exteroceptive pain manipulating sensory and illusory nociception with centrofrontal synchronisation in various frequency bands. These supported our hypothesis on the exteroception-modulated sensory and illusory nociception with electroencephalographic support, finalising a primary understanding of such a novel cognitive basis of pain.

Contents

List of Figures

Chapter 1

Introduction

1.1 Conceptual representation

Concepts, 'the units of thought', are considered a powerful approach in both human thought and artificial intelligence (AI) for implementing human high-level cognition [Carey, 2011, Fumagalli and Ferrario, 2019]. The representation of concepts has been greatly studied to explain various cognitive functions like learning, language, interference, decision-making, planning, etc. [Murphy, 2002], and has been largely applied to artificial agents with different conceptual systems for the corresponding cognitive functions [Vernon et al., 2007]. Despite a grand debate on the notion of 'concept' [Fumagalli and Ferrario, 2019], it is not enough to be merely regarded as a linguistic representation of knowledge categories in the context of generalisation learning, which contains knowledge transfer across multiple situations and domains [Banich et al., 2010]. This requires core cognition beyond language, such as agent-based goal-directed performance and object-based interference from causal or spatial relations [Carey, 2011], that necessitates a certain level of abstraction for analogy across multiple cognitive domains [Gallagher and Zahavi, 2012].

Connectionists have provided an implemental perspective on the notion of 'concept' that is compatible with the neural architecture of the brain with: The conceptual representation is distributed across multiple cognitive processing elements, or neural cells in the brain, that are connected with each other in a network [Fodor and Pylyshyn, 1988, Fumagalli and Ferrario, 2019]. From the perspective of neural networks in AI, the conceptual representation would be the activity of those interconnected hidden units [Goodfellow et al., 2016]. For agent-based goal-directed performance, the conceptual representation may be derived from predictive value-driven learning, which is implemented as reinforcement learning (RL) in AI [Sutton and Barto, 1998], with different value-based functionalities found across multiple brain regions [Botvinick

et al., 2020]. For object-based interference, the conceptual representation should also capture the structure of both incoming inputs and internal knowledge for knowledge transfer learning [Kemp and Tenenbaum, 2008]. In the brains across numerous species including humans, this is believed to be embedded in a 'cognitive map' in hippocampal regions (e.g. with hippocampal place cells and entorhinal grid cells among rodents), which interconnects with other relative brain regions, for encoding both spatial and causal relations among entities [Behrens et al., 2018, Hafting et al., 2005, O'Keefe and Nadel, 1978, Tolman, 1948]. Recent AI algorithms inspired by such a 'cognitive map' with entorhinal grid cells show considerable performance in vector-based navigation Banino et al. [2018] and structural knowledge generalisation [Whittington et al., 2020].

Among various human cognitive functions, the pain system is a highly skilled learning 'machine'. It is a fantastic model for understanding the structure of knowledge encoded in the brain because of its i) strong interaction between clear-cut performance (i.e. action) control and perceptual inference, ii) implementation of distinct information processing circuits across a broad range of brain regions, iii) high pervasion and saliency, and iv) crucial functionality of various human behaviours like bodily survival and social interaction [Mohamed and Ott, 2020, Seymour and Mancini, 2020]. Pain also owns unique properties that differ from simple sensory perception or 'negative' value signal (see the next subsection). In fact, numerous machine learning and AI algorithms have been applied to pain research and clinical application for data analysis [Lötsch and Ultsch, 2018]. In addition, AI approaches with safe RL have been designed for avoiding potential dangers and meanwhile maximising returns [García and Fernández, 2015], which resembles the functionality of the human pain system for self-protection [Seymour and Mancini, 2020]. Whereas pain itself as both perception and cognition so far has been less used to enlighten current AI algorithms [Mohamed and Ott, 2020], as we still have less knowledge on the conceptual representation of pain encoded in the human brain [Seymour and Mancini, 2020]: How does the human pain system extract information from multisensory cues to effectively and flexibly avoid potential complex dangers, probably via value-based aversive learning?

AI algorithms so far have yet to contain powerful conceptual representations that serve well for a considerable number of cross-domain cognitive functions [Fumagalli and Ferrario, 2019]. With a conceptual representation of pain discussed in the following subsection, a pain-inspired AI would be crucial for not only a high capability on self-protection against avoidable potential dangers that may cause painful experience, but also understanding the structural knowledge of pain itself and implementing generalisable cognitive functions in artificial agents, especially in single-experience

learning, novel stimuli generalisation, and knowledge transfer among multiple agents [Mohamed and Ott, 2020].

1.2 Cognitive basis of pain learning

Computational and cognitive models of sensory information processing assume that humans aim to infer the causes of afferent sensory input [Friston, 2005, Gregory, 1980, Hatfield, 1990, Körding et al., 2007, Seymour and Mancini, 2020, Tabor et al., 2017]. In the case of pain, this lies in two distinct categories of nociceptive causes: those external to the body that may cause potential harm which is avoidable (e.g. a burning stove in the real world) [Price et al., 2003, Sherrington, 1906], and those internal to the body that reflects potentially damaged body tissue inducing a persistent pain that is not avoidable (e.g. skin that is more sensitive to burn) [Craig, 2002, 2003, Geldard, 1972]. These provide a core nociceptive dichotomy between two aspects of pain: exteroception, of which the knowledge we extract from the external environment for protective behaviours against potential threats [Price et al., 2003]; and interoception, of which the knowledge we obtain from the internal body as physiological condition along with homeostatic behaviours [Craig, 2003, Wall, 1979]. Thus, a fundamental distinction could be figured out in such a conceptual basis of pain: motivational problems with the external world as exteroceptive pain, and nociceptive problems in the internal body as interoceptive pain [Seymour, 2019, Seymour and Mancini, 2020].

From a nociceptive perspective, the pain information of 'how does the brain realise that the body is injured' is inferred and affected from both exteroceptive phasic pain [Price et al., 2003] and interoceptive persistent (or chronic) pain [Di Lernia et al., 2016,]. They are separately processed in two spinothalamic pathways in the brain [Dostrovsky and Craig, 2006]. The exteroceptive pain signal is transmitted in a lateral pathway for sensation incorporating somatosensory discrimination [Melzack and Wall, 1965, Melzack and Casey, 1968]: Lamina neurons project from dorsal horn to primary somatosensory cortex (SI) and secondary somatosensory cortex (SII) via ventrolateral/ventroposterior thalamus for multisensory integration [Apkarian and Hodge, 1989, Craig, 2008, Mackel et al., 1992, Stevens et al., 1993]. Whilst the interoceptive pain signal is transmitted in a medial pathway for homeostatic emotions [Craig, 2003, Strigo and Craig, 2016]: Lamina neurons project from dorsal horn to insula via ventromedial thalamus for homeostatic sensation, as well as to the anterior cingulate cortex (ACC) [Banich et al., 2010] via dorsomedial thalamus for affective motivation [Craig et al., 1994, Craig, 1995, 2002, 2003]. These two pathways may share

bidirectional interactions with each other. On one hand, the persistent interoceptive pain may drive an endogenous control on the sensation of co-occurring exteroceptive pain, e.g. in the conditioned pain modulation [Kennedy et al., 2016]. On the other hand, learning the exteroceptive pain for avoidance learning requires an inhibition on interoceptive persistent pain that is abnormal within chronic pain conditions [Staud, 2012].

From a cognitive perspective, pain is processed with Bayesian interference for prediction, where the sensory encoding is biased by predictive information with statistical optimisation [Anchisi and Zanon, 2015, Seymour and Mancini, 2020, Tabor and Burr, 2019, Yoshida et al., 2013]. For exteroceptive pain with somatosensory discrimination, the lateral projection to SI and SII shows a topographical map for discriminating spatial differences of nociception across body parts [Mancini et al., 2012, 2014]. A previous study has also found evidence of predictive encoding in SI and SII for pain modulated by expectation and violation [Fardo et al., 2017]. That is, the somatosensory map may be responsible for both body-level representation and prediction of pain [Brecht, 2017, Seymour and Mancini, 2020]. For interoceptive pain with homeostatic sensation and affective motivation, the predictive signal has also been reported to correlate with unexpected pain (opposed to features) in the anterior insula [Geuter et al., 2017], in the dorsal putamen, and uncertainty in the pregenual ACC [Zhang et al., 2018].

To incorporate efficient pain avoidance learning within sensory-rich environments, we further propose that the exteroceptive pain could be inferred from the specific pattern of phasic exteroceptive pain stimuli: Whether they point to a primary problem either in subjective bodily experience or from objective external objects [Farkas, 2013]. The idea is similar to the spatial navigational system that switches between the viewpoint-dependent body-model (or self-oriented) frame encoded in parietal regions and viewpoint-independent world-model 'cognitive map' in hippocampal regions [O'Keefe and Nadel, 1978, Vann et al., 2009, Wolbers and Wiener, 2014], which has been applied to AI algorithms inspired by MEC grid cells for vector-based navigation [Banino et al., 2018]. For pain, this lies on a cognitive dichotomy between body-model and world-model aspects of pain, underlying another distinction in the basis of attributing the exteroceptive pain for its avoidance: world-model causes located within the experience-independent external world, and body-model causes induced by experience-dependent bodily movement. To avoid potential threats effectively in advance, cognitive knowledge of exteroceptive pain with world-model aspects from the

external world should be learned and inferred from body-model aspects that could be directly obtained from bodily experience during the early stage of exploration.

How does the brain know which aspect to attribute to the current exteroceptive pain? How do different brain regions encode each of them to infer the cause of exteroceptive pain for future avoidance, which enables an efficient avoidance learning from navigational environments with complex and dynamic dangers? How does the exteroceptive pain learning provide an endogenous control on the persistent sensation (including pain), which itself is not predictable for pain avoidance and meanwhile yielding a chronic sensation (or pain) state appropriate to injury, in order to keep the cognitive functionality of avoidance learning?

We assume the way to infer exteroceptive causes comes from whether the pain is specifically associated with external objective states or internal subjective actions. For instance, if a person only experiences the skin burn in a certain place just next to the burning stove, the pain is likely to be 'world-model' and associated with the stove. On the other hand, if the pain sensation is felt as more sensitive in the left arm only when the person stretches it, the pain is likely to be 'body-model' and associated with the left arm. In this way, a person could learn proper pain avoidance behaviours. Meanwhile, the additional interceptive persistent pain in the body may sensitise/inhibit these pain cognition processes within the relative body part so as to increase/decrease the corresponding pain avoidance behaviour [Moriarty et al., 2011], cf. conditioned pain modulation on nociception [Kennedy et al., 2016]. On the other hand, efficient exteroceptive pain learning and avoidance needs a certain level of endogenous control on interoceptive persistent pain to suppress the unlearnable 'background noise', cf. cognitive dysfunctionality when suffering chronic pain [Moriarty et al., 2011, Staud, 2012].

1.3 Algorithms of cognitive pain learning

From an algorithmic perspective, RL models with temporal difference (TD) may offer insight and draw on models of value-based action learning, such as Pavlovian and instrumental learning [Sutton and Barto, 1981], whilst keeping potential biological plausibility [Sutton and Barto, 1998]. For instrumental learning, the behaviour is modulated by two systems: stimulus-response habitual learning without a model of the external world, and value-based goal-directed learning with simulating the external world [Balleine and Dickinson, 1998, Gläscher et al., 2010]. Thus, to predict exteroceptive pain during navigation within complex environments, an RL model needs

to incorporate both model-based goal-directed learning for predicting world-model pain and model-free habitual learning for predicting body-model pain.

Previous studies have already applied a similar modelling architecture to computational models for decision-making [Lee et al., 2014], spatial navigation [Geerts et al., 2020], and pain avoidance under the decision-making task [Wang et al., 2018, Mahajan et al., 2024]. Neuroimaging studies in decision-making tasks have found model-free signals correlating unexpected pain in the putamen, whilst the combination of model-free and model-based signals correlating unexpected pain in the ventromedial prefrontal cortex (vmPFC) [Lee et al., 2014, Wunderlich et al., 2012], in which they use FORWARD learning and BACKWARD planning for model-based RL whilst SARSA for model-free RL [Gläscher et al., 2010, Lee et al., 2014, Wang et al., 2018]. In addition, hippocampal regions such as medial entorhinal cortex (MEC) and hippocampus (HPC) are suggested to encode a 'cognitive map' of the external world for model-based learning during spatial navigation [Behrens et al., 2018], in which the successor representation is suggested for such model-based RL whilst SARSA for model-free RL [Dayan, 1993, Stachenfeld et al., 2017, Geerts et al., 2020].

We also need to regard pain avoidance learning as not only a fundamental RL problem but also a knowledge representational problem. We notice that apart from traditional reward-based RL problems on learning from a single scalar input [Sutton and Barto, 1981], an effective pain learning can utilise the somatosensory body map as multidimensional inputs [Mancini et al., 2012, 2014]. Furthermore, when inferring exteroceptive pain, a low state-outcome contingency may be more likely to imply a body-model pain, especially if also with a high action-outcome contingency (although this may not exclusively be the case, e.g. if the state itself is body-model). We also suggest that both body-model and world-model pain may be equally controllable, apart from the uncontrollable persistent pain as interoception. As distinguishing world-model pain from body-model pain for modulating interoceptive nociception or sensation may require more sophisticated models of the interaction between external and internal world, the RL-based pain avoidance learning may also rely on accumulating bodily nociceptive evidence in accordance with Bayesian interference [Seymour and Mancini, 2020].

1.4 Aims and hypothesis

The goal of this research is to acquire new knowledge on the conceptual basis of pain, along with developing a new computational model on it, with the connection between

nociception and cognition. This underlies two fundamental distinctions: i) nociceptive problems from the external world (as exteroception) and in the internal body (as interoception) for cognitive modulation on nociception, and ii) cognitive problems located within the external world (as world-model pain) and induced by the bodily actions (as body-model pain) from nociception for pain avoidance learning.

We assume that the human brain could infer exteroceptive causes that come flexibly between external objective states and internal subjective actions. Meanwhile, these two different learning schemas of cognitive pain could either inhibit or exhibit a certain level of interoceptive (pain) sensation, potentially facilitating efficient aversive avoidance learning.

In Chapter 2 we build an RL model (Section 2.2), which incorporates such exteroceptive pain knowledge dissociation and interoceptive pain modulation, for predicting human pain avoidance behaviours during navigation. The model consists of both model-based and model-free RL for body-model and world-model exteroceptive pains, as well as the arbitration between them for flexible learning (Section 2.2.2). **With model simulation and data fitting in Chapter 2, we show that our model provides effective pain avoidance learning for navigational environments that contain complex and dynamic dangers.**

We also design two human experiments that allow us to discover the behavioural and neural evidence of these hypotheses to discover i) **the computational and neural dissociation between body-model and world-model aspects of pain encoded in the brain within an MRI study incorporating a desktop-based Virtual Reality (dVR) game (Section 3.2)**, and ii) **the interaction of persistent nociception and sensation modulated by these two pain cognition processes within an EEG study via an immersive Virtual Reality (iVR) game (Section 4.2)**, in which the iVR technique enables an immersive pain learning environment for human participants to conduct natural bodily movement for realistic pain avoidance behaviour, which has been recently used for phasic and tonic pain valuation [Hewitt et al., 2025, Tong et al., 2025] as well as chronic pain management and treatment [Li et al., 2011, Birckhead et al., 2021, Goudman et al., 2022, Teh et al., 2024]. **In Chapter 3 we show the dissociated exteroceptive pain learning with neuroimaging support**, and **in Chapter 4 we show the sensory and illusory nociception modulated by exteroception with electroencephalographic support.**

In this way, we provide a large-scale neural architecture of the body-model and world-model conceptual basis of exteroceptive pain cognition, as well as the exteroception-modulated interoceptive nociception and sensation, for understanding

the neural mechanism of conceptual pain in the brain within sensory-rich environments. It is possible to provide AI with considerable human cognitive abilities on pain cognition for efficient self-protection as well as interaction among multiple agents and human beings, so as to reinforce industrial productivity and social efficiency. In addition, such a conceptual and computational approach to the pain could provide a new theoretical perspective of pain, benefiting translational research for patients in reality who are suffering various sorts of pain, e.g. by helping them to maintain cognitive function performance that could have been impaired by suffering chronic pain [Moriarty et al., 2011], or to relieve interoceptive noxious sensations given the appropriate high-level cognitive learning task in immersive Virtual Reality.

Chapter 2

Model simulation with arbitrative reinforcement learning (aRL)

2.1 Introduction

Learning to survive the environment includes not only maximising expected rewards but also minimising predicted dangers [Dayan and Daw, 2008, Mobbs et al., 2020]. Such instrumental behaviour is usually believed to be under control of two dichotomic learning systems for evaluating actions [Balleine and Dickinson, 1998, Dayan and Daw, 2008, Drummond and Niv, 2020]: 'automatic' habitual behaviour, which learns a direct stimulus-response contingency, thus provides quick learning; and 'deliberative' goal-directed behaviour, which additionally recruits the internal structural knowledge of the external environment from the input pattern of stimuli, that provides relatively slow but large-scale and generalisable learning. These require both agent-based goal-directed performance and object-based interference from causal or spatial relations [Carey, 2011], which necessitates a certain level of abstraction for analogy across multiple cognitive domains [Gallagher and Zahavi, 2012]. The agent-based goal-directed performance may be derived from predictive value-driven learning, such as reinforcement learning (RL) [Sutton and Barto, 1998], with different value-based functionalities found across multiple brain regions, such as prefrontal and striatal regions [Botvinick et al., 2020, Daw et al., 2005]. Specifically, habitual behaviour is typically described as a model-free learning without explicit models for evaluating stimuli-response associations [Daw et al., 2005, Dayan and Daw, 2008], whilst deliberate behaviour as a model-based learning with an internal model on environmental inputs for online action evaluation [Doya et al., 2002, Daw et al., 2005, Dayan and Daw, 2008]. This duality implies some arbitration that permits flexible learning in complicated environments [Daw et al., 2005, Lee et al., 2014].

Pain, an unpleasant sensory and emotional experience acquired through life experience [Raja et al., 2020], is crucial not only to perceive current (invisible) threats but also to learn to avoid future threats. The pain system is a good model for understanding the structure of knowledge learnt by the brain because of its strong interaction between clear-cut action control and perceptual inference implemented in distinct information processing circuits across a broad range of brain regions [Mohamed and Ott, 2020, Seymour and Mancini, 2020]. Pain also owns unique multidimensional properties that differ from simple sensory perception or unidimensional 'negative' value signals. However, with threat prediction and minimisation, pain is studied insofar as driving protective responses either through directly eliciting action (e.g. through withdrawal reflexes) or by directing learning (as a teaching signal) to avoid pain [Seymour, 2019, Seymour and Mancini, 2020]. Under the framework of dichotomic learning systems of decision-making, what is missing is how we respectively represent and learn pain as a cognitive concept from its perception.

To incorporate efficient pain avoidance learning within complex risky environments, the exteroceptive pain stimuli should be inferred from the specific pattern of input as abstract representations across various levels [Seymour, 2019]. Fast habitual behaviour could be supported by learning a direct association between proximal bodily pain perception and response actions, described as model-free learning [Seymour, 2019, Drummond and Niv, 2020]. On the other hand, slow deliberative behaviour necessitates an internal model with hidden states to encode distal spatial information as additional representations to support an enormous capacity for avoidance performance, described as model-based learning [Tolman, 1948, Dayan and Daw, 2008]. Thus, for learning the exteroceptive pain to capture and avoid external dangers, this lies on a cognitive learning dichotomy underlying the basis of its attribution: a body-model cause induced by experience-dependent bodily movement, and a world-model cause located within the experience-independent external world. That is, whether pain is regarded as a cognitive learning problem, either in the subjective bodily experience that is quick to learn directly, or from objective external objects that necessitate an internal state model of the world [Farkas, 2013]. The idea is similar to the spatial navigational system that flexibly switches between the viewpoint-dependent egocentric frame encoded in parietal regions and viewpoint-independent allocentric 'cognitive map' in hippocampal regions for large-scale navigation [O'Keefe and Nadel, 1978, Vann et al., 2009, Wolbers and Wiener, 2014]. Previous studies have already applied a similar modelling architecture to computational models for decision-making [Lee et al., 2014], spatial navigation [Geerts et al., 2020], and pain avoidance under the decision-making task [Wang et al.,

2018]. Whereas in pain, the somatosensory regions, directly interconnected with motor cortices, as a 'body-model' may be responsible for both body-level representation and prediction of pain for learning its avoidance behaviour [Brecht, 2017, Seymour and Mancini, 2020].

We assume the way to infer exteroceptive causes comes from whether the pain is specifically associated with external objective states or internal subjective actions. From an algorithmic perspective, we need to regard pain avoidance learning as not only a fundamental RL problem but also a knowledge representational problem. When inferring exteroceptive pain, a low state-outcome contingency may be more likely to imply an action-induced pain, especially if also with a high action-outcome contingency (although this may not exclusively be the case, e.g. if the state itself is embedded in the body-model). As distinguishing world-model from body-model pain under interoceptive modulation may require more sophisticated models of the interaction between the external and internal world, the RL-based pain avoidance learning may also rely on accumulating bodily nociceptive evidence with Bayesian inference [Seymour and Mancini, 2020]. We also notice that apart from traditional reward-based RL problems on learning from a single scalar input [Sutton and Barto, 1981], effective pain learning can utilise the somatosensory body map as multidimensional inputs [Mancini et al., 2012, 2014].

Thus with a navigational avoidance task, we suggested a reinforcement learning (RL) model with multidimensional bodily pain input. The model arbitrated between model-free and model-based learning for efficient pain avoidance behaviour. Simulation results showed its effective predictive error-driven avoidance learning over time, akin to human behaviour in different pain conditions, as well as flexible learning in varying pain conditions. We also showed the necessity of input pain perception as a multisensory vector embedded in a body map for effective pain avoidance, as well as partially dysfunctionalised pain avoidance learning following the extent of perceptive bodily impairment. Overall, our work suggests a single arbitrative learning system for human cognitive pain representation learning both external threatening environments and internal bodily integrity.

2.2 Methods

2.2.1 Model simulation task design

The simulations were conducted within a navigational task akin to our human experiment with a computer-based virtual spatial navigation task with a first-person view

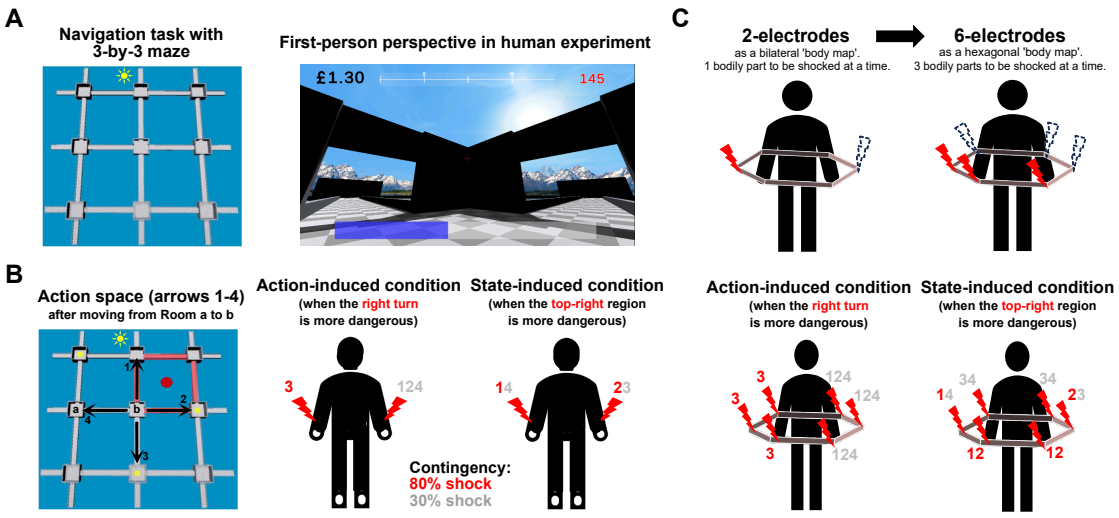

Figure 2.1: **Spatial aversive learning task within the virtual environment with multidimensional negative inputs (cf. electric shocks onto the human body) given under different pain conditions that required the corresponding learning strategies.**

(A) A virtual environment scenery consistent with the human experimental setup in Section 3.2 where participants conducted the task with the first-person perspective. **(B)** Example of different 'pain' (as a 2-element vector with negative values) stimulation patterns in the action-induced condition (right turn as dangerous turn) and state-induced condition (dangerous centre in the first quadrant) when agents moved from Room a to b. Red rectangles stand for dangerous aisles, with a red circle as their dangerous centre. Black numbers beside the electrode on the corresponding arm (right) refer to the action (left) leading to a high shocking probability, whilst grey numbers refer to a low shocking probability. **(C)** Example of 'pain' which was the same as shown in B, but instead receiving a 6-element 'pain' stimulation vector with somatosensory topology. Notice each body part had an independent shock-receiving possibility from others.

(Figure 2.1A). The virtual environment had stably dangerous turning directions or spatial places that might cause electric 'painful shocks' as multidimensional negative 'reward' inputs. Figure 2.1B provided an example of a 2-element vector input, corresponding to the bilateral electrodes applied to the human skin surface of both arms, with 1 for each side on the arm in our human experiment. Figure 2.1C provided an example of a 6-element vector input as a more complicated multidimensional bodily perception input with somatosensory topology.

Each place in the virtual environment was considered as a simple square open room, with each neighbouring room connected by a narrow open aisle, so that the whole environment becomes a 3×3 square grid (Figure 2.1A, left). During the navigational

task in each simulation episode block with 48 room visits, agents might receive a 'shock' that projects onto their virtual topological 'body map' as a multidimensional negative 'reward' input as soon as they leave the room to a dangerous aisle (Figure 2.1B, left), which depended on the movement and pain conditions described below in details.

2.2.1.1 Action-induced pain condition

For the action-induced pain condition with body-model pain, agents had an action contingency q among different actions to receive a 'shock', depending on whether to do a quarter turn when leaving the room (Figure 2.1B, middle). The quarter turn referred to either the left turn or the right turn. Agents had the conditioned probability $q_+ = 0.8$ to receive a shock when doing a dangerous quarter turn, whilst having the conditioned probability $q_- = 0.3$ to receive a shock when not doing a dangerous quarter turn, so that the action contingency $q = q_+ - q_- = 0.5$. With a specific (i.e. either left or right) quarter turn in the corresponding block, half of the body parts within the ipsilateral side might receive the shock; Otherwise, those within the contralateral side might receive the shock (Figure 2.1C, left).

2.2.1.2 State-induced pain condition

For the state-induced pain condition with world-model pain, participants had a state contingency p among different states to receive a 'shock', depending on the position of dangerous places (Figure 2.1B, right). Here, invisible dangerous places composed a cluster of 4 adjacent aisles that connected a 2×2 region of rooms. More specifically, agents had the conditioned probability $p_+ = 0.8$ to receive a shock when entering a dangerous aisle, whilst having the conditioned probability $p_- = 0.3$ to receive a shock when entering a non-dangerous aisle, so that the state contingency $p = p_+ - p_- = 0.5$. Only half of the body parts (as a hemisphere of the body) might receive a shock if those body parts were more exposed to the dangerous centre in the virtual environment (Figure 2.1C, right).

2.2.1.3 Pain avoidance performance (PAP) quantification

Agents conducted aversive learning on the task with pain avoidance performance over time. We defined pain avoidance performance (PAP) as the action that led to a lower possibility of receiving shocks when leaving the room. Each action could be a left turn, a right turn, moving forward, or moving backwards. For the action-induced condition, PAP was the other three turning directions apart from the dangerous one. For the

state-induced condition, PAP was the turning direction that did not directly lead to a dangerous aisle. To show the temporal learning progress at the behavioural level, we calculated PAP as a percentage of PAP per se over all actions separately within 4 sequential segments in a simulation episode, each for the corresponding segment with 12 room visits. This was consistent with our human experiment, which allowed 4 minutes for each navigational block that allowed participants to have around 40-50 room visits (Section 3.2), so that the simulation required fast learning within limited room visits rather than hundreds of simulation trials. Chance levels were then regarded as a pure action guess strategy calculated over all pairs of room and action. That is, $52/68 = 76.47\%$ for action-induced condition guess rate and $44/68 = 64.71\%$ for state-induced condition guess rate.

2.2.1.4 Simulation protocols

Two different simulation protocols were conducted, depending on whether the pain condition was varying during the whole episode. For the simulation within a stable pain condition, a simulation episode included 4 consecutive segments, with each segment summarising PAP over 12 steps. For the simulation with varying pain conditions, a simulation episode included 8 consecutive segments (12 steps per segment), with each condition lasting for 2 segments before switching to the other, so that the agent could have enough steps to learn pain avoidance against environment instability. Simulation episodes were averaged for each of the $N = 30$ simulations with different random initial parameters as different random agents. The whole learning simulation would then include 8 episodes for pain avoidance learning. Other simulation protocol changes, including 2 or 6 input body parts, perceptive bodily impairment on pain signal inputs, and cognitive bodily impairment on state pain coding, would be specified in the next subsection.

2.2.2 Computational Modelling

We suggested reinforcement learning (RL) underlying arbitration between two pain maps (Figure 2.2): A model-free RL for the bodily pain map, in parallel with a model-based RL for the world pain map, with both coding value function for deciding navigational actions in each room (cf. striatal model-free and hippocampal model-based RL unification for navigation in Geerts et al. [2020]). We used successor representation for predictive coding of navigational states (here refer to rooms) in the model-based RL [Stachenfeld et al., 2017]. Compared to the classic model-based RL that slowly

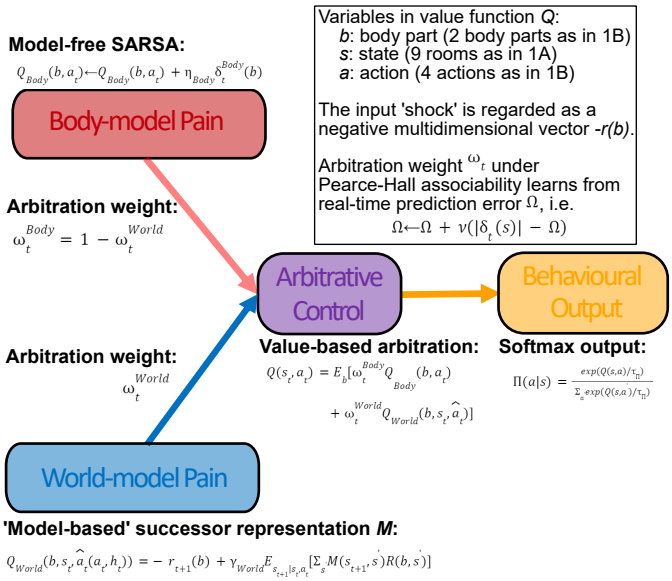

Figure 2.2: **RL model with arbitration between model-free and model-based learning that provided a computational architecture of dichotomic cognitive learning on pain within complex environments.**

The model incorporated a model-free SARSA (red square) in parallel with a model-based successor representation RL (Stachenfeld et al. [2017]; blue square) for learning state transition in the external world, with both coding value functions arbitrating with each other (purple square) for deciding navigational actions in each room (yellow square). See Seymour and Mancini [2020] for the neurobiological support. Arbitrating weights ω_t were learnt based on the real-time prediction error under Pearce-Hall associability [Krugel et al., 2009]. The model also contained the input pain perception as a multisensory vector embedded in a body map (i.e. b in the value function), rather than a single scalar for classical reward-based task modelling.

updates the state transition upon each state visit, the successor representation provides a fast and efficient approximation to the state transition with state occupancy in spatial navigation tasks [Gershman, 2018], so that it models the predictive relationship among external spatial states factorised out of state-based pain value coding.

2.2.2.1 Input pain signals

Based on the topological encoding of body parts in the somatosensory cortex, we defined the input firing-rate vector

$$\Lambda = (r(1), r(2), ..., r(n)),$$

where the total number of body parts $N_{body} = |\Lambda|$ in the simulation. The total number of actions $N_{action} = 4$ for turning left, turning right, moving forward, and turning back

in each room. Each body part b was facing towards $\Phi(b) \in [-\pi, \pi)$ (minus for the left and plus for the right) for the allocentric direction under the initial subject head direction $\theta_0 = 0$ (i.e. towards north) at the start of each episode. Depending on the simulation setup, we set

$$\Phi = (\Phi(1), ..., \Phi(N_{body})) = (\Phi(1), \Phi(2)) = (-\pi/2, \pi/2)$$

when the agents had $N_{body} = 2$ different body parts to perceive shocks as a simple 2-element vector pain input apart from a single scalar reward, and

$$\Phi = (\Phi(1), ..., \Phi(N_{body})) = (\Phi(1), ..., \Phi(6)) = (-\pi/2, -\pi/6, \pi/6, \pi/2, 5\pi/6, -5\pi/6)$$

when the agents had $N_{body} = 6$ different body parts to perceive shocks as a more complicated multidimensional bodily perception input with somatosensory topology.

When receiving a shock, each element in Λ at time t referred to the normalised somatosensory neuron firing $r_t(b) = 1/\sum_b r_t(b)$ for a specific part of the body b to receive a real electric shock, and $r_t(b) = 0$ otherwise. When not receiving a shock, all elements in Λ remained as $r_t(b) = 0$. Moreover, for simulations addressing perceptive bodily impairment, a set of perceptively impaired bodily parts Φ_I were prevented from receiving pain signal inputs, so that

$$r_t(b) = 0, \forall b \in \Phi_I. \tag{2.1}$$

2.2.2.2 Model-free RL

For each body part b, as well as for each room state s_t (i.e. when leaving the current room) and the corresponding action a_t every time, the state-independent somatosensory map encoded the corresponding value function as (Figure 2.2; red square)

$$Q_{Body}(b, a_t) \leftarrow Q_{Body}(b, a_t) + \eta_{Body}\delta_t^{Body}(b), \tag{2.2}$$

where the temporal difference (TD) was calculated via the on-policy SARSA

$$\delta_t^{Body}(b) = -r_{t+1}(b) + \gamma_{Body}Q_{Body}(b, a_{t+1}) - Q_{Body}(b, a_t). \tag{2.3}$$

The vector $-r_{t+1} = -(r_{t+1}(1), ..., r_{t+1}(N_{body}))$ referred to a multidimensional 'negative reward' as pain input rather than a single scalar as in traditional RL models.

2.2.2.3 Model-based RL

The allocentric map encoded the corresponding value function with the factorisation into spatial successor representation [Stachenfeld et al., 2017] and state pain function as (Figure 2.2; blue square)

$$Q_{World}(b, s_t, \hat{a}_t(a_t, h_t)) = -r_{t+1}(b) + \gamma_{World}\mathbb{E}_{s_{t+1}|s_t,a_t}[\sum_{s'} M(s_{t+1}, s')R(b, s')], \quad (2.4)$$

where $\hat{a}_t(a_t, h_t)$ referred to the current allocentric moving direction (NEWS) decided by the head direction h_t and egocentric turn a_t made in the room. The spatial successor representation

$$M(s_t, s') \leftarrow M(s_t, s') + \eta_{SR}\delta_t^{SR}(s') \quad (2.5)$$

was updated with the corresponding TD as

$$\delta_t^{SR}(s') = \mathbb{1}_{s'=s_t} + \gamma_{SR}M(s_{t+1}, s') - M(s_t, s') \quad (2.6)$$

where $\mathbb{1}_{s'=s_t}$ was a binary variable for whether s' is s_t, and the state pain function

$$R(b, s_t) \leftarrow R(b, s_t) + \eta_R\delta_t^{Pain}(b, s_t) \quad (2.7)$$

was updated via

$$\delta_t^{Pain}(b, s_t) = -r_{t+1}(b) - \sum_{b'} F(\Phi(b'), \Phi(b); \kappa)R(b', s_t), \quad (2.8)$$

where $\Phi(b) \in [-\pi, \pi)$ was the simplified 1-dimensional angular position of the body part b ($\Phi(b) = 0$ for the left arm and goes positive clockwise), $\Delta\Phi = 2\pi/N_{body}$ for the angular interval, and the angular bodily weights

$$F(\Phi(b'), \Phi(b); \kappa) = \int_{\Phi(b')-\Delta\Phi/2}^{\Phi(b')+\Delta\Phi/2} f(\varphi, \Phi(b); \kappa)d\varphi \quad (2.9)$$

depending on the different prior selections of the bodily weight function $f(\varphi, \varphi_0; \kappa)$ described below.

2.2.2.4 Bodily weight function

We discovered different ways of projecting bodily perception to the state pain coding in Equations 2.7 and 2.8 by selecting the following bodily weight functions:

- The von Mises distribution addressed the bodily perception projecting more to the state pain coding with closer angular distance controlled by an encoding precision κ as

$$f(\varphi, \varphi_0; \kappa) = \frac{e^{\kappa cos(\varphi - \varphi_0)}}{2\pi I_0(\kappa)}, \tag{2.10}$$

  where $I_0(\kappa)$ was the modified Bessel function of the first kind with order 0. A special case when $\kappa \to +\infty$ referred to the strict 1-to-1 projection from the bodily perception to the state pain coding (disregarding the influence of mismatched body parts) as

$$f(\varphi, \varphi_0) = \mathbb{1}_{\varphi=\varphi_0}. \tag{2.11}$$

  Here $\mathbb{1}_{\varphi=\varphi_0}$ was a binary variable of the body part correspondence for whether φ is φ_0; and

- The cognitive bodily impairment addressed a set of cognitively impaired bodily parts Φ_I without receiving bodily perception projection to the state pain coding, apart from the perceptive bodily impairment that simply prevented the corresponding input signals (Equation 2.1). The other intact bodily parts received the same bodily perception projection as the von Mises distribution in Equation 2.10, so that

$$f(\varphi, \varphi_0; \kappa) = \begin{cases} 0 & \varphi \in \Phi_I \\ \frac{g(\varphi, \varphi_0; \kappa)}{\sum_{\theta \notin \Phi_I} g(\theta, \varphi_0; \kappa)} & \varphi \notin \Phi_I \end{cases}, \tag{2.12}$$

  where $g(\varphi, \varphi_0; \kappa)$ was the von Mises distribution as Equation 2.10.

2.2.2.5 TD-based weighted arbitration

Performing navigational action in each room was based on a TD-based weighted value function (Figure 2.2; purple square)

$$Q(s_t, a_t) = \mathbb{E}_b[\omega_t^{Body} Q_{Body}(b, a_t) + \omega_t^{World} Q_{World}(b, s_t, \hat{a}_t(a_t, h_t))], \tag{2.13}$$

where the bodily prior $p_{body}(b) = \frac{1}{N_{body}}$. The weight of body-model map ω_t^{Body}, which was designed to increase when the corresponding prediction error $|\delta_t|$ (as an estimate of uncertainty or reliability, see Lee et al. [2014] and Geerts et al. [2020]) decreased, was updated via

$$\omega_t^{Body} \leftarrow \omega_t^{Body} + \tau_\omega(\frac{\alpha_{Body}}{1 + e^{\beta_{Body}\chi_{Body}}}\omega_t^{Body} - \frac{\alpha_{World}}{1 + e^{\beta_{World}\chi_{World}}}\omega_t^{World}), \tag{2.14}$$

and the transition rates of Pearce-Hall associability for reliability estimation [Sutton, 1992, Le Pelley, 2004, Li et al., 2011] were based on the unsigned prediction error [Krugel et al., 2009] and updated via

$$\chi_{Body} \leftarrow \chi_{Body} + \rho_{Body}(1 - \frac{\Omega_{Body}}{max\delta^{Body}} - \chi_{Body}) \tag{2.15}$$

and

$$\chi_{World} \leftarrow \chi_{World} + \rho_{World}(1 - \frac{\Omega_{World}}{max\delta^{World}} - \chi_{World}), \tag{2.16}$$

with the temporal average of absolute prediction errors

$$\Omega_{Body} \leftarrow \Omega_{Body} + \nu_{Body}(|\delta_t^{Body}(b)| - \Omega_{Body}) \tag{2.17}$$

and

$$\Omega_{World} \leftarrow \Omega_{World} + \nu_{World}(|\delta_t^{World}(s)| - \Omega_{World}), \tag{2.18}$$

meanwhile

$$\omega_{World} = 1 - \omega_{Body}. \tag{2.19}$$

2.2.2.6 Action output

The navigational action in each room followed a softmax policy (Figure 2.2; yellow square)

$$p(a|s) = \frac{e^{Q(s,a)/\tau_\Pi}}{\sum_{a'} e^{Q(s,a')/\tau_\Pi}}, \tag{2.20}$$

where

$$\tau_\Pi = \tau_\Pi^{(0)} e^{-\lambda_\Pi t} \tag{2.21}$$

was the temperature parameter compensating for spatial exploration and exploitation.

2.2.2.7 Alternative models

We also compared various alternative models based on whether the real-time bodily 'shock' information (i.e. which body part gets shocked) and current spatial information (i.e. which room) was included in the Q-value learning. These included:

- `mf`: The pure model-free RL learning Q-value on body parts and actions $Q_{Body}(b, a)$ in Equation 2.2. $\omega_{World} \equiv 0$ was set to invalidate model-based learning;

- `mf_noB`: The pure model-free RL learning Q-value only on actions $Q_{Body}(a)$ as

$$Q_{Body}(a_t) \leftarrow Q_{Body}(a_t) + \eta_{Body}\delta_t^{Body}, \tag{2.22}$$

  so that the shock input was regarded as a single scalar $-r_{t+1}$ no matter which body part received the shock, rather than a vector $-(r_{t+1}(1), ..., r_{t+1}(N_{body}))$. $\omega_{World} \equiv 0$ was set to invalidate model-based learning;

- `mb`: The pure model-based RL learning Q-value on body parts, spatial states, and actions $Q_{World}(b, s, a)$ in Equation 2.4. $\omega_{Body} \equiv 0$ was set to invalidate model-free learning;

- `mb_noB`: The pure model-based RL learning Q-value on spatial states and actions $Q_{World}(s, a)$ as

$$Q_{World}(s_t, \hat{a}_t(a_t, h_t)) = -r_{t+1} + \gamma_{World}\mathbb{E}_{s_{t+1}|s_t,a_t}[\sum_{s'} M(s_{t+1}, s')R(s')], \tag{2.23}$$

  so that the shock input was regarded as a single scalar $-r_{t+1}$ no matter which body part received the shock, rather than a vector $-(r_{t+1}(1), ..., r_{t+1}(N_{body}))$. $\omega_{Body} \equiv 0$ was set to invalidate model-free learning;

- `arb`: The arbitrative RL learning on both $Q_{Body}(b, a)$ (`mf`) and $Q_{World}(b, s, a)$ (`mb`) in Equation 2.13;

- `arb_noMFB`: The arbitrative RL learning on both $Q_{Body}(a)$ (`mf_noB`) and $Q_{World}(b, s, a)$ (`mb`), so that the shock input to model-free learning was regarded as a single scalar $-r_{t+1}$ no matter which body part received the shock; and

- `arb_noMBB`: The arbitrative RL learning on both $Q_{Body}(b, a)$ (`mf`) and $Q_{World}(s, a)$ (`mb_noB`), so that the shock input to model-based learning was regarded as a single scalar $-r_{t+1}$ no matter which body part received the shock.

Table 2.1: Free parameters used for various pain learning simulations, chosen based on the data fitting results from our human experiments (Section 3.2 & Chapter 3) with the same navigational task and computational models.

Types	Descriptions	Parameters	Values
Model-free learning	Discount	γ_{Body}	0.53
	Learning rate	η_{Body}	0.49
		ν_{Body}	0.48
Model-based learning	Discount	γ_{World}	0.53
		γ_{SR}	0.53
	Learning rate	η_{SR}	0.52
		η_{R}	0.48
		ν_{World}	0.48
	Pain state encoding precision	κ	0.48
Arbitrative learning	TD-based weighted association	τ_{ω}	1
		ρ_{Body}	1
		ρ_{World}	1
	Tuning parameters in	α_{Body}	2.74
	the weights of pain	α_{World}	2.4
	maps	β_{Body}	0.44
		β_{World}	0.38
Behavioual output	Tuning parameters in	$\tau_{\Pi}^{(0)}$	2.46
	the softmax function	$\lambda_{\Pi}*$	0.05

*Apart from no decay setup used for behavioural data fitting in Chapter 3, the action temperature decay here was set to a small positive number to ensure simulation convergence.

2.2.2.8 List of parameters

See Table 2.1 for a list of the model parameters chosen for all simulations in Chapter 2. These parameters were chosen based on data fitting results as between-subject averages of fitted parameters derived from our human experiments in Section 3.2 & Chapter 3 with the same 2-electrode navigational task and computational models. All free model parameters included:

- η_{Body}, γ_{Body}, and ν_{Body} as learning rates and discounts in the model-free RL;
- γ_{World}, η_{SR}, γ_{SR}, η_{R}, and ν_{World} as learning rates and discounts in the model-based RL;
- κ as temperature parameter in von Mises distribution referring to the pain state encoding precision;
- τ_{ω}, ρ_{Body}, and ρ_{World} as learning rates in TD-based weighted association;
- α_{Body}, α_{World}, β_{Body}, and β_{World} as tuning parameters in the weights of pain maps; and
- $\tau_{\Pi}^{(0)}$ (initial action temperature) and λ_{Π} (decay action temperature) as those controlling temperature parameters for the action with softmax policy. Note λ_{Π} here was set to a small positive number to ensure simulation convergence, apart from the zero decay used for behavioural data fitting with real human behaviours.

2.3 Results

2.3.1 Efficient error-driven pain learning in separate conditions

We first conducted the simulation on $N = 30$ random agents, averaging 8 episodes within a simple navigational task, the same as that used in our human experiment (Figure 2.1A; also see Section 3.2). That is, the arbitrative learning model receiving a 2-electrode pain input vector tried to learn the task either in the action-induced condition (with right-turn as a dangerous action) or the state-induced condition (with top-left quadrant II as a dangerous region). We found the significant temporal increase of PAP overall all agents in both action-induced condition (Figure 2.3A, top-right; $k = 2.51\%, r = .504, p < .001$) and state-induced condition (Figure 2.3B, top-right; $k = 5.50\%, r = .628, p < .001$). Moreover, PAP after the early learning

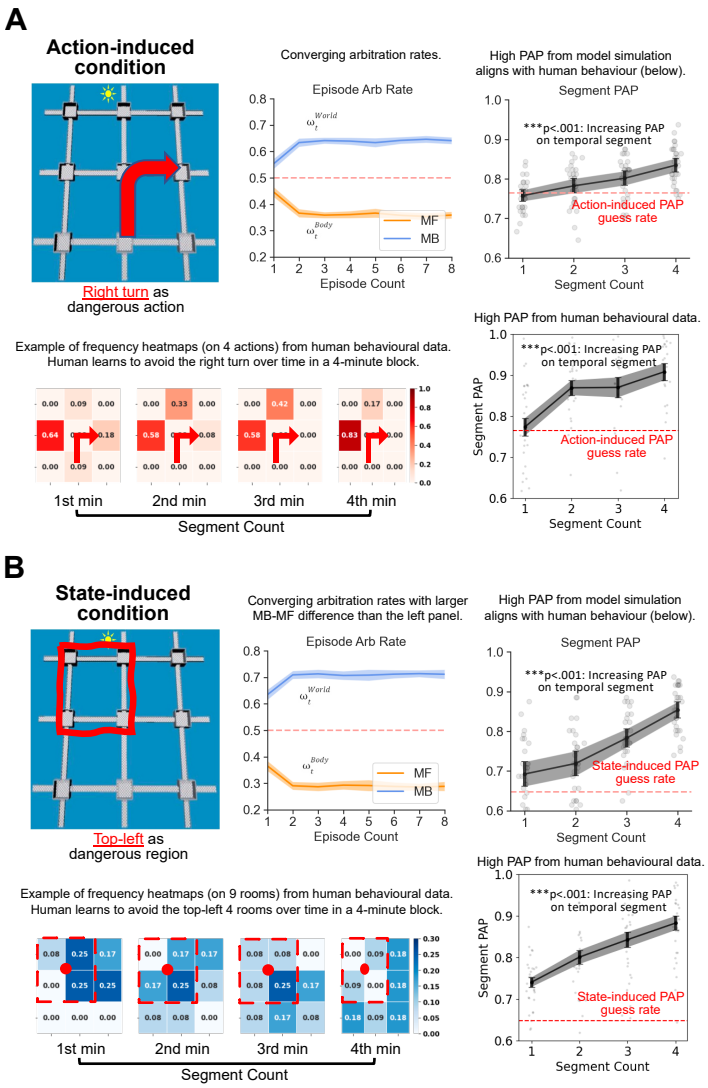

Figure 2.3: **Model simulations showed significant predictive error-driven aversive learning on pain avoidance performance (PAP) in both pain conditions in accordance with human behavioural data.**

(A) An example of an action-induced condition with the dangerous right turn (big red arrow on top-left), with visit frequency heatmaps in each minute on 4 actions in the action-induced condition (bottom). We showed a significant temporal increase of PAP over all participants (top-right) in the action-induced condition. **(B)** An example of a state-induced condition with the dangerous front-left region (red square on top-left), with visit frequency heatmaps in each minute on 9 rooms in the state-induced condition (bottom). We showed a significant temporal increase of PAP over all participants (top-right) in the state-induced condition. Error bars stand for 95% CI. FDR-corrected p-values showed the result of the Wald Test on each segment PAP in the corresponding condition with t-distribution against the zero-slope null hypothesis. *** for $p < .001$.

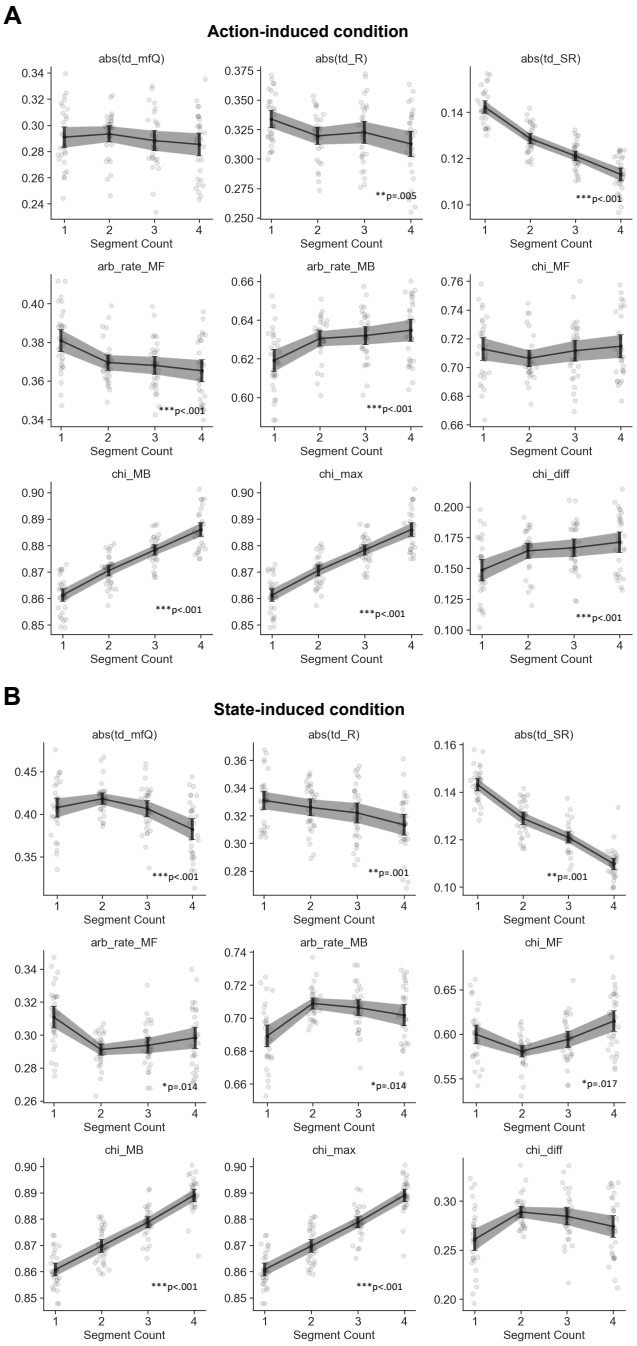

Figure 2.4: **Model simulations suggested an efficient predictive error-driven aversive learning, particularly effectively decreasing learning errors in both pathways when learning state-induced pain.**

Model variables in both conditions (i.e. A for action-induced condition, and B for state-induced condition) over 4 temporal segments showed significant temporal decrease of absolute TD errors and temporal increase of reliabilities that were based on predictive error-driven TD for learning. See Section 2.2.2 for the corresponding abbreviations of each variable. Error bars stand for 95% CI. FDR-corrected p-values showed the result of the Wald Test on each model variable in the corresponding condition with t-distribution against the zero-slope null hypothesis. *** for $p < .001$, ** for $p < .01$, and * for $p < .05$.

stage significantly exceeded the chance level (i.e. pure guess rate of the corresponding pain condition). This could be seen in action-induced condition (Figure 2.3A, top-right) at the 2nd segment ($\Delta = 1.76\%, t(29) = 1.84, p = .076$), 3rd segment ($\Delta = 3.70\%, t(29) = 3.94, p < .001$), and 4th segment ($\Delta = 7.00\%, t(29) = 7.74, p < .001$), as well as in state-induced condition (Figure 2.3B, top-right) at the 1st segment ($\Delta = 4.53\%, t(29) = 2.82, p = .009$), 2nd segment ($\Delta = 7.20\%, t(29) = 4.43, p < .001$), 3rd segment ($\Delta = 13.63\%, t(29) = 11.31, p < .001$), and 4th segment ($\Delta = 20.71\%, t(29) = 19.66, p < .001$). These findings suggested that our models could conduct effective cognitive aversive learning over time under both action-induced and state-induced conditions, which was akin to human behaviours (Figures 2.3A-B, bottom).

We further showed that the arbitrative model could stably learn pain avoidance with converging arbitration rates over training episodes in both action-induced condition (Figure 2.3A, top-middle) and state-induced condition (Figure 2.3B, top-middle). Furthermore, the difference between model-based and model-free arbitration rates was larger in the state-induced condition than the action-induced condition, suggesting a stronger shift towards model-based learning in the corresponding state-induced condition that required state contingency knowledge of the world-model pain.

When looking deeper into simulated model variables (Section 2.2.2) over temporal segments, we found significant temporal decrease of absolute TD errors and temporal increase of reliability that were based on error-driven TD for learning (Figure 2.4). The Wald tests on these model variables were corrected by positive false discovery rates (FDR) with the Benjamini-Hochberg procedure for multiple hypotheses testing [Benjamini and Hochberg, 1995]. The absolute predictive error of model-free Q-values were decreased in state-induced condition ($|\delta^{Body}|$: $k = -.01, r = -.314, p < .001$ in state-induced condition), while model-based Q-values were decreased in both conditions ($|\delta^{Pain}|$: $k = -.01, r = -.261, p = .005$ for action-induced and $k = -.01, r = -.314, p < .001$ for state-induced condition; $|\delta^{SR}|$: $k = -.01, r = -.833, p < .001$ for action-induced and $k = -.01, r = -.865, p = .001$ for state-induced condition). Therefore, the error-based learning reliabilities showed significant increase in both pain conditions (χ^{max}: $k = .01, r = .821, p < .001$ in action-induced condition and $k = .01, r = .849, p < .001$ in state-induced condition; χ^{Body}: $k = .01, r = .220, p = .017$ in state-induced condition; χ^{World}: $k = .01, r = .821, p < .001$ in action-induced condition and $k = .01, r = .849, p < .001$ in state-induced condition). These suggested an efficient predictive error-driven aversive learning of our arbitrative model when facing different pain conditions, particularly effectively decreasing learning errors in both pathways when learning state-induced pain.

2.3.2 Flexible pain learning under varying conditions

To show how our model with a single arbitrative system could learn pain avoidance under different conditions based on the previous simulation, we compared the pain avoidance performance (PAP) between our arbitrative model (`arb`) and other learning models with a single learning strategy as either pure model-free (`mf`) or pure model-based (`mb`) under the corresponding conditions (Figure 2.5A). Under the action-induced condition that required body-model pain learning, the pure `mf` model learnt a significantly higher PAP than pure `mb` ($\Delta = .03, t(29) = 11.92, p < .001$) and `arb` ($\Delta = .03, t(29) = 8.69, p < .001$) models (Figure 2.5A, top). On the other hand, under the state-induced condition that required world-model pain learning, both the pure `mb` ($\Delta = .14, t(29) = 13.06, p < .001$) and `arb` ($\Delta = .10, t(29) = 9.24, p < .001$) models learnt a significantly higher PAP than the pure `mf` model, whilst the pure `mb` model ($\Delta = .04, t(29) = 3.85, p = .001$) learnt a significantly higher PAP than `arb` model (Figure 2.5A, bottom). These suggested pure MF or MB learning strategies favouring the environmental condition with corresponding pain learning knowledge, as well as the mediated pain avoidance learning efficacy in a single arbitrative system when dealing with both conditions compared to single MF/MB learning.

We further conducted a similar simulation in a complex environment with varying conditions to show how our single arbitrative model could have flexible pain avoidance learning against environmental contingency changes (Figure 2.5B). The new simulation was extended to 8 segments, with each condition lasting for 2 segments before switching to the other (first launching as a state-induced condition), so that the agent could have enough steps to learn pain avoidance under varying pain conditions. We still found the overall higher PAP of `arb` models than pure `mf` ($\Delta = .06, t(29) = 12.20, p < .001$) model, as well as a near-significantly higher PAP than pure `mb` ($\Delta = .01, t(29) = 1.89, p = .069$) model (Figure 2.5B, top), along with FDR-corrected Wald tests on model variables showing significantly decreased absolute TD errors ($|\delta^{Body}|$: $k = -.02, r = -.714, p < .001$; $|\delta^{Pain}|$: $k = -.01, r = -.671, p < .001$; $|\delta^{SR}|$: $k = -.01, r = -.962, p < .001$) and increased reliabilities ($|\chi^{Body}|$: $k = .02, r = .712, p < .001$; $|\chi^{World}|$: $k = .01, r = .953, p < .001$; $|\chi^{max}|$: $k = .01, r = .953, p < .001$) over time in both learning pathways of `arb` model under such dynamic environment (Figure 2.5B, bottom). These indicated flexible learning of pain avoidance via arbitrating between predictive error-driven MF and MB learning within the dynamic and complex dangerous environment.

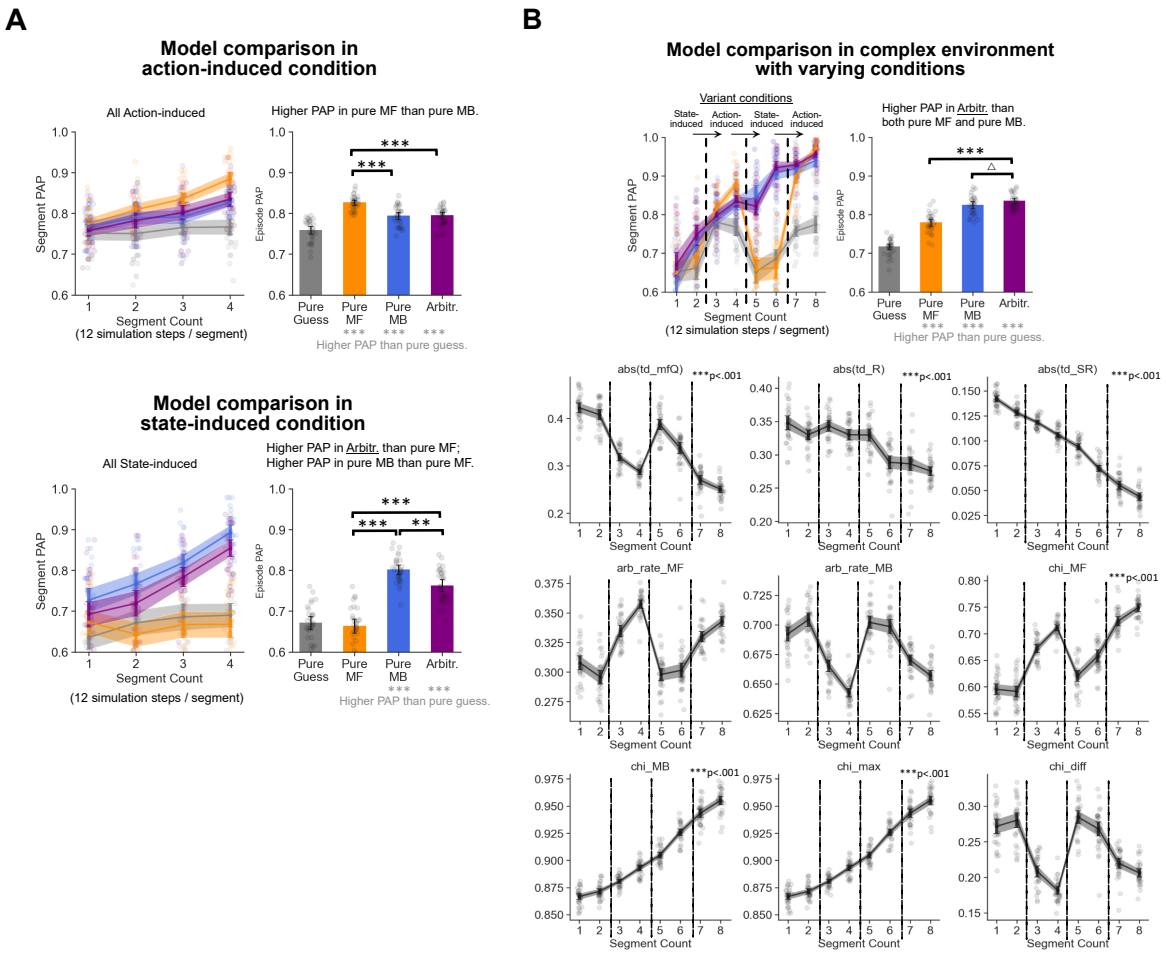

Figure 2.5: **The model outperformed both pure model-based and pure model-free learning across distinct and varying environments.**

(A) The corresponding pain learning models could best perform avoidance within the respective environmental threatening conditions when the rule of shocks was stable. **(B)** When the shocks varied between body-model and world-model pain (8 segments per episode), our arbitrative RL model outperformed both pure model-free and model-based learning, suggesting flexible learning of predictive error-driven pain avoidance within dynamic and complex dangerous environments. Error bars stand for 95% CI. FDR-corrected p-values show the result of the Wald Test on each model variable with a t-distribution against the null hypothesis. *** for $p < .001$, * for $p < .05$, and Δ for $p < .1$.

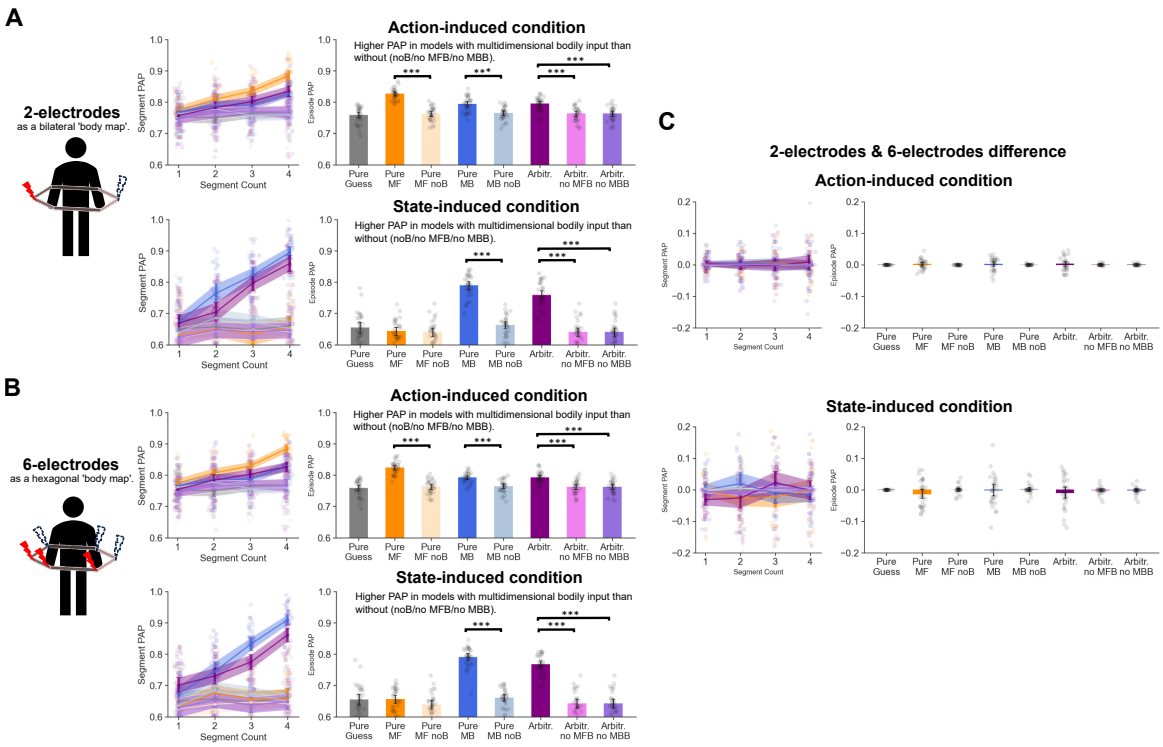

Figure 2.6: **The model necessitated input pain perception as a multisensory vector embedded in a body map for effective pain avoidance learning.**

(A) Apart from a single scalar input for classical reward-based task modelling, models with multidimensional bodily pain inputs outperformed those without (i.e. no body parts in model-free or model-based learning circuits). **(B)** Same simulations but with a 6-electrode input as a hexagonal 'body map' shock input, implying a potential generalisation to the continuous bodily space as topological somatosensory inputs for pain learning. **(C)** No significant PAP decrease was found when generalising the simulation from 2-electrodes to the 6-electrode scenario. Error bars stand for 95% CI. *** for $p < .001$.

2.3.3 Multidimensional bodily inputs as body map

We here aimed to show whether utilising the somatosensory body map as multidimensional inputs was necessary for efficient pain avoidance learning, apart from traditional reward-based RL problems on learning from a single scalar input. We deprived the differentiability of 2 bodily parts for the pain input vector in pure MF learning (`mf_noB`), pure MB learning (`mb_noB`), MF learning pathway of the arbitrative model (`arb_noMFB`), and MB learning pathway of the arbitrative model (`arb_noMBB`) to test this idea (Figure 2.6A). Within action-induced conditions (Figure 2.6A, top), we found higher PAP for `mf` versus `mf_noB` ($\Delta = .06, t(29) = 16.71, p < .001$), higher PAP for `mb` versus `mb_noB` ($\Delta = .03, t(29) = 7.41, p < .001$), as well as higher

PAP for `arb` versus `arb_noMFB` ($\Delta = .03, t(29) = 7.84, p < .001$) and `mf_noMBB` ($\Delta = .03, t(29) = 7.84, p < .001$). Within state-induced conditions (Figure 2.6A, bottom), we also found higher PAP for `mb` versus `mb_noB` ($\Delta = .13, t(29) = 2.74, p < .001$), as well as higher PAP for `arb` versus `arb_noMFB` ($\Delta = .12, t(29) = 10.60, p < .001$) and `mf_noMBB` ($\Delta = .12, t(29) = 10.60, p < .001$). These suggested the necessity of input pain perception as a multisensory vector embedded in a body map for effective pain avoidance learning.

For a more generalised situation, we also showed the simulation results when the number of 'body parts' (cf. electrodes conveying shocks to different parts of the body) for the pain input vector was increased to 6 (Figure 2.6B). Notice each body part had an independent shock receiving possibility from others, implying more complicated and random perceptive pain inputs. Within action-induced conditions (Figure 2.6B, top), we found higher PAP for `mf` versus `mf_noB` ($\Delta = .06, t(29) = 19.67, p < .001$), higher PAP for `mb` versus `mb_noB` ($\Delta = .03, t(29) = 8.33, p < .001$), as well as higher PAP for `arb` versus `arb_noMFB` ($\Delta = .03, t(29) = 8.95, p < .001$) and `mf_noMBB` ($\Delta = .03, t(29) = 8.95, p < .001$). Within state-induced conditions (Figure 2.6B, bottom), we also found higher PAP for `mb` versus `mb_noB` ($\Delta = .13, t(29) = 18.75, p < .001$), as well as higher PAP for `arb` versus `arb_noMFB` ($\Delta = .12, t(29) = 16.45, p < .001$) and `mf_noMBB` ($\Delta = .12, t(29) = 16.45, p < .001$). Furthermore, simulation with 6 body parts showed a certain generalisation capability from those with 2 body parts without deteriorating PAP performance in both conditions (Figure 2.6C). Although still with discrete bodily pain inputs, these would imply a potential generalisation towards a continuous bodily space as topological somatosensory inputs for pain learning.

2.3.4 Dysfunctionalised learning following the extent of perceptive bodily impairment

We further used our arbitrative model `arb` to simulate two types of bodily pain learning impairment. The perceptive bodily impairment on pain nociception prevented some specific bodily parts from receiving pain input signals (Equation 2.1), whilst the cognitive bodily impairment on state pain coding deprived some specific bodily parts of projections from intact sensory pain input in the model-based learning pathway (Equation 2.12). The simulation was conducted with a 6-electrode input so that the impairment on any single body part would not lead to a permanent single scalar input. For the limited body impairment (Figure 2.7, left), we chose to impair either the far-left bodily input $\Phi_I = \{-\pi/2\}$ (deep orange bars as 'Per Left' versus purple bars as 'Intact' body parts) or the far-right bodily input $\Phi_I = \{\pi/2\}$ (Figure 2.7 left, deep blue bar as

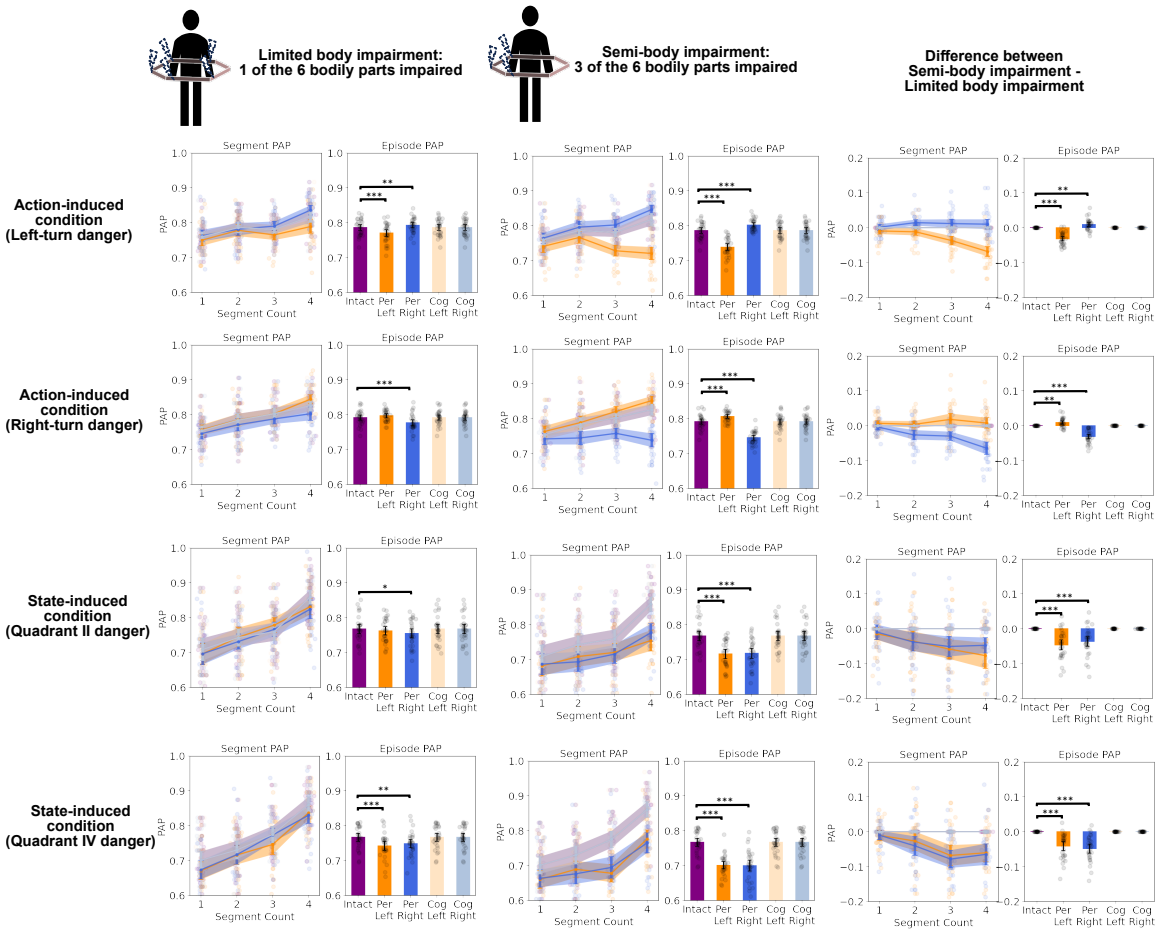

Figure 2.7: **The arbitrative model showed partially dysfunctionalised pain avoidance learning following the extent of perceptive bodily impairment.**

Top 2 rows within action-induced condition showed a significant PAP decrease when learning the ipsilateral quarter turn as a dangerous action, which facilitated the contralateral dangerous quarter turn avoidance, of which the effect was more significant in semi-body impairment than limited body impairment. Bottom 2 rows within the state-induced condition showed decreased PAP against bodily impairment when learning the state-induced exteroceptive pain, of which the effect was more significant in semi-body impairment than limited body impairment. 'Left' and 'Right' for the corresponding bodily impairment. 'Per' for perceptive impairment. 'Cog' for cognitive impairment. Error bars stand for 95% CI. *** for $p < .001$, ** for $p < .01$, and * for $p < .05$.

'Per Right' versus purple bars as 'Intact' body parts) akin to the ipsilateral impairment on human hand as injury. For the semi-body impairment (Figure 2.7, middle), we chose to impair either the whole semi-left bodily input $\Phi_I = \{-\pi/2, -\pi/6, -5\pi/6\}$ or the whole semi-right bodily input $\Phi_I = \{\pi/2, \pi/6, 5\pi/6\}$ akin to the ipsilateral bodily impairment such as hemiplegia.

For the action-induced condition with limited bodily impairment (Figure 2.7, left, 1st & 2nd rows), the arbitrative model showed a decreased PAP to learn the left-turn danger when the left body parts were impaired (Figure 2.7, left, 1st row, Intact - Per Left $\Delta = .02, t(29) = 5.89, p < .001$), as well as a decreased PAP to learn the right-turn danger when the right body parts were impaired (Figure 4.3, left, 2nd row, Intact - Per Right $\Delta = .01, t(29) = 4.35, p < .001$). It also showed an increased PAP to learn the left-turn danger when the right body parts were impaired (Figure 2.7, left, 1st row, Per Right - Intact $\Delta = .01, t(29) = 2.85, p = .008$). This suggested a significant PAP decrease under unilateral bodily impairment when learning the ipsilateral quarter turn as a dangerous action, which facilitated the contralateral dangerous quarter turn avoidance. Similar results were also found for the semi-body impairment (Figure 2.7, middle, 1st row, Intact - Per Left: $\Delta = .05, t(29) = 13.50, p < .001$, Per Right - Intact: $\Delta = .02, t(29) = 5.42, p < .001$; 2nd row, Per Left - Intact: $\Delta = .01, t(29) = 3.98, p < .001$, Intact - Per Right: $\Delta = .05, t(29) = 11.51, p < .001$). The ipsilateral learning inhibition was more significant in semi-body impairment than limited body impairment (Figure 2.7, right, 1st row, Intact - Per Left: $\Delta = .03, t(29) = 9.69, p < .001$; 2nd row, Intact - Per Right: $\Delta = .03, t(29) = 8.88, p < .001$), as well as the contralateral learning facilitation (Figure 2.7, right, 1st row, Per Right - Intact: $\Delta = .01, t(29) = 3.12, p = .004$; 2nd row, Per Left - Intact: $\Delta = .01, t(29) = 3.66, p = .001$). These suggested a model-free learning driven avoidance dysfunctionalising ipsilateral bodily danger learning whilst facilitating contralateral bodily danger learning, of which the effect would be enhanced with broader bodily impairment.

For the state-induced condition with limited bodily impairment (Figure 2.7, left, 3rd & 4th rows), the arbitrative model showed a decreased PAP to learn the exteroceptive state-induced dangers given either side of the impairment (Figure 2.7, left, 3rd row: Intact - Per Right $\Delta = .01, t(29) = 2.27, p = .03$; 4th row: Intact - Per Left $\Delta = .02, t(29) = 3.79, p < .001$, Intact - Per Right $\Delta = .02, t(29) = 3.47, p = .002$). Similar results were also found for the semi-body impairment (Figure 2.7, middle, 3rd row, Intact - Per Left: $\Delta = .05, t(29) = 7.14, p < .001$, Per Right - Intact: $\Delta = .05, t(29) = 7.2, p < .001$; 4th row, Intact - Per Left: $\Delta = .07, t(29) = 8.61, p < .001$,

Intact - Per Right: $\Delta = .07, t(29) = 10.09, p < .001$). The decrease of PAP was more significant with semi-body impairment than limited body impairment (Figure 2.7, right, 3rd row, Intact - Per Left: $\Delta = .05, t(29) = 6.24, p < .001$, Per Right - Intact: $\Delta = .04, t(29) = 5.04, p < .001$; 4th row, Intact - Per Left: $\Delta = .04, t(29) = 6.22, p < .001$, Intact - Per Right: $\Delta = .05, t(29) = 7.29, p < .001$). These suggested a significant universal decrease of PAP under unilateral bodily impairment when learning the exteroceptive state-induced dangers apart from bodily attribution, of which the effect would again be enhanced with broader bodily impairment. Moreover, no dropped PAP was found under the cognitive bodily impairment (Figures 2.7, shallow orange and blue bars as 'Cog Left' and 'Cog Right' respectively) in all situations when the perceptive pain input signals were still intact per se. Altogether, these revealed a partially dysfunctionalised pain avoidance learning within a single arbitrative system against perceptive bodily impairment under various pain conditions and dangerous situations.

2.4 Discussion

We aimed to provide a computational architecture of dichotomic cognitive learning on pain for understanding effective cognitive pain learning within complex and dynamic environments, utilising pain to evaluate both the external world and bodily integrity (Section 1.4). The computational modelling was based on a fundamental dichotomy within cognitive pain learning: the cognitive distinction for avoiding exteroceptive pain learning actions between the objective experience-independent world (i.e. world-model) or the subjective bodily space (i.e. body-model). To address this, we suggested a computational framework with RL incorporating a multidimensional representation of pain for learning pain avoidance during navigation in complex environments. The model consisted of both model-based and model-free RL for body-model and world-model exteroceptive pain, as well as the arbitration between them for flexible learning. Simulation results show its effective predictive error-driven avoidance learning over time, akin to human behaviour in different pain conditions, as well as flexible learning in varying pain conditions. We also showed the necessity of input pain perception as a multisensory vector embedded in a body map for effective pain avoidance, along with partially dysfunctionalised pain avoidance learning following the extent of perceptive bodily impairment. In this way, we suggested a single arbitrative learning system for human cognitive pain representation learning both external threatening environments and internal bodily integrity.

2.4.1 Exteroceptive pain akin to egocentric and allocentric frameworks

The body-model and world-model exteroceptive pain dissociation we suggested both in the model and experiments were enlightened from egocentric vs allocentric spatial representations [Klatzky, 1998]: The self-centred body-model pain as the predictive sensorimotor 'body' map [Brecht, 2017], and the world-centred world-model pain as the predictive external state-oriented 'cognitive' map [Tolman, 1948]. We also used model-free and model-based arbitration as a computational interpretation of their interactions, akin to the similar computational approaches for spatial representation encoding [Geerts et al., 2020]. These implied that exteroceptive pain learning and spatial representation learning might share a certain degree of commonality, though we also hypothesised the cognition-level exteroceptive pain modulating perception-level nociception as the unique information for the brain to realise 'how the body is injured'.

2.4.2 Arbitrative pain learning with neural support

We suggested a reinforcement learning model arbitrating between model-free (MF) and model-based (MB) learning the efficient pain avoidance behaviour. Apart from pain-oriented learning, the neural substrates correlating MF and MB learning were also largely found within the cognitive tasks on learning single scalar rewards [Daw et al., 2005, Lee et al., 2014, Huang et al., 2020]. A quantitative meta-analyses research conducted by Huang et al. [2020] on various fMRI studies showed higher MF learning correlation with the brain activity in caudate head and superior temporal gyrus, whilst medial frontal and anterior cingulate activity correlates more with MB learning. In Chapter 3 we would show that these neural correlations within a reward-oriented learning context accorded with our previous fMRI study in pain-oriented learning, with the neural signature of MF pain learning on the anterior caudate activity for the arbitrative reliability difference correlating with more body-model pain than world-model pain condition, as well as exceedance probabilities favouring pure MF learning correlates more with body-model pain condition in the superior temporal regions, whilst favouring pure MB learning correlated more with world-model pain condition in the parietal regions.

The arbitration between two learning circuits also had various neural supports, such as the dorsal-medial prefrontal activity difference when switching between MF and MB learning in the 2-stage reward-oriented task [Lee et al., 2014], anterior cingular activity facilitating state-induced pain learning pathway via arbitrating with action-induced

pain learning (Chapter 3) as an MF-MB discrepancy on the reliability-driven anterior cingular signals on pain learning [Zhang et al., 2018], and insula activity correlating with world-model condition in learning the state-independent pain temporal difference (Chapter 3) following other studies showing predictive learning signals specific to pain in insula cortex [Brown et al., 2008, Geuter et al., 2017, Fazeli and Büchel, 2018, Koppel et al., 2022, Hoskin and Talmi, 2023]. Chapter 3 would further show that this correlated more with world-model pain in dorsolateral and superior prefrontal regions, in comparison with the medial prefrontal TD learning body-model pain. With successor representation that factorised spatial learning and pain state learning (cf. Stachenfeld et al. [2017], Geerts et al. [2020]), the arbitrative system might manifest a dorsolateral-medial prefrontal difference in the TD learning specific to pain between world-model and body-model learning.

2.4.3 Generalised body map and world map

We showed models with multidimensional bodily pain inputs performing more efficient avoidance behaviour than those without (i.e. no body parts in model-free or model-based learning circuits) even when the number of considered bodily parts was 6. The simulation results implied the bodily parts as a necessary multidimensional input for pain prediction rather than the single scalar input in traditional reward-oriented tasks. This supported a predictive 'body map' responsible for both body-level representation and prediction of pain [Brecht, 2017], with neural evidence on predictive pain signals found among somatosensory and motor regions in our human experiment (Chapter 3) and other human studies [Fardo et al., 2017]. Though the discrete 6-electrode input in our simulation setup was still too simplified to consider a topological somatosensory pain perception in a 2-dimensional surface space covering the whole body, and one of the future directions would be to incorporate a continuous bodily topological space for a more realistic pain avoidance learning [Mancini et al., 2012, 2014].

The successor representation we used for model-based learning addressed a factorisation between spatial-state predictive coding and pain-state evaluation, which enabled flexible learning across multiple environments by explicitly encoding and storing predictive relationships among spatial states as a 'world map' [Tolman, 1948]. It was also suggested as a hippocampal mechanism for between-state predictive learning without explicit planning [Stachenfeld et al., 2017, Geerts et al., 2020], where the hippocampus had been reported showing neuroimaging activations in various reward-oriented model-based learning tasks [Bornstein and Daw, 2012, Sebold et al., 2017]. Future research would consider more complex environmental setups with multiple

scenarios, so that to discover pain avoidance learning under large-scale navigation across sensory-rich environments.

Chapter 3

The dVR-fMRI exteroceptive pain dissociation

3.1 Introduction

Previous cognitive models assume pain processing as Bayesian inference to infer the causes of afferent sensory input, which is biased by predictive information with statistical optimisation [Yoshida et al., 2013, Anchisi and Zanon, 2015, Tabor and Burr, 2019, Seymour and Mancini, 2020]. This lies in two distinct categories of nociceptive causes: exteroception, extracted from the external environment for protective behaviours against avoidable potential threats [Sherrington, 1906, Price et al., 2003]; and interoception, obtained from the internal body as an unavoidable physiological condition along with homeostatic behaviours [Geldard, 1972, Wall, 1979, Craig, 2003, 2008]. Exteroceptive causes of pain (with stimuli arising from the outside world) involves a topographical map for discriminating spatial differences of nociception across body parts, encoded in somatosensory regions [Melzack and Wall, 1965, Melzack and Casey, 1968, Apkarian and Hodge, 1989, Mackel et al., 1992, Stevens et al., 1993, Craig, 2008, Mancini et al., 2012, 2014], with predictive pain encoding modulated by expectation and violation [Fardo et al., 2017]. Interoceptive causes of pain (with stimuli arising inside the body) subserves behavioural homeostatic functions in insula [Craig et al., 1994, Craig, 1995, 2002, 2003, Banich et al., 2010, Strigo and Craig, 2016], of which the predictive signal has been shown in anterior insula [Brown et al., 2008, Geuter et al., 2017, Fazeli and Büchel, 2018, Koppel et al., 2022, Hoskin and Talmi, 2023], dorsal putamen, and pregenual anterior cingulate cortex [Zhang et al., 2018]. These two pathways share bidirectional interactions with each other [Staud, 2012, Kennedy et al., 2016].

Neuroimaging studies in decision-making tasks have also found model-free signals correlating with unexpected pain in the putamen, whilst the combination of model-free and model-based signals correlating with unexpected pain in the ventromedial prefrontal cortex (vmPFC) [Lee et al., 2014, Wunderlich et al., 2012], in which they use FORWARD learning and BACKWARD planning for model-based RL whilst SARSA for model-free RL [Gläscher et al., 2010, Lee et al., 2014, Wang et al., 2018]. In addition, hippocampal regions such as medial entorhinal cortex (MEC) and hippocampus (HPC) are suggested to encode a 'cognitive map' of the external world for model-based learning during spatial navigation [Behrens et al., 2018], in which the successive representation is suggested for such model-based RL whilst SARSA for model-free RL [Dayan, 1993, Geerts et al., 2020, Stachenfeld et al., 2017].

In the previous study (Chapter 2), we have proposed a reinforcement learning (RL) model with multidimensional bodily pain input, which arbitrates between body-model and world-model pain avoidance behaviour. We further ask whether this conceptual difference is biologically reflected in the involvement of dissociable brain mechanisms. To do this we implemented an fMRI-based virtual-reality maze task, in which pain was associated either with a particular spatial location or a certain type of movement. We found that the model explained the data from the two conditions. Neuroimaging data revealed differentiable neural correlates of avoidance along a temporal-parietal axis, with a dorsolateral-medial prefrontal difference in world-model and body-model pain learning errors, sensorimotor and insula activations with pain-predictive signals, and an anterior cingulate circuit covering medial prefrontal and anterior striatal regions that correlated with reliability-based arbitration. Overall, our work suggests an arbitrative learning system in human brains that governs conceptually distinct pain representations relating to different exteroceptive contexts.

3.2 Methods

3.2.1 Desktop-based Virtual Reality (dVR) task design

A computer-based virtual spatial navigation task was created in Unity 2019.4 with a first-person view (Figure 2.1A), where participants were asked to avoid as many potential shocks as possible by conducting safer navigational actions that they needed to learn in the task. With Digitimer DS7A providing electric shocks, 2 electrodes were applied to the skin surface of both arms, with 1 for each side on the arm (Figure 2.1B, right).

The task consisted of a 3×3 square grid, with each square being an open room connected to neighbouring rooms with a narrow open aisle (Figure 2.1A, left). Participants were asked to learn safer navigational actions for receiving fewer shocks, subject to various pain conditions. During the 4-minute task in each block, participants continually moved forward at a constant velocity, with a maximum of 5.2 seconds moving from one room to the other. A trackball allows a 360-degree change of direction within the first-person view (Figure 2.1A, right). Participants used their right hands to roll the trackball to change the virtual facing direction, thus changing the virtual movement direction for conducting 1 of the 4 turning actions (i.e. left turn, right turn, forward move, and U-turn) in each room. A countdown was always visible in the top-right of the screen, showing how many seconds were left until the end of each block when navigating the virtual maze. A sun was visible in a specific direction of the sky for participants to locate themselves in the virtual maze. Participants could obtain invisible rewards once they enter the centre of a room where the invisible reward exists, which they would know from the screen, both the top-left reward calculator and the bottom reward progress bar (Figure 2.1A, right). They were encouraged to keep moving to obtain as many rewards as possible, as they would count for the final monetary outcome (10p for each) that participants obtained at the end of the task, which was independent of the shocks they might receive.

Depending on the movement and experimental condition, participants might receive a 40-millisecond electric shock on their arms as soon as they leave the room for a dangerous aisle (Figure 2.1B, left). Participants learnt with pre-instructed dangerous turns and regions with real-time visual stimuli during training rather than real shocks during the main session. Participants were always informed of which one of the 2 pain conditions (see below) it was at the start of each block. After each block, participants were asked to select which dangerous turning direction or region they had experienced in the corresponding block (i.e. post-block selection). They were told not to focus on finding out the exact danger but rather merely avoiding dangers, as actively finding it might bring more unnecessary shocks, given the task purpose of pain avoidance. They would not obtain any additional reward or punishment for answering the post-block selection questions.

3.2.1.1 Action-induced pain condition

For the action-induced pain condition with body-model pain, participants have an action contingency q among different actions to receive an electric shock, depending on the turning direction (i.e. turning left, turning right, moving forward, or a U-turn)

when leaving the room (Figure 2.1B, middle). Participants were told that 1 of the 4 turning actions could be more dangerous than the others, though either the left turn or the right turn could be dangerous, which was randomly counterbalanced across blocks. Participants have the conditioned probability $q_+ = 0.8$ to receive a shock when doing a dangerous turn, whilst having the conditioned probability $q_- = 0.3$ to receive a shock when not doing a dangerous turn, so that the contingency $q = q_+ - q_- = 0.5$. With a specific (i.e. either left or right) turn in the corresponding block, the ipsilateral body part may receive the shock; Otherwise, the contralateral body part may receive the shock.

3.2.1.2 State-induced pain condition

For the state-induced pain condition with world-model pain, participants had a state contingency p among different states to receive an electric shock in reality, depending on the position of dangerous places (Figure 2.1B, right). Here, invisible dangerous places composed a cluster of 4 adjacent aisles that connected a 2×2 region of rooms, which were randomly selected and counterbalanced across blocks. More specifically, participants had the conditioned probability $p_+ = 0.8$ to receive a shock when entering a dangerous aisle, whilst having the conditioned probability $p_- = 0.3$ to receive a shock when entering a non-dangerous aisle, so that the state contingency $p = p_+ - p_- = 0.5$. Only one part of the body might receive an electric shock if that part of the body was facing towards the dangerous centre in the virtual environment.

3.2.2 Neuroimaging (fMRI) experimental design

3.2.2.1 Participants

28 participants (10 females) attended the study after submitting informed consent. They were all pain-free, without neurological or psychiatric conditions, fluent English-speaking people aged 18-35 years (Mean: 30.56, SD: 7.08) with prior experience in video games, using their right hands to control the mouse, and not pregnant. Participants were screened to exclude any potential risks in the MRI scanner. The study was approved by the Central University Research Ethics Committee (CUREC; Approval Reference: R77882/RE003) of the University of Oxford.

3.2.2.2 Experimental procedures

Participants first took part in a training process in the preparation room, sitting comfortably at least 60 centimetres away from the computer screen to play the video

game. Detailed instructions were given to participants to help them get familiar with the navigation task. At the end of the training session, participants conducted two test blocks on different pain conditions to make sure they learn the task before moving to the scanner. Participants were asked to repeat the test blocks until they could provide the correct post-block selection in 2 recent blocks on different pain conditions. During the scan, participants experience two 20-minute scan repetitions (Figure 3.1A). Each session consisted of four 5-minute scan blocks, and each scan block consisted of a 4-minute navigational block and a following 1-minute interval for participants to report the pain rating and post-block selection. Participants had a minimum 1-minute break between each scan repetition, which could be extended if required by participants. Two pain conditions were counterbalanced within each scan repetition and between scan repetitions across all participants, which was to prevent the temporal effect when comparing between two pain conditions. Half of the participants had `World->Body->World->Body` pain condition as MRI scan repetition 1, whilst `Body->World->Body->World` pain condition in MRI scan repetition 2, the other participants conducted repetition 2 prior to repetition 1.

3.2.2.3 Pain ratings

Participants reported the intensity of electrical shock pain experienced on a VAS (visual analogue scale) in a 0–10 range (0: no pain, 4: start to feel pain, 10: worst pain imaginable) at the time of testing. Before the training session at the start of the experiment, participants experienced the pain calibration process, where the current intensity was increased from 0 mA in steps of 1-10 mA until the participant reported the VAS as 7. The calibration procedure was repeated three times, with the median current thresholds used for formal experimental sessions later. At the end of each experimental block, participants were also required to report VAS on the shock pain they had experienced. At the start of each MRI scan repetition, a re-calibration process is conducted to fine-tune the current intensity from the last setting until the participant reported the VAS as 7. This was to make sure that the subjective feeling of shock pain intensity was relatively stable across the whole experiment (Figure 3.1B), as we verified via the Walt test on reported VAS pain report over experimental blocks both on the left arm ($k = -.04, t(277) = -1.52, p = .130$) and on the right arm ($k = -.03, t(277) = -1.11, p = .267$).

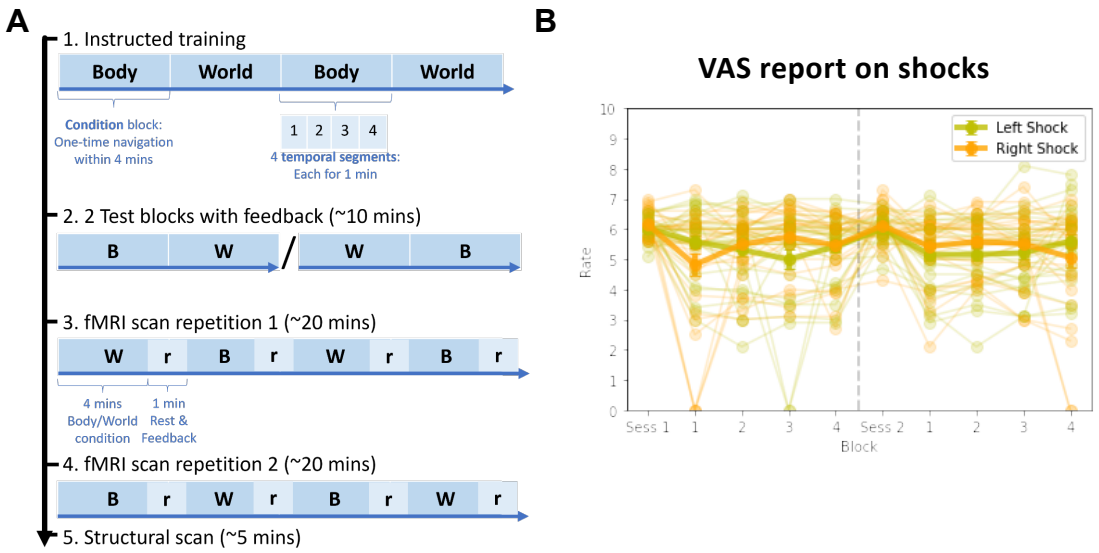

Figure 3.1: **Experiment protocol counterbalancing two pain conditions within and across MRI scan repetitions with stable pain rating reports.**

(A) Participants first conducted the instructed training in the preparation room, followed by two test blocks on different pain conditions to make sure they learn the task before moving to the scanner. During the scan, participants experienced two 20-minute scan repetitions. Each repetition consisted of four 5-minute scan blocks, and each scan block consisted of a 4-minute navigational block and a following 1-minute interval for participants to report the pain rating and post-block selection. Two pain conditions were counterbalanced within each scan repetition and between scan repetitions across all participants, so that half of the participants had the same fMRI session protocol in this figure, meanwhile the other participants conducted scan repetition 2 before scan repetition 1. **(B)** Participants reported stable VAS on painful shocks throughout the whole scan, which included 2 scan repetitions $\times$ 4 blocks. Participants reported VAS both at the start of each repetition as a recalibration process and at the end of each block after finishing the 4-minute virtual navigation. Dark yellow curves refer to the VAS report on the left arm, whilst shallow yellow curves refer to the VAS report on the right arm.

3.2.2.4 Electrical stimuli

Electrodes were applied to the volar surface of the participant's forearms both outside and inside the MRI scanner. Controlled current was applied only to the prepared surface area, without passing internally into the body. A Digitimer DS7A (Hertfordshire, UK) was used to elicit a low level of electrical output sufficient to induce a moderate-to-strong pain sensation calibrated to each participant. Participants received the output current (Mean: 141.4 mA, SD: 115.97 mA, max: 600 mA) with a series of 10 pulses in 40 milliseconds as a single shock (250 Hz, 2 ms/pulse) with a source voltage of 400 V.

3.2.2.5 MRI setup

Functional MRI imaging data were acquired on a 3T Siemens Magnetom Prisma scanner with a Siemens standard 32-channel phased array head coil. Functional images were collected using an interleaved MB3 Pat2 sequence (repetition time TR = 1,235 ms; echo time TE = 20 ms; field of view = 216 mm; flip angle = 65°), with 72 contiguous oblique-axial slices (2-mm isotropic voxels) parallel to the AC-PC line acquired. At the end of the scan, whole-brain high-resolution T1-weighted structural images were obtained using a single-shot ascending MPRAGE sequence (repetition time TR = 1,900 ms; echo time TE = 3.97 ms; field of view = 192 mm; flip angle = 8°; 1-mm isotropic voxels).

3.2.2.6 MRI preprocessing

The preprocessing of neuroimaging data was conducted via FSL 6 with standard procedures (Wellcome Centre for Integrative Neuroimaging, University of Oxford; `https://fsl.fmrib.ox.ac.uk`). With a 6-parameter rigid body transformation, motion artefacts were removed by realigning images to the first scan during the first session. The slice timing correction was then applied, followed by the co-registration with the corresponding structural image, and subsequently normalised into a MNI template space by resampling into 2-mm isotropic voxels. The functional images were then smoothed by a Gaussian kernel with full-width at half maximum (FWHM) of 8 mm.

3.2.3 Computational modelling

We used the same model as in the simulation task (Section 2.2.2, Figure 2.2) when $N_{body} = 2$.

3.2.3.1 Data fitting

The whole model with free parameters could thus fit behavioural data respectively under different pain conditions by maximising the likelihood $\Pi p(a|s)$ via Dual Annealing optimisation [Xiang et al., 1997] that searched the global maximum on targeted locations by combining the generalisation of Classical Simulated Annealing and Fast Simulated Annealing [Tsallis, 1988, Tsallis and Stariolo, 1996]. All free model parameters included:

- η_{Body}, γ_{Body}, and ν_{Body} as learning rates and discounts in the model-free RL;
- γ_{World}, η_{SR}, γ_{SR}, η_R, and ν_{World} as learning rates and discounts in the model-based RL;
- κ as temperature parameter in von Mises distribution;
- τ_ω, ρ_{Body}, and ρ_{World} as learning rates in TD-based weighted association;
- α_{Body}, α_{World}, β_{Body}, and β_{World} as tuning parameters in the weights of pain maps; and
- $\tau_\Pi^{(0)}$ (initial action temperature) and λ_Π (decay action temperature) as those controlling temperature parameters for the action with softmax policy.

For the limited number of subject actions in the real data, we further shrank the number of free parameters as follows (cf. Lee et al. [2014], Geerts et al. [2020]):

- $\gamma = \gamma_{Body} = \gamma_{World} = \gamma_{SR}$;
- $\nu = \nu_{Body} = \nu_{World}$;
- $\tau_\omega = \rho_{Body} = \rho_{World} = 1$;
- $\beta_{Body} = log(\frac{\alpha_{Body}}{\alpha_{B0}} - 1)$ where $\alpha_{Body} \geq 2\alpha_{B0}$;
- $\beta_{World} = log(\frac{\alpha_{World}}{\alpha_{W0}} - 1)$ where $\alpha_{World} \geq 2\alpha_{W0}$;
- $\alpha_{B0} = \alpha_{W0} = 0.1$; and
- $\lambda_\Pi = 0$ for no action temperature decay.

Thus, the free parameters (all of them are positive real numbers) used for data fitting were:

- γ as discount no larger than 1;
- η_{Body}, η_{SR}, η_R, and ν as learning rates that were all restricted to be lower than 1;
- κ as temperature parameter in von Mises distribution;
- $\alpha_{Body} \geq 0.2$ and $\alpha_{World} \geq 0.2$ as tuning parameters in the weights of pain maps; and
- $\tau_{\Pi}^{(0)}$ as action temperature for the action with softmax policy,

which resulted in the fitted model parameters (Table 2.1) also for the model simulation task (Section 2.2 & Chapter 2).

3.2.3.2 Parameter priors

To calculate the posterior $\Pi p(a|s)\Pi p_{parameters}$, we further assumed the following priors on the free parameters listed above:

- $\gamma \sim U(0,1)$ as the uniform distribution;
- $\eta_{Body}, \eta_{SR}, \eta_R, \nu \sim \Gamma(2,5)$ as the gamma distribution;
- $\alpha_{Body}, \alpha_{World} \sim \Gamma(2,1) + 0.2$ as the gamma distribution; and
- $\kappa, \tau_{\Pi}^{(0)} \sim N(2,1)$ as the normal distribution.

3.2.3.3 Alternative models

We also compared various alternative models based on whether the real-time bodily shock information (i.e. which body part gets shocked) and current spatial information (i.e. which room) were included in the Q-value learning. These included:

- `mf`: The pure model-free RL learning Q-value on body parts and actions $Q_{Body}(b,a)$ in Equation 2.2. $\omega_{World} \equiv 0$ was set to invalidate model-based learning;
- `mf_noB`: The pure model-free RL learning Q-value only on actions $Q_{Body}(a)$ as

$$Q_{Body}(a_t) \leftarrow Q_{Body}(a_t) + \eta_{Body}\delta_t^{Body}, \tag{3.1}$$

  so that the shock input was regarded as a single scalar $-r_{t+1}$ no matter which body part received the shock, rather than a vector $(-r_{t+1}(1), -r_{t+1}(2))$. $\omega_{World} \equiv 0$ was set to invalidate model-based learning;

- `mf_withS`: The pure model-free RL learning Q-value on body parts, spatial states, and actions $Q_{Body}(b, s, a)$ as

$$Q_{Body}(b, s_t, a_t) \leftarrow Q_{Body}(b, s_t, a_t) + \eta_{Body}\delta_t^{Body}(b), \tag{3.2}$$

  so that the model-free RL also considered the current spatial information input. $\omega_{World} \equiv 0$ was set to invalidate model-based learning;

- `mb`: The pure model-based RL learning Q-value on body parts, spatial states, and actions $Q_{World}(b, s, a)$ in Equation 2.4. $\omega_{Body} \equiv 0$ was set to invalidate model-free learning;

- `mb_noB`: The pure model-based RL learning Q-value on spatial states and actions $Q_{World}(s, a)$ as

$$Q_{World}(s_t, \hat{a}_t(a_t, h_t)) = -r_{t+1} + \gamma_{World}\mathbb{E}_{s_{t+1}|s_t,a_t}[\sum_{s'} M(s_{t+1}, s')R(s')], \tag{3.3}$$

  so that the shock input was regarded as a single scalar $-r_{t+1}$ no matter which body part received the shock, rather than a vector $(-r_{t+1}(1), -r_{t+1}(2))$. $\omega_{Body} \equiv 0$ was set to invalidate model-free learning;

- `arb`: The arbitrative RL learning on both $Q_{Body}(b, a)$ (`mf`) and $Q_{World}(b, s, a)$ (`mb`) in Equation 2.13;

- `arb_noMFB`: The arbitrative RL learning on both $Q_{Body}(a)$ (`mf_noB`) and $Q_{World}(b, s, a)$ (`mb`), so that the shock input to model-free learning was regarded as a single scalar $-r_{t+1}$ no matter which body part received the shock;

- `arb_noMBB`: The arbitrative RL learning on both $Q_{Body}(b, a)$ (`mf`) and $Q_{World}(s, a)$ (`mb_noB`), so that the shock input to model-based learning was regarded as a single scalar $-r_{t+1}$ no matter which body part received the shock; and

- `arb_withMFS`: The arbitrative RL learning on both $Q_{Body}(b, s, a)$ (`mf_withS`) and $Q_{World}(b, s, a)$ (`mb`), so that the model-free RL also considered the current spatial information input.

After fitting to human behavioural data, models were compared via Akaike information criterion (AIC; Akaike [1974]), Bayesian information criterion (BIC; Schwarz [1978]), and posterior exceedance probability [Stephan et al., 2009].

3.2.4 Neuroimaging data anaylsis

3.2.4.1 GLM design

In FSL 6, we used the general linear model (GLM) to generate voxelwise statistical parametric maps (SPM) from learning signal regressors. Subject to various regressing requirements, we provided multiple subject-specific design matrices on the corresponding regressors with population-free parameters:

- `BOLD activity in a specific phase`: (R1) regressors encoding the average BOLD response in the corresponding phase; and (R2) 6 motion regressors encoding movement displacement. Possible phases include: `Pain Onset` as a 50-millisecond time window after a shock is generated; `Pre-Room` as a 1-second time window before entering into the room; `Room Onset` during the movement in a room; and `Post-Room` as a 1-second time window after leaving the room.

- `Posterior exceedance probabilities`: (R1) regressors encoding the average BOLD response in the Post-room onset phase; (R2) real-time posterior exceedance probabilities favouring pure MF learning; (R3) real-time posterior exceedance probabilities favouring pure MB learning; and (R4) 6 motion regressors encoding movement displacement. The posterior exceedance probabilities were calculated over pure MF (`mf`), pure MB (`mb`), and arbitrative learning (`arb`) corresponding to each action.

- `Pure MF learning model parameters`: (R1) regressors encoding the average BOLD response in the Post-room onset phase; (R2) 2 regressors as the real-time temporal difference (TD) of model-free Q-value $\delta_t^{Body}(LEFT)$ and $\delta_t^{Body}(RIGHT)$ (corresponding to different electrodes) in the pure MF learning model (`mf`); and (R3) 6 motion regressors encoding movement displacement. Model variables were separated into different body parts as different regressors due to the asymmetric movement of both hands in reality (rather than in the virtual maze), as only the right hand was moving the trackball in the scanner.

- `Pure MB learning model parameters`: (R1) regressors encoding the average BOLD response in the Post-room onset phase; (R2) 2 regressors as the real-time TD of model-based R-value $\delta_t^{Pain}(LEFT)$ and $\delta_t^{Pain}(RIGHT)$ (corresponding to different electrodes) in the pure MB learning model (`mb`); (R3) real-time absolute mean TD of model-based M-value $\mathbb{E}_s[|\delta_t^{SR}(s)|]$ in the pure MB learning model (`mb`); and (R4) 6 motion regressors encoding movement displacement.

- `Arbitrative learning model parameters`: (R1) regressors encoding the average BOLD response in the Post-room onset phase; (R2) 2 regressors as the TD of model-free Q-value $\delta_t^{Body}(LEFT)$ and $\delta_t^{Body}(RIGHT)$ in the pure MF learning pathway (`mf`); (R3) 2 regressors as the real-time TD of model-based R-value $\delta_t^{Pain}(LEFT)$ and $\delta_t^{Pain}(RIGHT)$ in the MB learning pathway (`mb`); (R4) TD learning reliability scale $max(\chi_{Body}, \chi_{World})$ and reliability difference $\chi_{World} - \chi_{Body}$ in the arbitrative learning model (`arb`); (R5) real-time absolute mean TD of model-based M-value $\mathbb{E}_s[|\delta_t^{SR}(s)|]$ in the MB learning pathway (`mb`); and (R6) 6 motion regressors encoding movement displacement.

3.2.4.2 Multi-level analysis

All regressors were compiled into the corresponding GLM for the first-level analysis of individual participants and blocks. The results were then concatenated over different blocks into their corresponding 2 conditions in each participant's data. Resulting contrasts were used in third-level one-sample t-tests to consider population inference, and were compared between Body-model and World-model pain conditions as well as between each pain condition and 0. These activations were tested via the cluster correction with the whole-brain FWER (Family-Wise Error Rate) with corrected threshold ($z = 2.3, p = .05$) [Woo et al., 2014].

We also carried out ROI analysis on various specific brain regions via small volume correction performed within a 5-millimetre sphere [Poldrack, 2007], with resulting contrasts averaged from the second-level analysis to third-level one-sample t-tests. These brain regions included primary somatosensory cortex (SI) [Fardo et al., 2017], prefrontal cortex [Lee et al., 2014], cingulate cortex [Zhang et al., 2018], insular cortex [Wunderlich et al., 2011, Geuter et al., 2017], thalamus [Craig et al., 1994], amygdala [Neugebauer, 2015], and caudate [Wunderlich et al., 2011], etc.

3.3 Results

We included data from $N = 28$ participants for various data analyses. The behavioural data did not show any significantly biased laterality on the number of left/right arm shocks ($\Delta = -.22, t(27) = -.79, p = .434$; Figure 3.2A) or the number of left/right turning actions ($\Delta = .98, t(27) = .88, p = .380$; Figure 3.2B) in each block. These results suggested that no significant laterality bias on bodily shock or bodily turn was detected during the experiment.

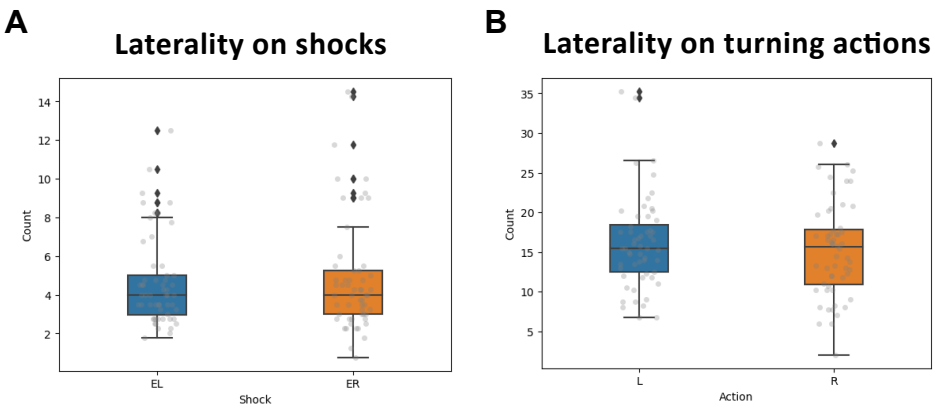

Figure 3.2: **No significant laterality bias on bodily shock or bodily turn detected during the navigational task from behavioural data.**

(A) Box plot on the average number of shocks received on the left arm ('EL' as the left bar) or right arm ('ER' as the right bar) in each block. **(B)** Box plot on the average number of bodily turns conducted as a left turn ('L' as the left bar) or right turn ('R' as the right bar) in each block. Black square dots stand for extreme values beyond 1.5 IQR (interquartile range boxes, as the percentile from 25th to 75th).

For the post-block selection, participants were asked to select the dangerous turning direction or region they had experienced after each block in formal sessions, so that each participant answered 8 corresponding questions in total during the whole scan. All 28 participants gave an overall post-block accuracy of 78.12% (SD: 20.76%), although they were told not to focus on the correct answer to these post-block selection questions, but rather only on shock avoidance. Participants reported a mean higher post-block selection accuracy in action-induced than state-induced conditions ($\Delta = 24.11\%, t(27) = 4.93, p < .001$), suggesting higher difficulty in learning world-model pain in our task with spatial inconsistency between virtual navigation and participants lying still in the MRI machine.

3.3.1 Human behavioural data showed considerable pain avoidance performance

Participants conducted aversive learning on the task with pain avoidance performance over time. We defined pain avoidance performance (PAP) as the action that led to a lower possibility of receiving shocks when leaving the room. Each action could be a left turn, a right turn, moving forward, or moving backwards. For the action-induced condition, PAP was the other three turning directions apart from the dangerous one (Figure 3.3A shows a dangerous right turn as an example). For the state-induced condition, PAP was the turning direction that did not directly lead to a dangerous

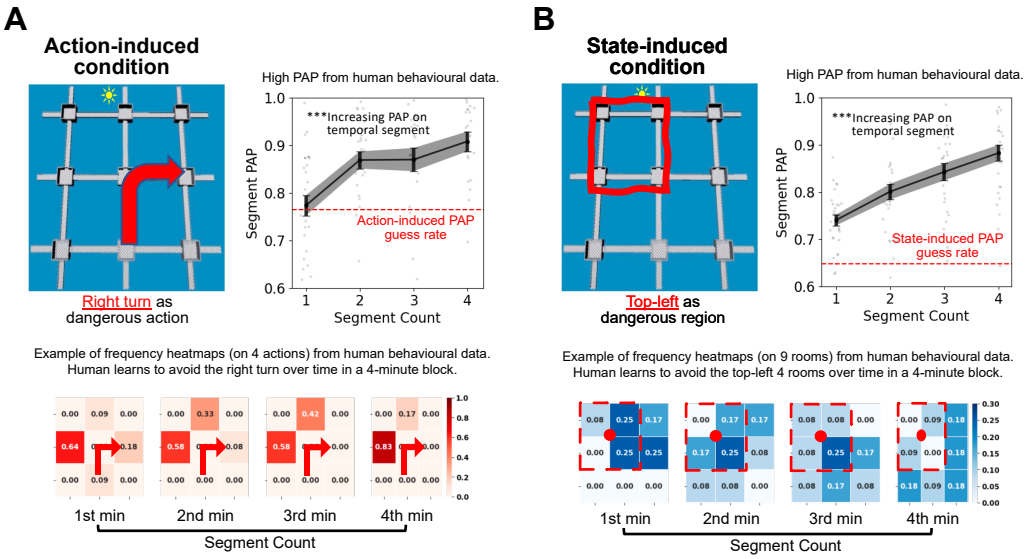

Figure 3.3: **Behavioural data analysis showed significant aversive learning on pain avoidance performance (PAP) in both pain conditions.**

(A) An example of an action-induced condition with the dangerous right turn (big red arrow on top-left), with visit frequency heatmaps in each minute on 4 actions in the action-induced condition (bottom). We showed a significant temporal increase of PAP over all participants (top-right) in the action-induced condition. **(B)** An example of a state-induced condition with the dangerous front-left region (red square on top-left corresponding to the visible 'yellow' sun), with visit frequency heatmaps in each minute on 9 rooms in the state-induced condition (bottom). We showed a significant temporal increase of PAP over all participants (top-right) in the state-induced condition. Error bars stand for 95% CI. *** for $p < .001$.

aisle (Figure 3.3B shows a dangerous top-left region as an example). To show the temporal learning progress at the behavioural level, we calculated PAP as a percentage over all actions separately within 4 temporal segments, each for the corresponding minute in a 4-minute block (see Figure 3.3A). Chance levels were then regarded as a pure action guess strategy calculated over all pairs of room and action. That is, $52/68 = 76.47\%$ for action-induced condition (Figure 3.3A with the red dashed line as action-induced PAP guess rate) and $44/68 = 64.71\%$ for state-induced condition (Figure 3.3B with the red dashed line as state-induced PAP guess rate). For each pain condition, the bottom row of Figure 3.3 provide the visualised example of visit frequency heatmaps to show how participants learnt to avoid dangerous actions or regions on each 1-minute segment in a single block, with deeper colours referring to a higher visit frequency which tends to avoid the danger in later segments.

We conducted a two-way repeated-measure ANOVA on PAP over all $N = 28$ participants and 4 blocks, with pain condition (2 conditions with action-induced

and state-induced) and temporal segment (4 segments from the 1st to 4th minute) as two main factors. A Huynh-Feldt correction was applied to correct sphericity [Mauchly, 1940, Huynh and Feldt, 1976], as the segment term and the interaction term failed to pass Mauchly's sphericity test (segment: Mauchly's $W = .84$, Huynh-Feldt's $\varepsilon = .94, df = 5, p = .002$; condition $\times$ segment: Mauchly's $W = .75$, Huynh-Feldt's $\varepsilon = .85, df = 5, p < .001$). We found significant main effects on both pain condition ($F(1, 111) = 9.17, p = .003, \eta_p^2 = .08$) and temporal segment ($F(3, 311.55) = 40.79, p < .001, \eta_p^2 = .27$), whilst no significant interaction effect ($F(3, 281.80) = 1.40, p = .246, \eta_p^2 = .01$). These results suggested that participants learnt aversive performance over time, which was more efficient in action-induced condition than in state-induced condition due to the naturally higher difficulty of learning state-induced pain with a lower guess rate than the action-induced pain.

We also found the significant temporal increase of PAP averaged on all 4 blocks overall all participants in both action-induced condition (Figure 3.3A, top-right; $k = 4.05\%, t(110) = 4.16, p < .001$) and state-induced condition (Figure 3.3B, top-right; $k = 4.69\%, t(110) = 6.48, p < .001$), as well as over both conditions ($k = 4.37\%, t(222) = 7.12, p < .001$). Moreover, PAP after the early learning stage significantly exceeded the chance level (i.e. pure guess rate of the corresponding pain condition). This could be seen in action-induced condition (Figure 3.3A, top-right) at the 2nd minute ($\Delta = 10.41\%, t(27) = 6.38, p < .001$), 3rd minute ($\Delta = 10.51\%, t(27) = 5.49, p < .001$), and 4th minute ($\Delta = 14.33\%, t(27) = 8.53, p < .001$), as well as in state-induced condition (Figure 3.3B, top-right) at the 1st minute ($\Delta = 9.30\%, t(27) = 7.80, p < .001$), 2nd minute ($\Delta = 15.35\%, t(27) = 10.47, p < .001$), 3rd minute ($\Delta = 19.54\%, t(27) = 14.33, p < .001$), and 4th minute ($\Delta = 23.53\%, t(27) = 17.87, p < .001$). These results again suggested that participants conduct effective cognitive aversive learning over time under both action-induced and state-induced conditions.

To find out whether the behavioural performance could be best described by the corresponding action-induced or state-induced pain condition, we also conducted Bayesian model selection (with Bayes factor; Kass and Raftery [1995]) over three ideal action strategies: guess, body-model, and world-model. The guessing strategy had a uniform action probability in all turn directions when leaving a certain room. The body-model and world-model strategies had a different action probability on PAP (corresponding to the action-induced/state-induced condition) compared to the uniform action probability on all other turn directions, of which the ratio was equal to the reversed ratio of shocking probability, i.e. $p_- : p_+ = q_- : q_+ = 0.3 : 0.8$.

Table 3.1: Bayes factors (BF) of avoidance actions over temporal segments showed learnt action strategy in the corresponding pain condition.

		BF in each segment			
Condition	Strategies	1st min	2nd min	3rd min	4th min
Action-induced	B over C	9.6×10^{-47}	3.3×10^{-8}	1.3×10^{-8}	$3.4 \times 10^{10}*$
	B over W	5.7×10^{-27}	$2.4 \times 10^{14}*$	$1.8 \times 10^{9}*$	$6.1 \times 10^{17}*$
State-induced	W over C	5.5×10^{-2}	5.9×10^{0}	$6.7 \times 10^{12}*$	$5.5 \times 10^{16}*$
	W over B	$1.1 \times 10^{98}*$	$3.0 \times 10^{96}*$	$6.2 \times 10^{96}*$	$6.0 \times 10^{117}*$

*Decisive evidence with BF > 100.
B: Body-model; W: World-model; C: Chance level.

Given the same prior among three action strategies, Table 3.1 shows the Bayes factors (BF) in 4 temporal segments, where likelihoods were calculated across all participants and blocks within the same condition. These Bayes factors were from the model comparison for: 1) body-model action over chance level within action-induced condition, 2) body-model action over world-model action strategy within action-induced condition, 3) world-model action over chance level within state-induced condition, and 4) world-model action over body-model action strategy within state-induced condition. All of these comparisons showed decisive evidence (BF > 100) at the later stage of navigation, suggesting that participants successfully learnt the correct corresponding action strategy when conducting the task within each condition.

3.3.2 Model comparison with temporal-posterior difference in explaining neuroimaging data

Based on the fitted model parameters (Figure 3.4), we compared alternative models under different pain conditions to see which model best accounts for the corresponding avoidance behaviour. Table 3.2 shows their AIC [Akaike, 1974] and BIC [Schwarz, 1978] under different pain conditions. Models with body part variable b in their Q-value function always outperformed those without b, suggesting the necessity of multidimensional bodily inputs for efficient pain learning rather than regarding pain as a single negative scalar. Furthermore, pure MF learning fitted better without having spatial information s in the Q-value, indicating that efficient body-model pain learning might be independent of external spatial knowledge, which was handled by world-model pain learning.

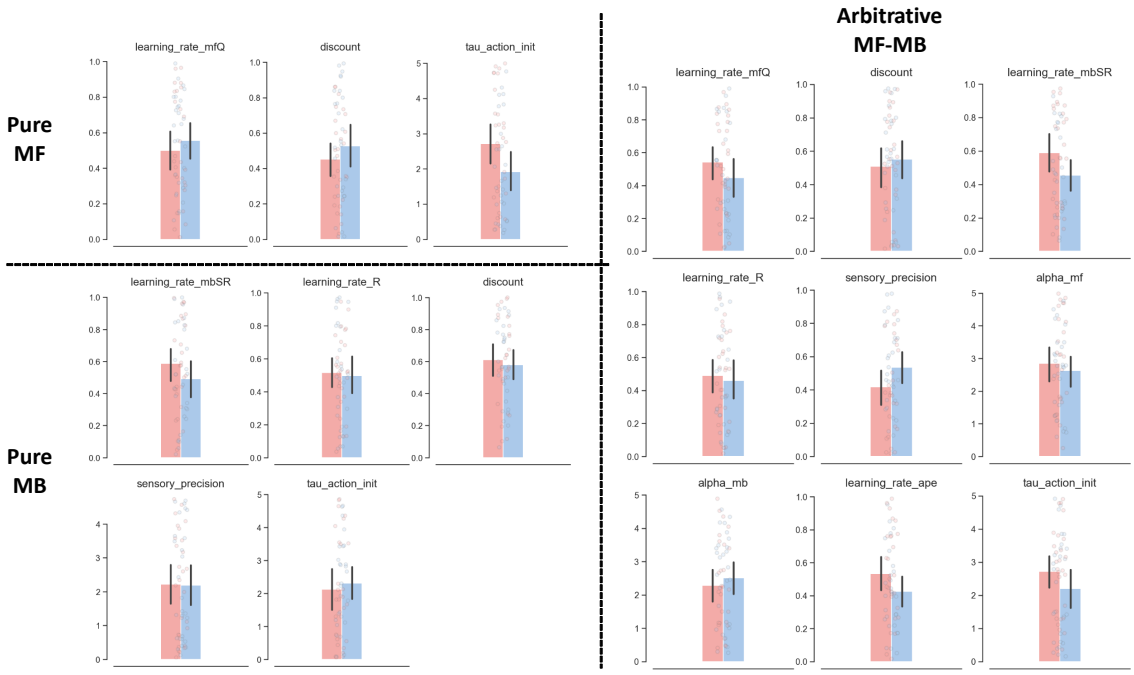

Figure 3.4: **Fitted model parameters did not show significant difference between action-induced and state-induced conditions.**

Bar plots of model parameters from data fitting under action-induced condition (red bars) and state-induced condition (blue bars) on pure MF, pure MB, and arbitrative learning model. Red bars for action-induced condition and blue bars for state-induced condition. See Section 2.2.2 for the corresponding abbreviations of each model and parameter. Error bars stand for 95% CI.

Table 3.2: Model comparisons among alternative models under different pain conditions, indicating that the pain learning strategy was fitted to the behavioural data subject to the specific environmental condition.

		BIC		AIC	
Model*	Q-function^{+}	Body	World	Boby	World
`mf`	$Q_{Body}(b, a)$	11954.9586	14443.9597	11693.7135	14182.9845
`mf_noB`	$Q_{Body}(a)$	14973.1418	20842.5906	14711.8959	20581.6161
`mf_withS`	$Q_{Body}(b, s, a)$	18150.3920	19449.8428	17889.1467	19188.8679
`mb`	$Q_{World}(b, s, a)$	12942.7218	13777.4575	12507.3125	13342.4995
`mb_noB`	$Q_{World}(s, a)$	17515.2005	19325.0055	17166.8726	18977.0385
`arb`	$Q(b, s, a)$	19108.5980	20415.2238	18324.8610	19632.2992
`arb_noMFB`	$Q_{Body}(a)$	31627.1482	34262.0889	30843.4116	33479.1644
`arb_noMBB`	$Q_{World}(s, a)$	31486.0664	34121.0975	30789.4116	33425.1644
`arb_withMFS`	$Q_{Body}(b, s, a)$	18270.2806	20600.5454	17515.7197	19817.6207

Models highlighted with light grey were used for neuroimaging analysis.
*See Section 2.2.2 for model details.
^{+}b: Body part; s: Spatial state; a: Bodily action.
Body: Action-induced condition with body-model pain; W: State-induced condition with world-model pain.

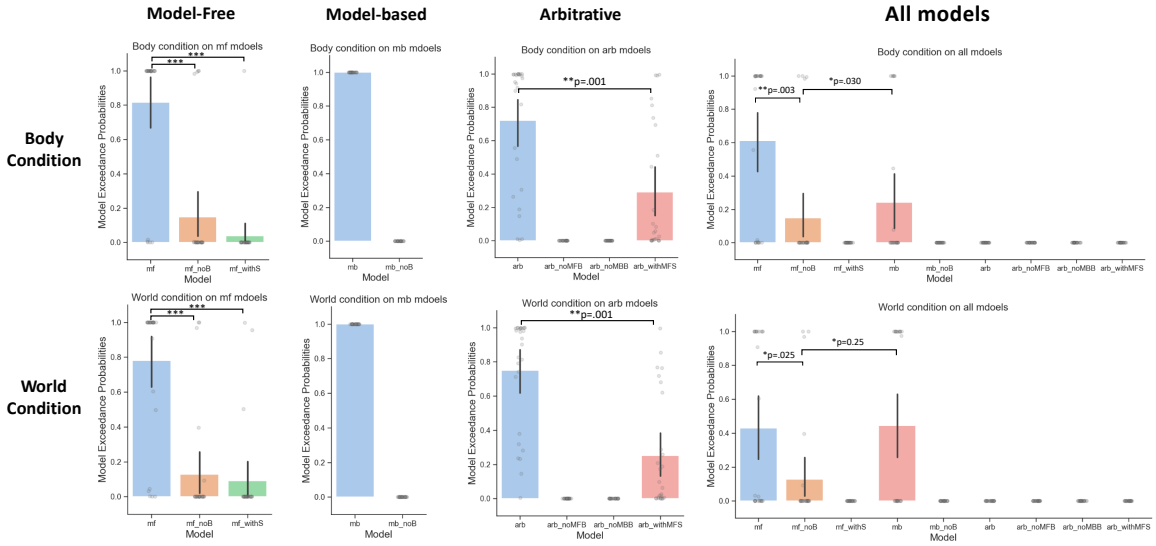

Figure 3.5: **Data fitting comparisons among alternative models showed the pure MF or MB learning in favour of the corresponding pain condition performance.**

The posterior exceedance probabilities were calculated under different model families (columns) on the behavioural data from respective pain conditions (rows). The pure MF, pure MB, and Arb learning models best described the avoidance performance within their own model families, whilst pure MF best described the body-model pain avoidance over all of the other alternative models. *** for $p < .001$, ** for $p < .01$, and * for $p < .05$. Error bars stand for standard deviation. Error bars stand for 95% CI.

We noticed that pure MF learning fitted better than pure MB learning in the body conditions under Wilcoxon signed-rank test [Wilcoxon, 1945] on both AIC ($T = 91, p = .019$) and BIC ($T = 88, p = .015$), whilst pure MB learning generally fitted better in the world conditions (Table 3.2). This suggested the corresponding pain learning strategy was fitted to the behavioural data, subject to the specific environmental condition. The arbitration model, on the other hand, did not fit well compared to pure MF/MB learning models, which might be because participants were always informed of the current pain conditions when the block started, so that they could learn and perform the task well within the restricted scan time. However, we still included results from the arbitrative model for neuroimaging analysis later, as it might capture brain activities relative to cognitive arbitration that the pure MF and MB learning models might fail to implement.

Using posterior exceedance probabilities (EP) [Stephan et al., 2009] also showed similar results (Figure 3.5A). Pure MF learning with bodily inputs and without spatial inputs described the most avoidance behaviour within the MF modelling family both in

action-induced condition (`mf-mf_noB`: $\Delta = .67, t(26) = 4.76, p < .001$; `mf-mf_withS`: $\Delta = .78, t(26) = 8.01, p < .001$) and state-induced condition (`mf-mf_noB`: $\Delta = .65, t(26) = 5.14, p < .001$; `mf-mf_withS`: $\Delta = .69, t(26) = 6.05, p < .001$) (Figure 3.5, first column), whilst pure MB learning with bodily inputs and spatial inputs described the most avoidance behaviour within the MB modelling family in both conditions ($EP_{mb\_noB} \equiv 0$) (Figure 3.5, second column). Despite not being directly shown in the AIC and BIC comparison, the arbitrative model without model-free spatial input described more avoidance behaviour than that includes model-free spatial input both in action-induced condition (`arb-arb_withMFS`: $\Delta = .48, t(26) = 3.90, p = .001$) and in state-induced condition (`arb-arb_withMFS`: $\Delta = .50, t(26) = 3.92, p = .001$) (Figure 3.5, third column), implying the independence between body-model pain learning and spatial knowledge under the single arbitrative system when the latter was handled by the other learning pathway (i.e. world-model pain learning). When calculating exceedance probabilities over all alternative models (Figure 3.5, fourth column), pure MF learning described the most avoidance behaviour in the body condition (`mf-mf_noB`: $\Delta = .046, t(26) = 3.25, p = .003$; `mf-mb`: $\Delta = .37, t(26) = 2.30, p = .030$; others: $EP \equiv 0$), whilst pure MB learning described one of the best in the world condition (`mf-mf_noB`: $\Delta = .30, t(26) = 2.39, p = .025$; `mb-mf_noB`: $\Delta = .32, t(26) = 2.38, p = .025$; `mb-mf`: $\Delta = .01, t(26) = .08, p = .936$; others: $EP \equiv 0$). Pure MB learning did not show significant exceedance over pure MF learning in world conditions, which might be because of higher difficulty in state-induced conditions for most participants. Along with the previous information criteria comparison, these altogether showed that the corresponding pain learning models could best describe the behavioural data within the respective environmental conditions for pain learning.

We also showed the neural correlates of the real-time posterior exceedance probabilities after data fitting over three models (those highlighted by light grey in Table 3.2): pure MF, pure MB, and the arbitrative model, to see any potential brain activity difference between two pain conditions when favouring a particular learning strategy (Figure 3.6). For EP on pure MF learning, the left superior temporal region was more activated in the body condition than the world condition ($z_{Peak} = 4.40, p = .019$, FWER cluster-level corrected), and the left primary motor cortex correlated the pure model-free EP in the body condition ($z_{Peak} = 3.98, p = .032$, FWER cluster-level corrected). On the other hand, for EP on pure MB learning, the precuneous was more activated in the world condition than the body condition ($z_{Peak} = 4.16, p = .004$, FWER cluster-level corrected), the region of which also covered the posterior cingulate cortex (PCC;$(-2, -48, 18)$, Sphere ROI $r = 5$ mm, $p = .022$) and left primary

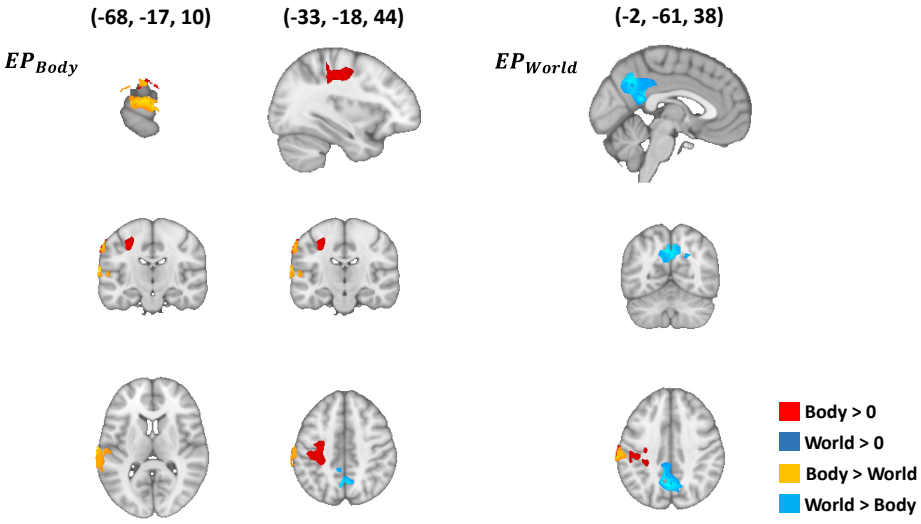

Figure 3.6: **Voxel-wise real-time model exceedance probabilities (EP) showed a temporal-parietal difference in favour of MF-MB pain learning strategies.** For EP on pure MF learning (left & yellow), the superior temporal region was more activated in the body condition than the world condition, whereas for EP on pure MB learning (right column & shallow blue), the parietal region was more activated in the world condition than the body condition. The first row shows the current MNI152 coordinates of the corresponding column. All shown activated regions were FWER cluster-level corrected ($z = 2.3, p = .05$).

somatosensory cortex (SI;$(-40, -25, 43)$, Sphere ROI $r = 5$ mm, $p = .031$). These implied a regional difference when using different pain learning strategies to describe brain activities, where the temporal region preferred MF learning whilst the posterior region preferred MB learning.

3.3.3 Phase-based neural correlation showed parietal and somatosensory difference between conditions

We first showed neural correlates of any specific phases during the virtual maze navigation when learning and avoiding shocks (Figure 3.7). These navigational phases included: `Pain Onset` as a 50-millisecond time window after a shock is generated; `Pre-Room` as a 1-second time window before entering into the room; `Room Onset` during the movement in a room; and `Post-Room` as a 1-second time window after leaving the room. Notice only the `Pain Onset` phase was attached to a shock existence, whilst the other three phases were attached to every room visit, no matter whether a shock is received.

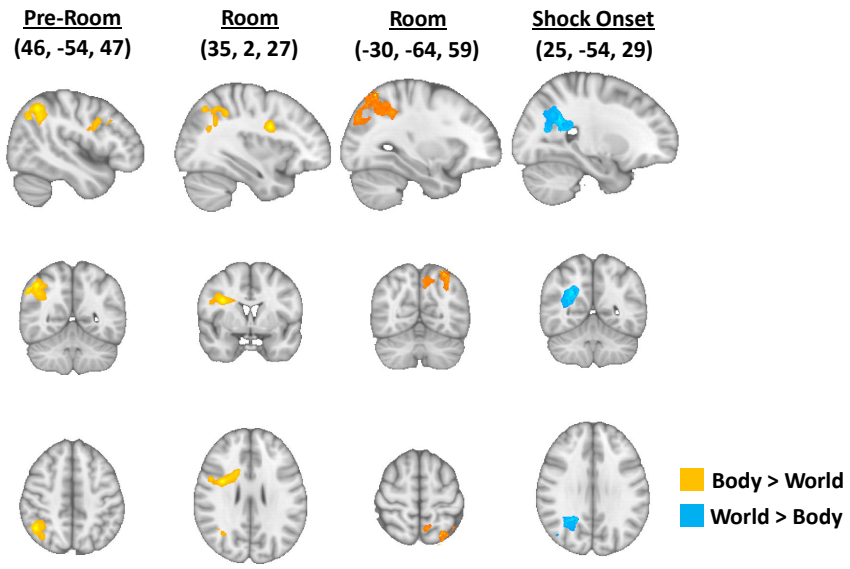

Figure 3.7: **The phase-based neural correlation showed a parietal difference between body-model and world-model conditions in various navigational phases, as well as a stronger motor-cortex activation during the phase that might be predicting the incoming body-model pain.**

The angular cortex and primary motor cortex correlated more with the body-model than world-model condition in the Pre-Room phase. The precuneus cortex correlated more with the world-model than the body-model condition when receiving the shock. The lateral occipital cortex correlated more with the body-model than the world-model condition in the room onset phase. The precuneus cortex correlated more with the world-model than the body-model condition in the shock onset phase. The first row indicates the corresponding navigational phases, and the second row shows the current MNI152 coordinates of the corresponding column. All shown activated regions were FWER cluster-level corrected ($p = .05$).

When comparing between the body-model and world-model conditions, the precuneus cortex showed a stronger correlation with the world-model than the body-model condition during the shock onset ($z_{Peak} = 3.98, p = .001$, FWER cluster-level corrected), indicating a shock-dependent parietal difference in pain perception as learning the world-model pain required a spatial knowledge of the external environment. Furthermore, we showed a stronger neural correlation in the body-model condition over the angular gyrus ($z_{Peak} = 3.60, p = .016$, FWER cluster-level corrected) before entering the room, the primary motor cortex (MI; $z_{Peak} = 3.54, p = .040$, FWER cluster-level corrected) before entering the room, and the lateral occipital cortex ($z_{Peak} = 5.51, p = .002$, FWER cluster-level corrected) when staying in the room. Besides the shock-independent parietal (angular gyrus) difference when making turning actions and predicting shocks, here we also showed a stronger MI activation during the phase that might be predicting the incoming body-model pain than world-model pain, which fitted the 'body-model' theory that the somatosensory-motor regions might be responsible for both body-level representation and prediction of pain for learning its avoidance behaviour [Brecht, 2017]. More evidence on both the shock-dependent and shock-independent neural differences would be given in the following neural correlation findings on data-fitted real-time model variables.

3.3.4 Real-time model variables indicated TD-driven learning for the corresponding condition

After behavioural data fitting, the model could provide real-time variables during the task, which we epoched into 4 segments (1 minute for each) to see their temporal tendencies over learning in two conditions (Figure 3.8). Two-way repeated ANOVA corrected by positive false discovery rates (FDR) with Benjamini-Hochberg procedure [Benjamini and Hochberg, 1995] showed significant segment main effect as temporal decrease in the absolute value of TD prediction errors (Pure MF $|\delta_t^{Body}(RIGHT)|, F(3, 81) = 4.97, p = .006$; Pure MB $|\delta_t^{SR}|, F(3, 81) = 22.81, p < .001$; Arb. $|\delta_t^{Body}(RIGHT)|, F(3, 81) = 4.34, p = .001$; Arb. $|\delta_t^{SR}|, F(3, 81) = 22.81, p < .001$), indicating pain avoidance learning with decreasing TD prediction errors over time. In detail, MF learning performed TD decreasing in action-induced condition (Pure MF $|\delta_t^{Body}(LEFT)|, r = -.203, p = .063$; Pure MF $|\delta_t^{Body}(RIGHT)|, r = -.344, p < .001$; Arb. MF $|\delta_t^{Body}(LEFT)|, r = -.206, p = .058$; Arb. MF $|\delta_t^{Body}(RIGHT)|, r = -.328, p = .001$), whist MB learning performed significant state-based TD decreasing in both action-induced condition (Pure MB $|\delta_t^{SR}|, r = -.411, p < .001$; Arb. MB $|\delta_t^{SR}|, r = -.411, p < .001$) and state-induced condition (Pure MB $|\delta_t^{SR}|, r =$

$-.324, p < .001$; Arb. MB $|\delta_t^{SR}|, r = -.324, p < .001$), suggesting a TD-decreasing MF/MB learning for the corresponding action-/state-induced condition.

For other variables attached to the arbitrative learning, we found significant interaction effects between condition and segment (Arb. rate: $\omega_t, F(3, 81) = 9.67, p < .001$; MF Reliability: $\chi_{Body}, F(3, 81) = 10.21, p < .001$; Reliability scale $max(\chi_{Body}, \chi_{World})$, $F(3, 81) = 14.55, p < .001$; Reliability difference: $\chi_{World} - \chi_{Body}, F(3, 81) = 11.10, p < .001$). These all revealed the tendency of a learning pathway outperforming the other over arbitration when learning the corresponding condition of exteroceptive pain, although MB learning also raised its reliability in learning MF-preferred action-induced pain ($\chi_{World}, r = .411, p < .001$) due to its decreasing predictive TD error.

3.3.5 Supplementary motor and medial prefrontal functionalisation on model-free learning

We then sought neural correlates on data-fitted real-time model variables during virtual navigation as the model-oriented statistical parametric maps (SPM). For the model-free learning variables, we conducted two independent GLMs on the pure MF learning model and the MF learning pathway of the arbitrative model (Figure 3.9). Regressors included the real-time temporal difference (TD) of model-free Q-value $\delta_t^{Body}(LEFT)$ and $\delta_t^{Body}(RIGHT)$ corresponding to the shocks onto different arms, as participants only used the right hand to control the trackball whilst keeping the left hand still during the scan. Two GLMs were conducted due to the potential information lost in the pure MF model from cognitive arbitration, even if the pure MF model fitted the behaviour better as the experimental pain condition was always informed before task blocks.

Under the pure MF learning model (Figure 3.9A), the prediction TD error on the left side correlated with body-model pain condition in the primary motor cortex (MI;$z_{Peak} = 4.33, p = .008$, FWER cluster-level corrected), whilst correlating with world-model pain condition in the supplementary motor area (MII;$z_{Peak} = 3.48, p = .010$, FWER cluster-level corrected). This showed a potential functional dualisation in motor regions between body-model and world-model MF learning, particularly corresponding to the still hand. On the other hand, the prediction TD error on the right side correlated more with body condition than world condition in left lateral occipital cortex (lOC;$z_{Peak} = 3.68, p = .023$, FWER cluster-level corrected) given pure MF learning model, and in medial prefrontal cortex (mPFC;$z_{Peak} = 4.43, p < .001$, FWER cluster-level corrected) given the arbitration model (Figure 3.9B). This difference might be due to a potential arbitrative mechanism that the pure MF learning model

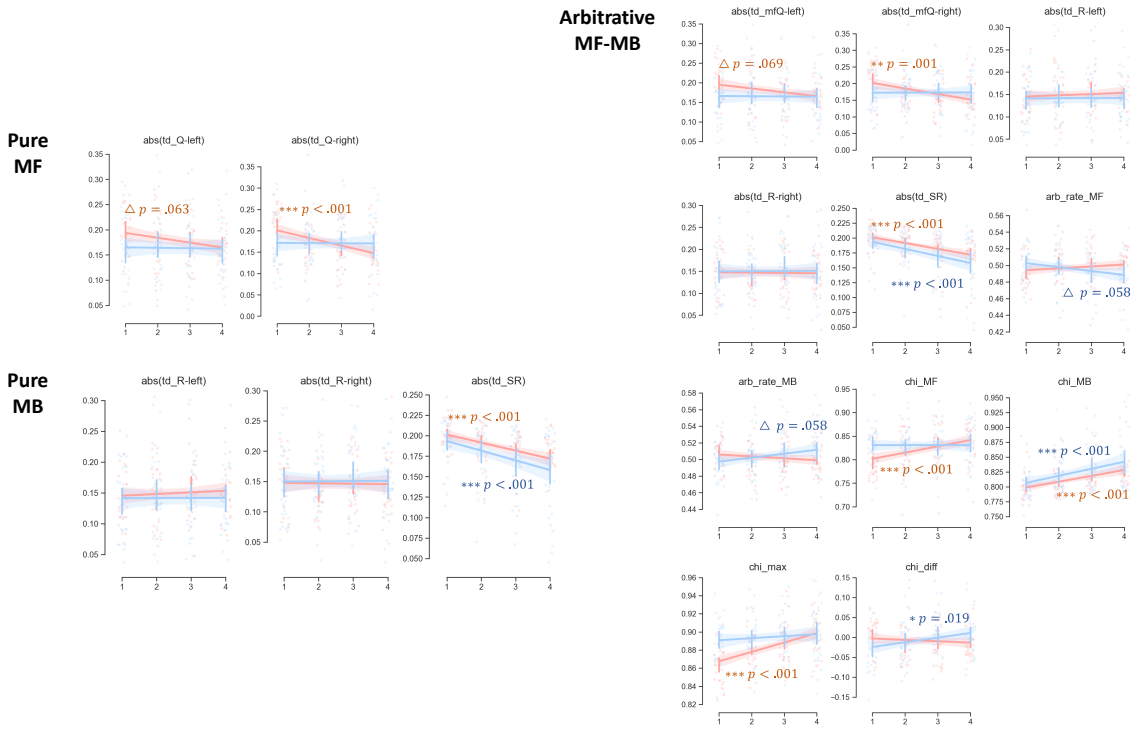

Figure 3.8: **Data-fitted real-time model variables showed predictive TD-driven learning with strong interaction effect as arbitration between condition and temporal segment.**

Scatter and regression plots of model variables from data fitting were shown under action-induced condition (red) and state-induced condition (blue) on pure MF (top-left), pure MB (bottom-left), and arbitrative learning model (right). These indicated a TD-decreasing MF/MB learning for the corresponding action-/state-induced condition, as well as the increased MB learning reliability even in learning MF-preferred action-induced pain. See Section 2.2.2 for the corresponding abbreviations of each model and variable. 'abs' means achieving the absolute value of temporal difference (i.e. 'td'). Error bars stand for 95% CI. FDR-corrected p-values with different colours show the result of the Wald Test in the corresponding condition with t-distribution against the zero-slope null hypothesis. *** for $p < .001$, ** for $p < .01$, * for $p < .05$, and Δ for $p < .1$.

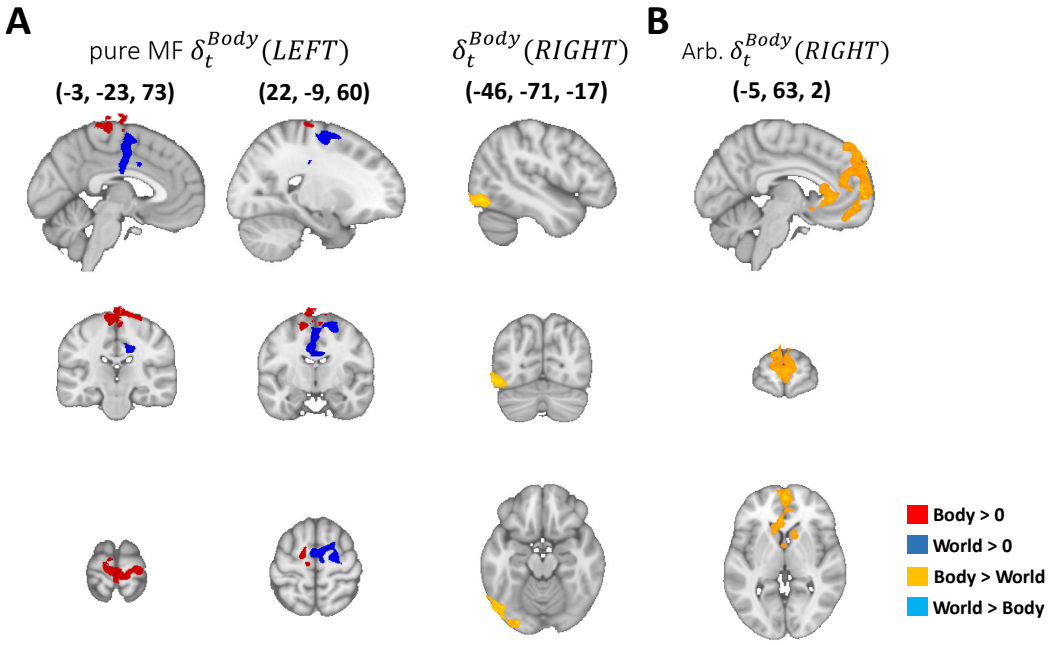

Figure 3.9: **The model-free temporal difference learning showed functional dualisation in motor cortex between body-model and world-model pain learning from the left arm, as well as medial prefrontal body-model pain learning from the right arm.**

(A) Under the pure MF learning model, the prediction TD error on the left side correlated with body-model pain condition in the primary motor cortex, whilst correlating with world-model pain condition in the supplementary motor cortex. The prediction TD error on the right side correlated more with body condition than world condition in the lateral occipital cortex, given a pure MF learning model. **(B)** Given the arbitrative learning model, the prediction TD error on the right side correlated more with body condition than world condition in the medial prefrontal cortex. The first row shows the current MNI152 coordinates of the corresponding column. All shown activated regions were FWER cluster-level corrected ($p = .05$).

failed to capture, so that the TD error of body-model pain learning correlated with the medial-prefrontal activity.

3.3.6 Model-based learning substrate ranging from prefrontal to cingulate cortical regions

For the model-based learning variables, we conducted two independent GLMs on the pure MB learning model and the MB learning pathway of the arbitrative model (Figure 3.10). Regressors included the real-time TD of model-based R-value $\delta_t^{Pain}(LEFT)$ and $\delta_t^{Pain}(RIGHT)$ corresponding to the shocks onto different arms, as well as real-time absolute mean TD of model-based M-value $\mathbb{E}_s[|\delta_t^{SR}(s)|]$ over 9 rooms. Two GLMs were conducted due to the potential information lost in the pure MB model from cognitive arbitration, even if the pure MB model fitted the behaviour better as the experimental pain condition was always informed before task blocks.

Under the pure MB learning model (Figure 3.10A), the state pain TD error on the left side correlated with world-model pain condition in the primary motor cortex (MI; $z_{Peak} = 3.81, p = .002$, FWER cluster-level corrected) and right putamen ($z_{Peak} = 3.49, p = .038$, FWER cluster-level corrected) covering right anterior insula cortex ($(30, 18, 10)$, Sphere ROI $r = 5$ mm$, p = .009$), whilst correlating more world-model pain than body-model pain condition in the superior frontal cortex (sFC; $z_{Peak} = 3.51, p < .001$, FWER cluster-level corrected) covering right dorsolateral prefrontal cortex (dlPFC; $(30, 40, 30)$, Sphere ROI $r = 5$ mm$, p = .011$). Under the arbitrative learning model (Figure 3.10B), on the other hand, the state pain TD error on the left side correlated with world-model pain more than body-model pain condition covering a high cluster including right lateral thalamus ($(23, -27, 4)$, Sphere ROI $r = 5$ mm$, p = .009$), right primary somatosensory cortex (SI; $(35, -34, 48)$, Sphere ROI $r = 5$ mm$, p = .006$), right posterior insula ($(35, -12, 13)$, Sphere ROI $r = 5$ mm$, p = .027$), etc. These suggested a broad range of neural substrates related to the state pain TD learning that is independent of spatial information.

We also noticed the spatial successor representation's absolute mean TD error showed a higher correlation with body-model pain than world-model pain in the right inferior frontal gyrus pars opercularis ($z_{Peak} = 3.49, p = .025$, FWER cluster-level corrected; Figure 3.10A, right) only in the pure MB learning model. This might be due to the loss of a potential arbitrative mechanism in the pure MB learning model, resulting in a worse successor representation learning (thus with larger absolute TD error) on merely spatial information in the 'contradictive' body-model condition.

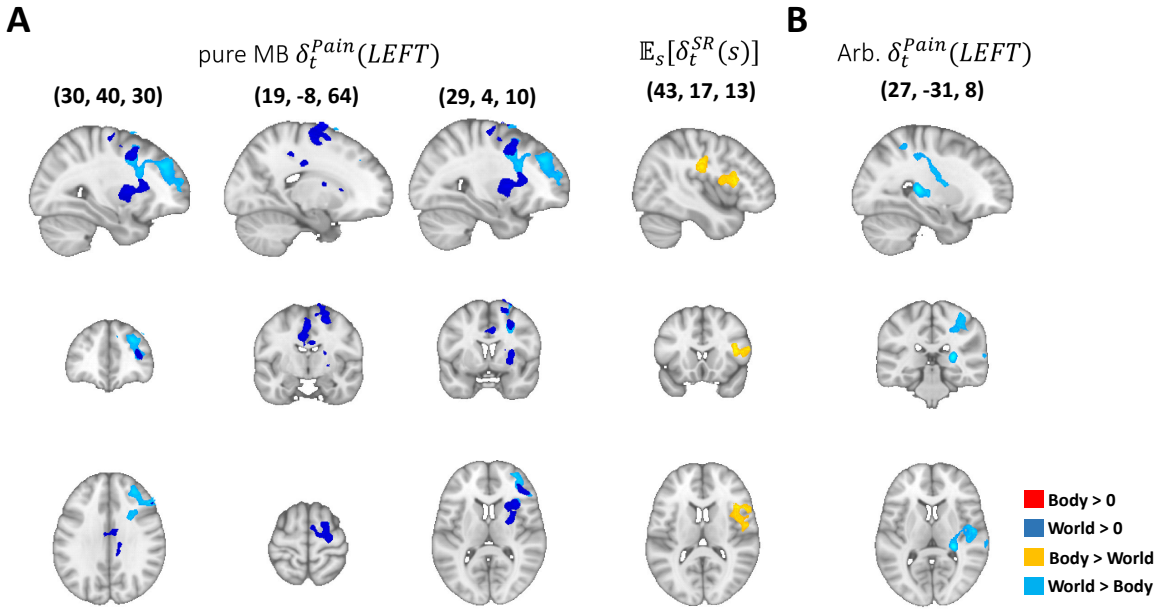

Figure 3.10: **The model-based temporal difference learning from the left arm supported a broad range of neural substrates for the state pain TD learning (which is independent of spatial information) correlating with world-model pain learning.**

(A) Under the pure MB learning model, the prediction TD error on the left side correlated with world-model pain condition in the primary motor cortex and putamen, whilst correlating more world-model pain than body-model pain condition in the supplementary motor cortex and dorsolateral/superior prefrontal cortex. The prediction TD error on the right side correlated more with body condition than world condition in the lateral occipital cortex, given a pure MF learning model. **(B)** Given the arbitrative learning model, the prediction TD error on the right side correlated more with world-model than body-model pain condition over a broad cluster covering lateral thalamus, posterior insula, and primary somatosensory cortex. The first row shows the current MNI152 coordinates of the corresponding column. All shown activated regions were FWER cluster-level corrected ($p = .05$).

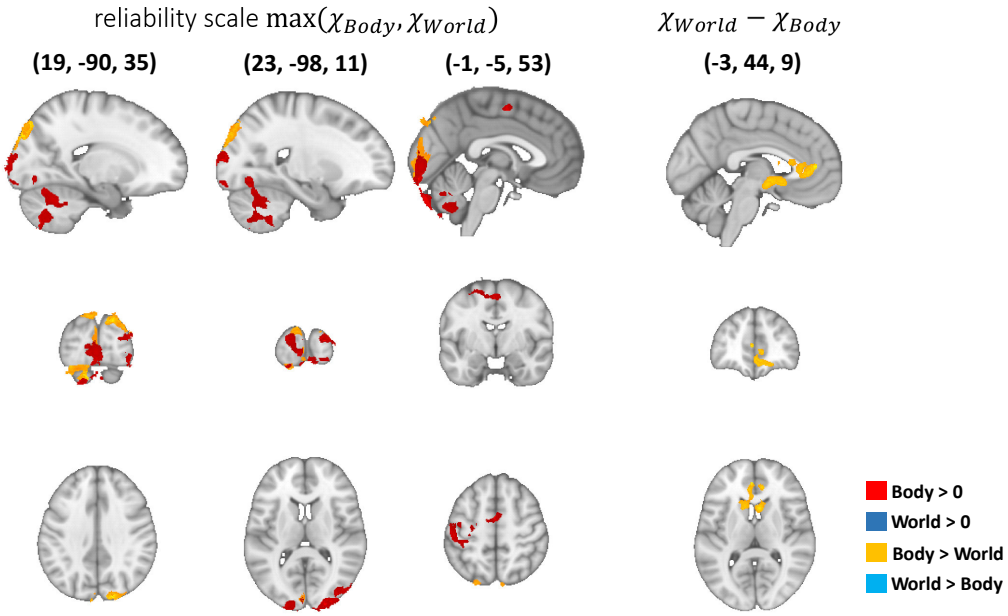

Figure 3.11: **A potential anterior cingulate circuit correlating with the arbitrative reliability between body-model and world-model pain learning.** **(A)** Given the arbitrative learning model, the reliability scale correlated with body-model pain in the dorsolateral occipital cortex and supplementary motor cortex, whilst correlating more body-model pain than world-model pain in the dorsolateral occipital cortex. **(B)** Given the arbitrative learning model, the reliability difference correlated with more body-model pain than world-model pain condition in a broad cluster around the anterior cingulate cortex. The first row shows the current MNI152 coordinates of the corresponding column. All shown activated regions were FWER cluster-level corrected ($p = .05$).

These findings altogether suggested a cognitive arbitration substrate ranging from prefrontal to cingulate cortical regions.

3.3.7 Anterior cingulate functionalisation on MB reliability enhancement via arbitration

Apart from pure MF and pure MB learning, the arbitrative model also implemented an explicit arbitrative mechanism between MF and MB learning pathways in a single system. To show its neural evidence, we conducted the GLM on arbitrative model with TD learning reliability scale $max(\chi_{Body}, \chi_{World})$ and reliability difference $\chi_{World} - \chi_{Body}$ that counted for the real-time arbitration (Figure 3.11).

The reliability scale correlated with body-model pain in the dorsolateral occipital cortex (dlOC; $z_{Peak} = 4.97, p = .019$, FWER cluster-level corrected) and supple-

mentary motor cortex (MII; $z_{Peak} = 4.37, p = .025$, FWER cluster-level corrected), whilst correlating more body-model pain than world-model pain in the dorsolateral occipital cortex (dlOC; $z_{Peak} = 4.11, p = .021$, FWER cluster-level corrected), which might be due to the visual-motor simplicity of the body-model condition per se for focusing on model-free learning more efficiently. The reliability difference ($\chi_{World} - \chi_{Body}$, $t = -.45$, Wilcoxon signed-rank test $p = .04$) correlated with more body-model pain than world-model pain condition in a broad cluster with the peak at the anterior cingulate cortex (ACC; $z_{Peak} = 3.76, p < .001$, FWER cluster-level corrected), covering medial prefrontal cortex (mPFC;$(5, 43, -12)$, Sphere ROI $r = 5$ mm$, p = .012$), anterior caudate ($(-12, 18, 5)$, Sphere ROI $r = 5$ mm$, p = .003$), anterior thalamus ($(4, -4, -1)$, Sphere ROI $r = 5$ mm$, p = .003$), and right amygdala ($(-18, -5, -16)$, Sphere ROI $r = 5$ mm$, p = .010$). This implied an anterior cingulate circuit correlating with the arbitrative reliability between body-model and world-model pain learning, with increased MB reliability even when learning MF-preferred action-induced pain (Figure 3.8, right). Given more learning periods, the reliability imbalance might thus be attenuated when both learning pathways approached much lower prediction errors [Geerts et al., 2020], whilst we suggested here as an anterior cingulate functionalisation on MB pathway learning originally unfavoured action-induced pain via arbitration with MF learning.

3.4 Discussion

We aimed to provide a large-scale neural architecture of dichotomic cognitive learning on pain for understanding the cognitive pain within complex environments: How and where a cognitive representation of pain is encoded in the brain, and what function it serves in pain-oriented learning behaviour (Section 1.4). We provided evidence for a fundamental dichotomy within cognitive pain learning: the cognitive distinction for avoiding exteroceptive pain learning actions between the objective experience-independent world (i.e. world-model) or the subjective bodily space (i.e. body-model). To address this, we designed a desktop-based virtual-navigation learning task in the MRI scanner with electric shocks delivered to both arms under different pain conditions that required the corresponding learning strategies. Participants showed significant aversive learning on pain avoidance performance (PAP) in both pain conditions. We then fitted our arbitrative RL model to human pain avoidance behaviours and the corresponding neuroimaging data during navigation. Model comparisons showed that model-free learning best fitted the body-model pain avoidance, whilst model-based

learning best fitted the world-model pain avoidance. This aligned with our model assumptions that the corresponding pain learning models could best describe the behavioural data within the respective environmental conditions for pain avoidance learning. Model variables showed predictive TD-driven learning with arbitrative reliabilities. Neuroimaging results further imply a temporal-parietal difference in favour of different pain learning pathways, a dorsolateral-medial prefrontal difference in predictive and arbitrative pain learning, sensory-motor and insula activations with various pain-predictive signals, and an anterior cingulate circuit correlating with reliability-based arbitration. Our findings implied a cortical-subcortical activity around the prefrontal, cingulate, and insula cortex correlating with cognitive pain learning arbitration.

3.4.1 Neural signatures compared to reward-oriented learning

We found the neural signature of MF pain learning on the anterior caudate activity for the arbitrative reliability difference $\chi_{World} - \chi_{Body}$ correlating with more body-model pain than world-model pain condition, as well as exceedance probabilities favouring pure MF learning correlates more with body-model pain condition in the superior temporal region. Apart from pain-oriented learning, the neural substrates correlating model-free (MF) and model-based (MB) learning were also largely found within the cognitive tasks on learning single scalar rewards [Daw et al., 2005, Huang et al., 2020]. A quantitative meta-analyses research conducted by Huang et al. [2020] on various fMRI studies showed higher MF learning correlation with the brain activity in caudate head and superior temporal gyrus, whilst medial frontal and anterior cingulate activity correlates more with MB learning. These neural correlations within a reward-oriented learning context accorded with our findings in pain-oriented learning.

The results of anterior caudate activity further indicated its computational mechanism with the TD-based reliability shifting towards the other 'incorrect' learning strategy when learning a specific pain condition, which we showed as an anterior cingulate circuit activation that accorded with another reward-oriented 2-stage learning task study using the similar arbitrative model [Lee et al., 2014]. This could also explain the general neural signature of MB learning correlating medial frontal and anterior cingulate activity, even though our findings only discovered the corresponding activity correlating more with the body-model pain condition. This was because the reliability difference ($\chi_{World} - \chi_{Body}$, $t = -.45$, Wilcoxon signed-rank test $p = .04$) in the world-model pain condition would generally step across 0 from negative to positive

so that the potentially significant correlation was hard to manifest. Exceedance probabilities favouring pure MB learning correlated more with world-model pain condition in the parietal region, probably because spatial knowledge learning was much more required in learning world-model than body-model pain in our task.

3.4.2 Neural signatures specific to pain-oriented learning

We found insula activity correlating with world-model condition in learning the state-independent pain TD on the left arm, in accordance with other studies showing predictive learning signals specific to pain in insula cortex [Brown et al., 2008, Geuter et al., 2017, Fazeli and Büchel, 2018, Koppel et al., 2022, Hoskin and Talmi, 2023]. We further showed this correlated more with world-model pain in dorsolateral and superior prefrontal regions, in comparison with the medial prefrontal TD learning body-model pain from the right arm. With successor representation that factorises spatial learning and pain learning, this implied a dorsolateral-medial prefrontal difference in the TD learning specific to pain between world-model and body-model learning.

The data fitting comparison indicated both body-dependent MF and body-dependent MB learning describing the corresponding avoidance behaviour, so that to imply the body parts as a necessary multidimensional input for pain prediction, rather than the single scalar input in traditional reward-oriented tasks. This supported a predictive 'body map' responsible for both body-level representation and prediction of pain [Brecht, 2017], with neural evidence on both the Room Onset phase activation in the primary motor region and laterality-biased predictive variables (e.g. TD errors and reliability scale) found in both pain conditions among somatosensory and motor regions in our findings. Though the 2-electrode setup in our experiment was still too simplified to consider a topological somatosensory pain perception, one of the future directions would be to incorporate a continuous bodily topological space for a more realistic pain avoidance learning.

3.4.3 Flexible learning under pain-oriented arbitration

A core idea of using an arbitrator over multiple learning pathways in a single system was to handle the varying environmental conditions for learning [Lee et al., 2014], aligned with the human learning capability on flexible shifting between 'automatic' habitual behaviour and 'deliberative' goal-directed behaviour for evaluating actions [Balleine and Dickinson, 1998, Dayan and Daw, 2008, Drummond and Niv, 2020]. Our virtual-navigation task did not implement such environmental variations, as

raising the task difficulty may obstruct their learning efficiency within the limited time in the scanner. Participants were always informed of the current pain condition to control the task difficulty, though this resulted in worse data fitting results of the arbitrative model compared to the pure MF or pure MB model with explicit prior knowledge in model selection. Future research would focus on the model simulation and experimental design with variant pain conditions in a single block to show the necessity of the arbitrative mechanism for pain avoidance learning on complex and varying dangers.

The successor representation we used for MB learning addressed a factorisation between spatial-state predictive coding and pain-state evaluation, which might enable flexible learning across multiple environments by explicitly encoding and storing predictive relationships among spatial states. It was suggested as a hippocampal mechanism for between-state predictive learning without explicit planning [Stachenfeld et al., 2017, Geerts et al., 2020], where the hippocampus had been reported showing neuroimaging activations in various reward-oriented MB learning tasks [Bornstein and Daw, 2012, Sebold et al., 2017]. Future research would consider more complex environmental setups with multiple scenarios to discover hippocampal circuit recruitment when learning pain avoidance under large-scale navigation across sensory-rich environments.

Chapter 4

The iVR-EEG interoceptive pain modulation

4.1 Introduction

Previous cognitive models of sensory information processing assume that humans aim to infer the causes of afferent sensory input [Gregory, 1980, Hatfield, 1990, Friston, 2005, Körding et al., 2007, Tabor et al., 2017, Seymour and Mancini, 2020]. In the case of pain, this lies in two distinct categories of nociceptive causes: those external to the body that may cause potential harm which is avoidable (e.g. a burning stove in the real world) [Sherrington, 1906, Price et al., 2003], and those internal to the body that reflect potentially damaged body tissue inducing persistent pain that is not avoidable (e.g., skin that is more sensitive to burning) [Geldard, 1972, Craig, 2002, 2003]. These provide a core nociceptive dichotomy between two aspects of pain: exteroception, of which the knowledge we extract from the external environment for protective behaviours against potential threats [Price et al., 2003]; and interoception, of which the knowledge we obtain from the internal body as a physiological condition along with homeostatic behaviours [Wall, 1979, Craig, 2003]. Thus, a fundamental distinction between exteroceptive and interoceptive pain could be figured out in such a conceptual basis of pain. The brain learns the motivational information of pain from the external world as exteroceptive pain, meanwhile learns the nociceptive information of pain in the internal body as interoceptive pain [Seymour, 2019, Seymour and Mancini, 2020].

From a nociceptive perspective, the pain information of 'how does the brain realise that the body is injured' is inferred and affected by both exteroceptive pain [Price et al., 2003] and interoceptive persistent (or chronic) pain [Di Lernia et al., 2016,]. They are processed separately in two spinothalamic pathways in the brain [Dostrovsky and Craig,

2006], while they may share bidirectional interactions. Persistent interoceptive pain may drive an endogenous control on the sensation of co-occurring exteroceptive pain, e.g. in the conditioned pain modulation [Kennedy et al., 2016]. On the other hand, learning the exteroceptive pain for avoidance requires an inhibition on interoceptive persistent pain that is abnormal within chronic pain conditions [Staud, 2012].

In the previous MRI study, we have shown that human brains handled cognitive problems located within the external world (as world-model pain) and induced by the bodily actions (as body-model pain) from nociception for pain avoidance learning (Chapter 3), facilitated by the model-free/model-based arbitrative reinforcement learning model (Chapter 2). Thus, the human brain could infer exteroceptive causes flexibly between external objective states and internal subjective actions. However, the MRI study used a desktop-based VR task that required participants to lie still in the scanner, rather than moving freely with natural bodily avoidance behaviour. The study did not focus on how the exteroceptive pain processing affected pain nociception, and the time course was not considered as a main interest due to the low temporal resolution of MRI data. Hence, we further investigated the connection between nociception and such exteroceptive cognition, which underlies another fundamental distinction: Nociceptive problems from the external world (as exteroception) and in the internal body (as interoception) for cognitive modulation on nociception and sensation. We also used an electroencephalograph (EEG) with high temporal resolution in an immersive Virtual Reality (iVR) game, allowing free-operant moving in the open arena. The iVR technique enables an immersive pain learning environment for human participants to conduct natural bodily movement for realistic pain avoidance behaviour, which has been recently used for phasic and tonic pain valuation [Hewitt et al., 2025, Tong et al., 2025] as well as chronic pain management and treatment [Li et al., 2011, Birckhead et al., 2021, Goudman et al., 2022, Teh et al., 2024].

How does learning exteroceptive pain provide an endogenous control on the persistent sensation (including pain) to keep the cognitive functionality of danger avoidance? The persistent sensation itself is not predictable for pain avoidance, meanwhile yields a chronic sensation (or pain) state appropriate to injury. We assume that these two different cognitive essences of exteroceptive pain could modulate interoceptive nociception or sensation, represented by certain brain activities. Previous EEG studies linked cognitive control to high frontal potentials [Kopp et al., 1996, Brass et al., 2005, Hilgard et al., 2014, Kopp et al., 2020] and mid-frontal theta activity [Cavanagh and Frank, 2014, Zavala et al., 2018, Eisma et al., 2021, Tan et al., 2024], and other cognitive task studies linked central alpha and beta oscillations to somatosensory

decision-making and sensorimotor processing [Witham et al., 2007, Haegens et al., 2011, Quandt et al., 2013, Brickwedde et al., 2019]. Therefore, we expect that the action-induced body-model pain would evoke more posterior EEG activities and relevant alpha and beta oscillations, whilst the state-induced world-model pain would evoke more anterior EEG activities with frontal theta synchronisation, for modulating subjective perceptive and illusory nociception.

We designed the experiment to discover the interaction of persistent nociception and sensation modulated by these two pain cognition processes within an iVR navigation task. We showed discrepant representations of body-model and world-model exteroceptive pain in EEG as posterior-anterior ERPs, and as body-model beta and world-model frontal theta power. We further discovered exteroceptive pain that manipulated sensory and illusory nociception with centrofrontal synchronisation in various frequency bands. These supported our hypothesis on the exteroception-modulated sensory and illusory nociception with electroencephalographic support.

4.2 Methods

4.2.1 Immersive Virtual Reality (iVR) task design

A computer-based immersive virtual reality (iVR) spatial navigation task was created in Unity 2019.4 with a first-person view (Figure 4.1A) with a VR headset (HTV Viva Pro Eye) connected to Steam VR. Participants were asked to avoid as many potential shocks as possible when bending their backs in the task. With Digitimer DS7A providing electric shocks, 1 electrode was applied to the skin surface of the back (Figure 4.1B), above which the menthol cream was applied on the back to provide a persistent and illusory bodily sensation which might be interpreted as noxious.

4.2.1.1 Navigational task in iVR

The iVR task consisted of a 3×3 square grid in a 4 m $\times$ 4 m open arena, with each square panel on the ground showing the potential target position that participants needed to reach for collecting crystals (Figure 4.1A). Participants could look around and move freely in this open arena set up in the iVR. The sun was visible in a specific direction of the sky, which participants always started to face towards, so participants could locate themselves in the virtual open arena. During the 3-minute task in each block, participants were asked to bend their backs when reaching the square panel, meanwhile stretching their right hand to collect crystals on the ground under various pain conditions. Participants could obtain crystals as visible rewards once they reached

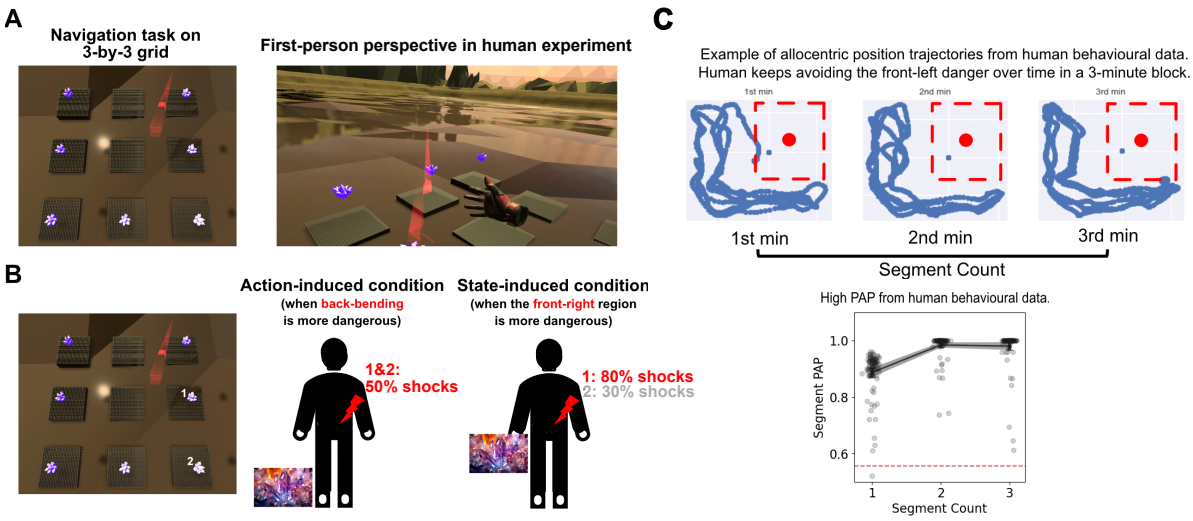

Figure 4.1: **Spatial moving task within the virtual environment with electrical stimuli delivered under different exteroceptive pain conditions.**

(A) A virtual environment scenery with the first-person perspective where participants could move freely around. **(B)** Example of different real pain stimulation patterns in the action-induced condition (back-bending as a dangerous action) and state-induced condition (front-right region as a dangerous place, indicated by the centred red beacon) when participants stayed in the middle square panel and faced towards north. Crystals were visible monetary rewards that required participants to bend their backs and stretch their arms to touch. For numbers indicating which crystal was touched (left), red numbers beside the electrode on the lower-right back led to a high shocking probability, whilst grey numbers refer to a low shocking probability (middle & right). In action-induced conditions with body-model pain, a back-bending action could lead to a potential shock, whilst in state-induced conditions with world-model pain, the shock was given when touching the crystal on a dangerous/non-dangerous square panel. **(C)** An example of the participant movement trajectory in the iVR open arena in a state-induced condition, avoiding the dangerous front-left region (top). We showed the significant PAP across all participants over the 55.56% guess rate in the state-induced condition (bottom), indicating that participants were indeed avoiding the dangerous region to suffer fewer shocks. Error bars stand for 95% CI.

the square panel where the crystal existed (Figure 4.1A). They were encouraged to keep moving to obtain as many crystals as possible, as they would count for the final monetary outcome (10p for each, 50 crystals as the upper limit they could obtain in a 3-min block) that participants obtained at the end of the task, which was independent of the shocks they might receive.

Depending on the movement and experimental condition, participants might receive electric stimuli on their backs, subject to the specific exteroceptive pain (i.e. body-model or world-model exteroceptive pain) as the experimental session condition (Figure 4.1B). Participants were always informed of which one of the 2 pain conditions (see below) it was at the start of each session for all the following 4 blocks. After each block, participants were asked to select which dangerous action or region they had experienced in the corresponding block to see if they successfully understood how to avoid visible dangers.

4.2.1.2 Action-induced body-model pain condition

For the action-induced pain condition with body-model pain, participants have an action contingency q to receive an electric shock as soon as they bend their back for more than 30 degrees (Figure 4.1B, middle). Participants had the conditioned probability $q_+ = 0.5$ to receive a shock when bending their backs, whilst having the conditioned probability $q_- = 0$ to receive a shock without bending the back, so that the action contingency $q = q_+ - q_- = 0.5$.

4.2.1.3 State-induced world-model pain condition

For the state-induced pain condition with world-model pain, participants had a state contingency p among different states to receive electric stimuli, depending on the position of dangerous places (Figure 4.1B, right). Here, the visible dangerous places composed a cluster of a 2×2 region of square panels, which were randomly selected and counterbalanced across blocks, with a visible red beacon located at the centre of the dangerous region. Participants had the conditioned probability $p_+ = 0.8$ to receive a shock as soon as touching a crystal on the dangerous square panel, whilst having the conditioned probability $p_- = 0.3$ to receive a shock as soon as touching a crystal on the non-dangerous square panel, so that the state contingency $p = p_+ - p_- = 0.5$.

4.2.2 Experimental design

4.2.2.1 Participants

32 participants (18 females) attended the study after submitting informed consent. They were all healthy people aged 18-45 years (Mean: 25.81, SD: 5.64) without significant mobility problems, diabetes, hypertension, hypotension, blood circulation problems, or existing painful conditions. The study was approved by the Central University Research Ethics Committee (CUREC; Approval Reference: R58778/RE005) of the University of Oxford. Due to the technical failure caused by large bodily movement with the mobile EEG and VR setup, we removed 9 participants' data prior to EEG preprocessing, so that the final dataset consisted of 23 participants' data (14 females) aged 18-45 years (Mean: 26.04, SD: 5.71) used for behavioural and electroencephalographic data anaylsis.

4.2.2.2 Experimental procedures

The experiment consisted of 2 visit days at least 1 week apart (Figure 4.2A). On each visit day, participants first took part in a training process in the VR room to play the video game wearing the VR headset, with detailed instructions to get familiar with the corresponding pain condition required for playing the navigational iVR task that day. After the training session, the 32-channel EEG cap and electrodes were put on the participants' heads. Participants then experienced one 20-minute EEG session with the instructed corresponding pain condition. Each session consisted of four 5-minute task blocks with the same pain condition, and each task block consisted of a 3-minute navigational block wearing the VR headset and a following 2-minute interval for participants to report ratings after removing the VR headset (Figure 4.2B). Two pain conditions were counterbalanced between EEG sessions on different days across all participants. Half of the participants had a body-model pain condition as the main EEG session on Day 1, whilst the world-model pain condition on Day 2, and others had the reversed pain condition design.

4.2.2.3 Pain ratings

Participants reported the intensity of electrical shock pain experienced on a VAS (visual analogue scale) in a 0–10 range (0: no pain, 4: start to feel pain, 10: worst pain imaginable) at the time of testing (Figure 4.2B, top, Q1). Before the EEG setup, participants experienced the pain calibration process, where the current intensity was increased from 0 mA in steps of 1-10 mA until the participant reported the VAS

Figure 4.2: **Experiment protocol counterbalancing two pain conditions between 2 EEG sessions on different days with questionnaires on various sensation ratings.**

(A) The experiment protocol separated body-model and world-model pain conditions in two different experimental days, counterbalanced among participants. **(B)** Questionnaires both on various sensation ratings for participants to report after each 3-min EEG block (top) and on their general illusory perception of the back after the whole EEG session (bottom). Participants were required to remove the VR headset before filling out the questionnaire on the laptop screen, and wear the VR headset again to continue the next EEG block.

as 7. The calibration procedure was repeated three times, with the median current thresholds used for the main experimental session later. At the start of each EEG session, a recalibration process was conducted to fine-tune the current intensity from the last setting until the participant reported the VAS as 7 again.

4.2.2.4 Electrical stimuli

A pair of snap electrodes was applied to the surface of the participant's lower-right back. Controlled current was applied only to the prepared surface area, without passing internally into the body. A Digitimer DS7A (Hertfordshire, UK) was used to elicit a low level of electrical output sufficient to induce a moderate-to-strong pain sensation calibrated to each participant. Participants received the output current (Day 1: Mean 81.41 mA, SD 35.09 mA, max 200 mA; Day 2: Mean 88.48 mA, SD 37.48 mA, max 200 mA) with a series of 10 pulses as a single shock (2 ms/pulse, 10 pulses/shock in 2 seconds, 250 Hz) with a source voltage of 400 V.

4.2.2.5 Menthol cream

The menthol cream was used to apply a sustained somatosensory stimulus to participants' backs. When such static contact cooling occurred after the existing contact between menthol and human skin, the menthol cream might induce not only cold sensations but also various sorts of nociceptive sensations (e.g. burning, stinging, or pricking) known as the innocuous cold nociception (ICN) via the menthol-sensitive channel TRPM8 [Green and Jennifer, 2003, Green and Kate, 2007]. Such ICN could also induce cold hyperalgesia [Namer et al., 2005]. The persistent sensation itself was not learnable for avoiding external dangers with phasic electrical shocks, so that we could find out how exteroceptive pain might modulate nociceptive sensations with broader meanings. Thus, our usage of menthol cream aimed to mimic a persistent bodily sensation which might be interpreted as noxious following the experimental manipulation.

4.2.2.6 Questionnaire ratings on sensation and illusory nociception

Before the start of the main session, The Dermacool Plus 2% Menthol Aqueous Cream was used to apply the menthol cream onto participants' backs, imitating a permanent nociception. The 8-ml menthol cream was spread on a 5 cm $\times$ 7 cm pad (Tegaderm 3582 Bandages, 3M) pasted onto the back immediately above the electrode. 5 minutes after menthol cream administration, participants completed a questionnaire to assess subjective sensations on their backs (Figure 4.2B, top, Q2-8) based on a VAS (visual

analogue scale) in instructed 0–10 range. This was to see if the menthol started to functionalise, so that participants would be required to wait for a longer time until the cream was rated as providing either a hot or cooling sensation of at least 4 out of 10. After each block, participants were also required to give ratings on the same questionnaire. At the end of each main session, participants also gave ratings on 3 extra questions (Figure 4.2B, bottom) about their general feelings on the illusory nociception on the backs (where the menthol cream was applied), as these questions were referring to the interoceptive back injury that should not have existed. After the whole main session, the menthol cream was removed from their backs.

4.2.2.7 EEG data acquisiton

Electroencephalography was recorded using a 32-channel Ag-AgCl LiveAmp system (BrainProducts GmbH, Munich, Germany) placed on the scalp according to the International 10-20 system [Jasper, 1958], with a sampling rate of 500 Hz. Reference was set up as FCz, and the ground was set up as FPz. All scalp electrodes were collected with an average reference.

4.2.3 Electroencephalographic data analysis

4.2.3.1 Preprocessing

We processed and analysed EEG data using EEGLab [Delorme and Makeig, 2004] in Matlab 2022a. The data were resampled to 250 Hz, followed by the high-pass filtering at 30 Hz and the low-pass filtering at 0.1 Hz using a finite impulse response (FIR) filter. The high-pass filter was selected due to the strong motion artefacts in the mobile EEG data, as participants conducted large body movements in our free-operant iVR task. Oculomotor and muscular artefacts were removed using a 2-stage independent component analysis (ICA) [Hjorth, 1975]. In each ICA stage, the data was filtered to 1-30 Hz and resampled at 100 Hz. Outliers with the absolute amplitude above 500 μV were then detected and removed before we conducted the ICA. Artifactual components highly related to muscle, eye blinks, eye movements, channel or line noise, and heartbeats were then manually detected and removed. Subsequently, ICA weights were exported and projected back onto the previous data, which was the data before being filtered to 1-30 Hz, but further removing outliers with the absolute amplitude above 750 μV. These steps were repeated twice as a 2-stage ICA. Electrodes with large artefacts remaining were interpolated ($< 10\%$ electrodes; 3 channels maximum), followed by all-channel average re-reference.

The preprocessed data at this stage were then epoched from -1 to 2 s around the following 3 events: shock onset, crystal touching, and back bending. Shock onset event referred to the time of the electric stimuli delivery. Crystal touching event referred to the time when participants used their virtual right hand to touch the crystal on the ground. Back bending event referred to the time when participants bent their backs more than 30 degrees from the perpendicular.

4.2.3.2 Disentangling overlapping events

Important recorded events (shock onset, crystal touching, and back bending) could sometimes be quite close to each other due to the free-operant design in our iVR task, rather than the restricted block design. We used the *Unfold* toolbox to adjust for overlap between subsequent potentials on the epoched EEG data using linear deconvolution [Ehinger and Dimigen, 2019]. Outliers with the absolute amplitude above 750 μV were first detected and dropped before fitting to the deconvolution model on all 3 events, where the stick function approach was used to model each time-point relative to the local event as represented by a unique predictor. The event-related potentials (ERP) corresponding to the 3 events showed decreased potential scale with linear deconvolution compared to the original GLM without detecting event potential overlaps (Figure 4.3), implying the disentangled event potentials. With linear deconvolution, the ERP related to shock onset also showed a steady N1-P1-N2 complex between 100 and 400 ms (Figure 4.3, top-left, dashed red box) believed to be associated with nociception [Lee et al., 2009], whereas the complex was highly distorted without deconvoluting overlapping event potentials (Figure 4.3, bottom-left).

4.2.3.3 Event-related potential analysis

With the baseline correction on $[-500, 0]$ ms, event-related potentials (ERP) after linear deconvolution with *Unfold* toolbox were then averaged over individual ERPs subject to different pain conditions (see an example of channel Cz in Figure 4.3, top). Cluster time intervals (required at least 50 ms) were detected by the analytic t-distribution cluster-based multiple comparison correction [Maris and Oostenveld, 2007] using *FieldTrip* toolbox [Oostenveld et al., 2011], followed by the paired t-test for the corresponding mean potential difference between body-model and world-model pain condition within the corresponding time interval.

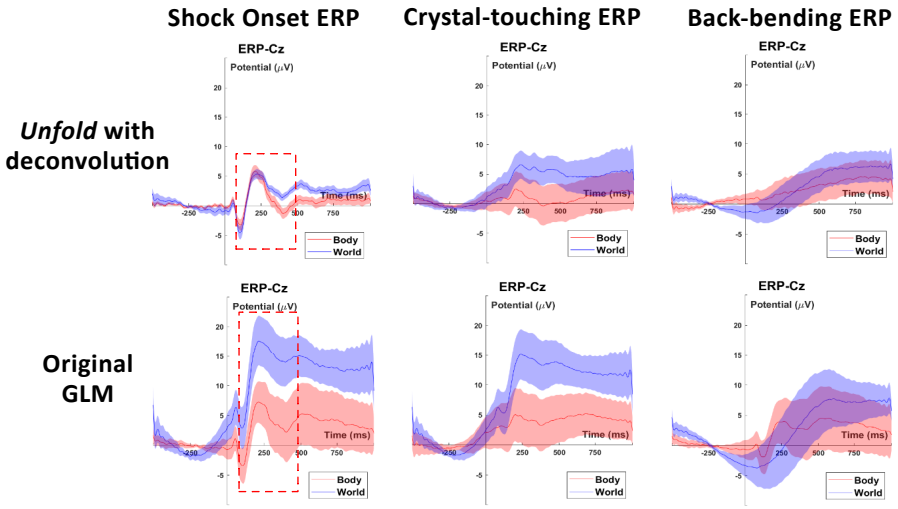

Figure 4.3: ***Unfold* toolbox with linear deconvolution enabled disentangling main event potentials with steady N1-P1-N2 complex associated with nociception.**

ERPs of channel Cz on 3 events (columns) were calculated either via linear deconvolution with *Unfold* toolbox (top) or the original GLM without detecting event potential overlaps (bottom). With the baseline correction on -500 - 0 ms, they were then averaged over participants subject to different pain conditions (red for body-model pain, and blue for world-model pain). ERPs with linear deconvolution showed decreased potential scale compared to the original GLM, meanwhile providing a steady N1-P1-N2 complex for shock-onset ERP between 100 and 400 ms. Filled shades around EEG curves indicate standard errors.

4.2.3.4 Time-frequency representation analysis

With the baseline correction on $[-500, 0]$ ms, time-frequency representation (TFR) was detected with transient event-related spectral perturbation [Makeig, 1993]. TFR was obtained using a Morlet 3-cycle wavelet transform (with a Hanning-tapered window applied) with cycle sizes of $0.8 * \mathbf{f}$, restricted in the frequency range as 1-30 Hz. The upper limit of this frequency range was set due to the strong motion artefacts in the mobile EEG data, as participants conducted high body movements in our free-operant iVR task. Relative band power was then extracted as theta power (4-7 Hz), alpha power (8-13 Hz), and beta power (14-29 Hz). Electrodes were selected among those with significant ERP differences. Besides, Fz was additionally included to specifically assess mid-frontal theta activity linked to cognitive control and evaluative processing [Cavanagh and Frank, 2014, Zavala et al., 2018, Eisma et al., 2021, Tan et al., 2024]. Cluster time intervals (required at least 50 ms) were detected by the analytic t-distribution cluster-based multiple comparison correction [Maris and Oostenveld, 2007] on one of the frequency bands using *FieldTrip* toolbox [Oostenveld et al., 2011], followed by the paired t-test for the mean power difference between body-model and world-model pain condition within the corresponding time interval before analysing the other 2 frequency bands with the same time interval.

4.2.3.5 Rating-signal correlation GLMs

General linear models (GLMs) were used to investigate how different pain conditions might manipulate the correlation between various sensation ratings and electroencephalographic signals. For each rating in the questionnaire (Figure 4.2B), the GLM was set up as

$$Rating \sim Signal + C(Condition) + Signal \times C(Condition)$$

where we had $C(Condition)$ as the categorical variable of the body-model/world-model pain condition, whilst having $Signal$ as either the mean ERP or the mean TFR within a given time interval. The GLM disentangled these two main effect terms from the interaction term $Signal \times C(Condition)$ that we were interested in, which indicated how different pain conditions would manipulate the changes of subjective ratings corresponding to electroencephalographic signals. The Wald test was then conducted on fitted coefficients of the corresponding independent variables in GLM, corrected by positive false discovery rates (FDR) with the Benjamini-Hochberg procedure for multiple hypotheses testing [Benjamini and Hochberg, 1995].

4.3 Results

Data from $N = 23$ participants were included for the whole data analysis procedure. Participants received an average of 37.79 ± 7.13 times of the electric stimuli during the 3-minute virtual navigational game. In the action-induced condition, participants experienced potential shocks when bending their backs to collect crystals. Whilst in the state-induced condition, crystals were dangerous if they were positioned in the current dangerous region, indicated by a visible beacon in the task. We defined pain avoidance performance (PAP) as touching the crystal, which led to a lower possibility of receiving shocks. Participants performed significantly higher PAP then the 55.56% guess rate in all three 1-min temporal segments in the state-induced condition (Figure 4.1C, bottom, $p < .001$), suggesting that participants were indeed avoiding dangerous regions all the time to experience fewer shocks in the state-induced condition.

4.3.1 Extracting subjective perception to principal components

During the task, participants reported their rated various subjective sensations (Figure 4.2B, top) every time after they finished a 3-minute task block (Figure 4.4A). As these 8 ratings were highly correlated among each other (Figure 4.4B), we performed a principal component analysis (PCA) to extract main components explaining more than 85% percent of the rating variance (Figure 4.4C, bottom), resulting in 4 principal components on sensation ratings (Figure 4.4C, top row for PCA weights, middle rows for the corresponding ratings). According to the contribution weights from original sensations to PC ratings, these 4 principal sensation components could be respectively described as perception (except coldness) (PC1), hot tonic pain (PC2), cold pain (PC3), and tonic-phasic nociception discrepency (PC4).

A two-way repeated-measure ANOVA was respectively conducted on those sensation PC ratings over blocks (including the calibration ratings before the first block, shown as 'C' in Figure 4.4C) and conditions. We found the main effect on some of these ratings over blocks (PC1: $F(4, 88) = 13.196, p_{FDR} < .001$; PC3: $F(4, 88) = 17.091, p_{FDR} < .001$; PC4: $F(4, 88) = 7.159, p < .001$) except on PC2 ($F(4, 88) = 1.819, p_{FDR} = .132$). This suggested general temporal drifts of subjective sensations across the task, including the increasing perception and tonic-phasic nociception discrepancy, as well as the decreasing cold pain.

After the whole EEG session, participants also rated their subjective illusory perception (including pain, see Figure 4.2B, bottom) on the back where the menthol cream was applied (Figure 4.5A). As these 3 ratings were highly correlated among

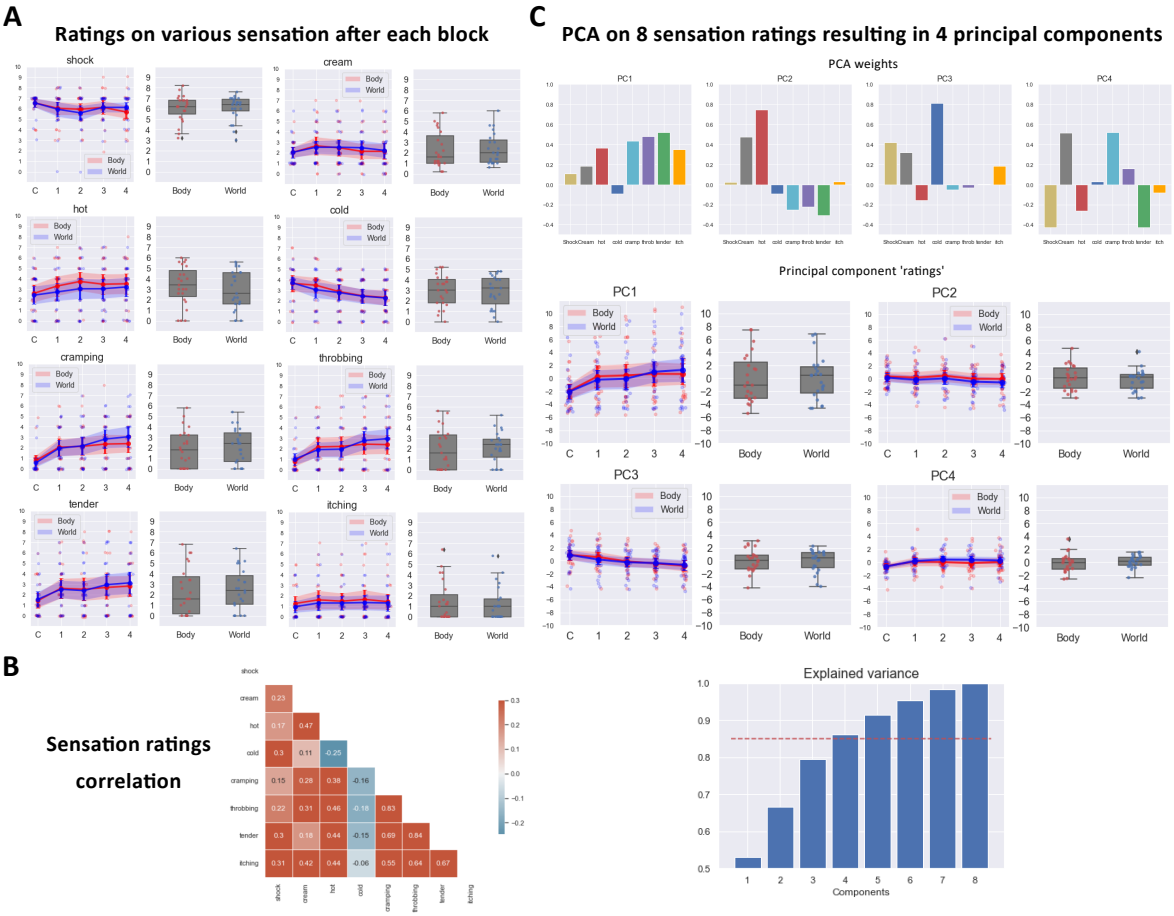

Figure 4.4: **8 subject sensation ratings were extracted to 4 principal components.**

(A) Subjective ratings over temporal blocks (left panel for each rating) and averaged across blocks (right panel for each rating) collected from the questionnaire (Figure 4.2B, top) on various sensations after the 3-minute task as each block. 'C' for the calibration process, where the menthol cream was just applied before the 1st block. **(B)** Correlation heatmap with Pearson correlation coefficients on sensation ratings in A showed considerable correlations among these sensation ratings. **(C)** PCA extracted main components from sensation ratings in A that explained more than 85% of the rating variance, resulting in 4 principal components for sensation ratings. They could be respectively described as perception (except coldness) (PC1), hot tonic pain (PC2), cold pain (PC3), and tonic-phasic nociception discrepancy (PC4). For box plots, the three quartile values of the distribution were plotted, with extreme values beyond 1.5 IQR marked as black dots. For curve plots, error bars stand for standard errors. For explained variance plots, dashed red lines stand for 85% of the explained variance.

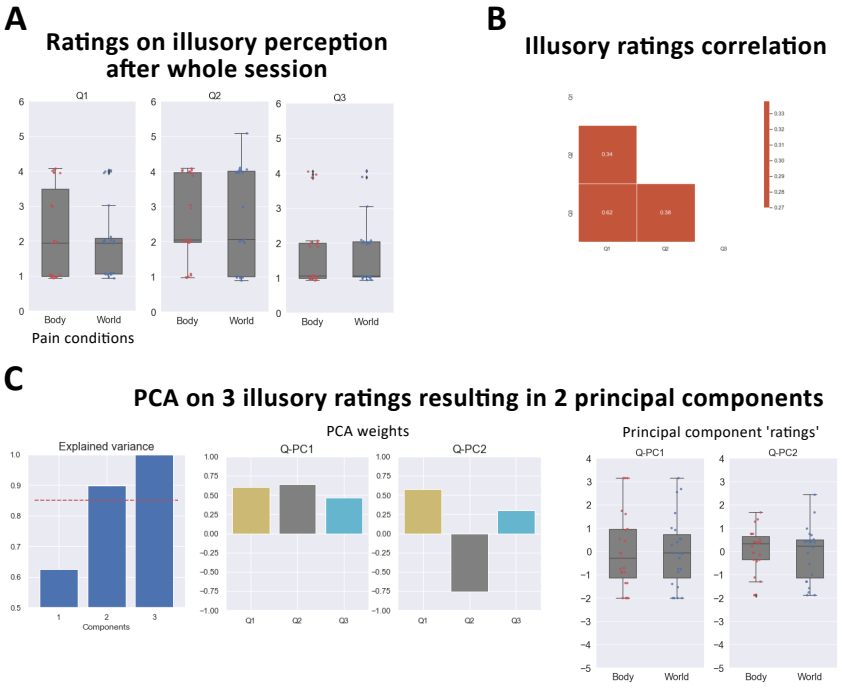

Figure 4.5: **3 illusory perception ratings were extracted to 2 principal components.**

(D) Subjective ratings on illusory perception (Figure 4.2B, bottom) after the 4-block session. **(E)** PCA extracted main components from illusory perception ratings in D that explained more than 85% of the rating variance, resulting in 2 principal components for illusory pain ratings. They could be respectively described as illusory pain (Q-PC1) and tonic-phasic illusory pain discrepancy (Q-PC2). **(F)** Correlation heatmap with Pearson correlation coefficients on illusory perception ratings in D showed considerable correlations among these illusory ratings. For box plots, the three quartile values of the distribution were plotted, with extreme values beyond 1.5 IQR marked as black dots. For curve plots, error bars stand for standard errors. For explained variance plots, dashed red lines stand for 85% of the explained variance.

each other (Figure 4.5B), we also performed a principal component analysis (PCA) to extract main components explaining more than 85% of the rating variance (Figure 4.5C, left), resulting in 2 principal components for illusory perception ratings (Figure 4.5C, middle for PCA weights, right for the corresponding ratings). According to the contribution weights from original illusory perception to PC ratings (Figure 4.5C, middle), they could be respectively described as illusory pain (Q-PC1) and tonic-phasic illusory pain discrepancy (Q-PC2). No significant main effects of pain condition were found on the illusory perception PCs (Figure 4.5C, right; Q-PC1: $p = .912$; Q-PC2: $p = .478$).

4.3.2 ERPs distinguished with posterior-anterior discrepancy between body-model and world-model pain

To examine how body-model and world-model pain were differentially represented in the brain, we conducted event-related potential (ERP) analysis on deconvolved EEG responses time-locked to three key events: shock onset, crystal touching, and back bending. Following shock onset, body-model pain elicited stronger ERPs at centrofrontal electrodes, whilst world-model pain elicited stronger ERPs at more anterior central and frontal electrodes. Topographic maps showed that brain activity during early latencies was centrally distributed across both conditions (Figure 4.6, top). However, while body-model pain led to a posterior and bilateral shift in activity over time, world-model pain maintained a more focal distribution across frontal and central electrodes (Figure 4.6, bottom). ERPs for body-model pain showed significantly greater amplitude than for world-model pain at centrofrontal electrodes (C4 and FC6). Specifically, greater amplitudes were observed at C4 during early latency ($[4, 216]ms : \Delta = 2.48, t(22) = 2.91, p = .008$) and late latency ($[540, 836]ms : \Delta = 3.56, t(22) = 2.35, p = .028$), and at FC6 during early latency ($[216, 280]ms : \Delta = 3.58, t(22) = 2.29, p = .032$) and mid latency ($[440, 464]ms : \Delta = 3.60, t(22) = 2.24, p = .035$) (Figure 4.6, left). In contrast, world-model pain evoked significantly larger amplitude ERPs than body-model pain at central and frontal electrodes (F4 and Cz), specifically at F4 during mid latency ($[276, 372]ms : \Delta = -3.80, t(22) = -2.36, p = .028$), and Cz during mid latency ($[360, 676]ms : \Delta = -2.40, t(22) = -3.05, p = .006$) and late latency ($[920, 1000]ms : \Delta = -2.3, t(22) = -2.44, p = .023$) (Figure 4.6, right).

In response to crystal touching, a motor action involving forward-reaching with the right arm, body-model pain again elicited stronger ERPs at right parietotemporal electrodes, whilst world-model pain elicited stronger ERPs in the centrofrontal electrode. The medial-bilateral drifting ERP amplitude was found for body-model

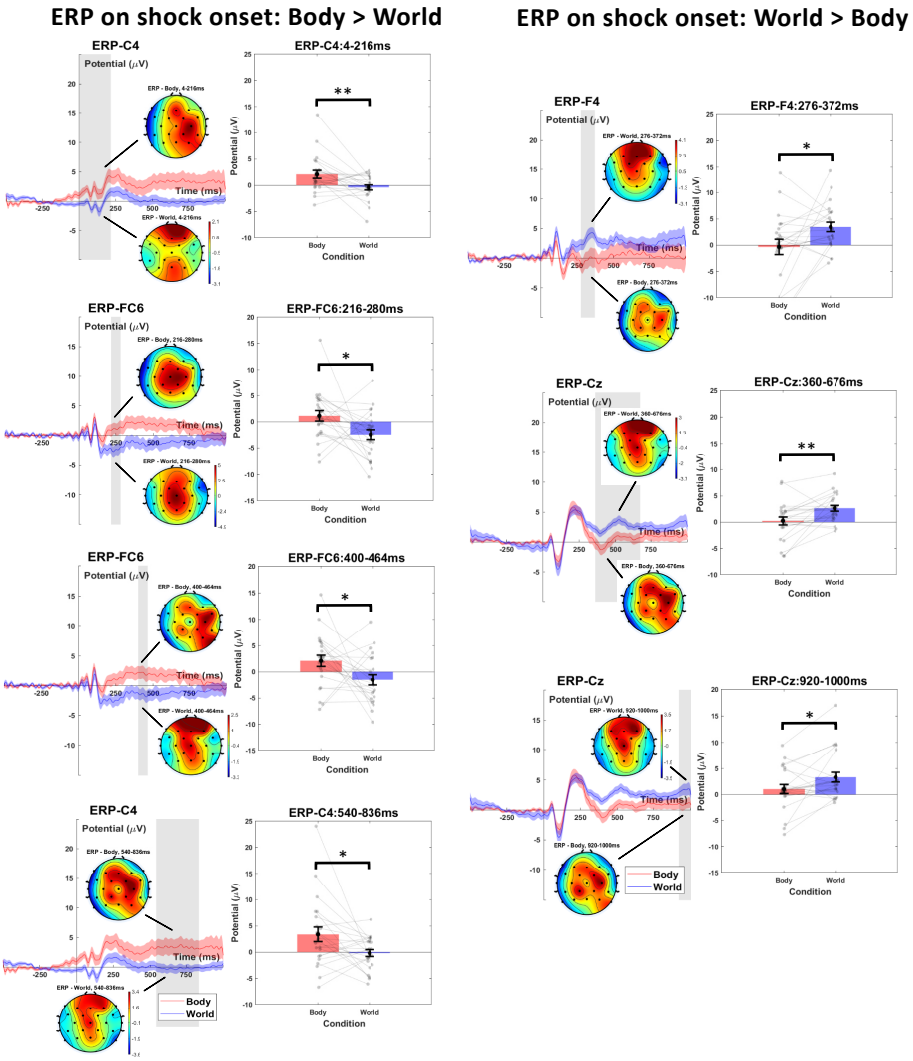

Figure 4.6: **ERPs distinguished with posterior-anterior discrepancy between body-model and world-model pain from early to late latencies after shock stimuli events.**

ERPs after shock onset for body-model pain showed significantly greater amplitude than world-model pain at right central electrodes, including C4 and FC6 (left), whilst with greater amplitude in world-model pain at central and frontal sites, including F4 and Cz (right). Topographic maps showed early medial activity that shifted bilaterally in mid-to-late latencies for body-model pain, whereas world-model pain maintained a more central focus. ERP plots were sorted in the temporal order (top for early latency and bottom for late latency), given specific ERP events and the dominant pain condition. For each ERP plot, topographic maps subject to pain conditions were given at each clustered time interval (highlighted as grey zones; see Section 4.2 for its achievement method) with significant mean potential difference shown in the corresponding bar plot. Error bars stand for standard errors. ** for $p < .01$ and * for $p < .05$.

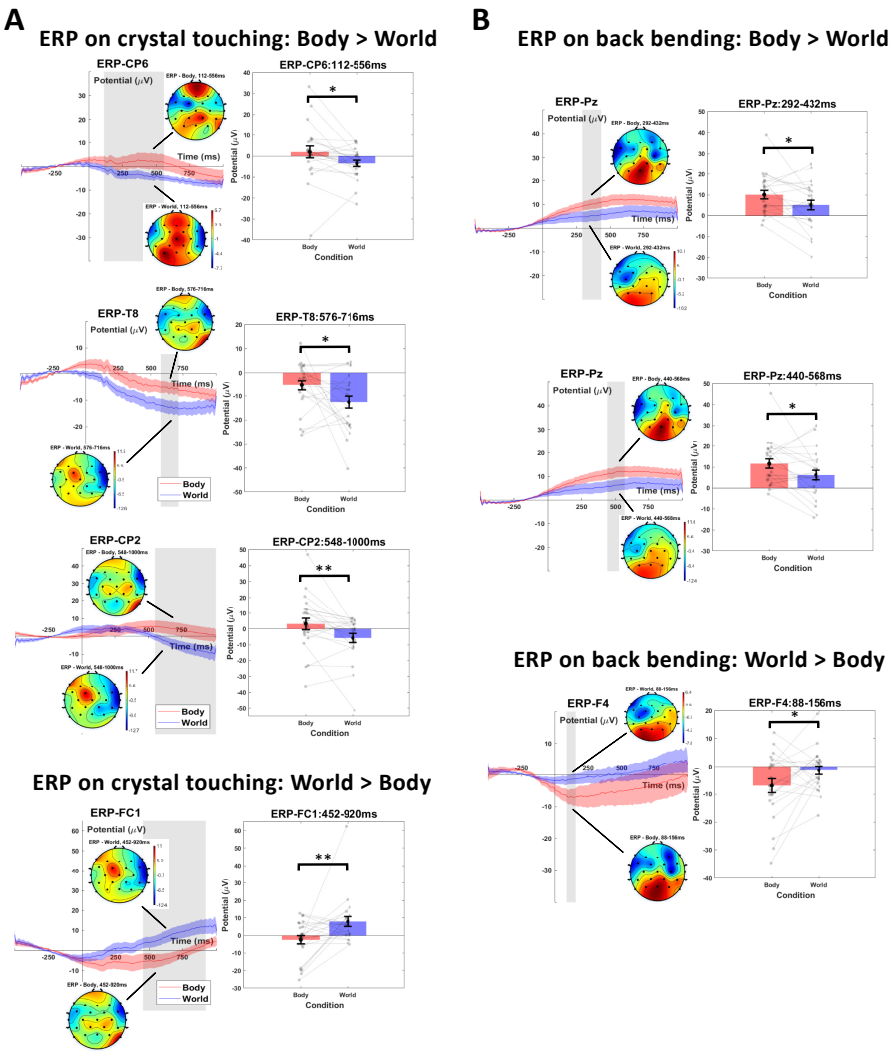

Figure 4.7: **ERPs distinguished with posterior-anterior discrepancy between body-model and world-model pain from early to late latencies after action events.**

(A) ERPs after crystal touching for body-model pain showed significantly greater amplitude than world-model pain at CP6, T8, and CP2 (top 3 rows), whilst with greater amplitude in world-model pain at FC1 (bottom-left). Given the moving right arm and the still left arm, the medial-bilateral drifting greater amplitude was also found for body-model pain, with greater ipsilateral amplitude shown in the topological maps, whilst greater contralateral amplitude was found for world-model pain. **(B)** ERPs after back bending for body-model pain showed significantly greater amplitude than world-model pain at Pz (top 2 rows), whilst with greater amplitude in world-model pain at F4 (bottom-right). Topological maps showed greater medial-centralised amplitude for body-model pain, whilst the greater ipsilateral amplitude for world-model pain after back-bending, which was apart from the greater contralateral amplitude after crystal-touching. For each ERP plot, topographic maps subject to pain conditions were given at each clustered time interval (highlighted as grey zones; see Section 4.2 for its achievement method) with significant mean potential difference shown in the corresponding bar plot. Error bars stand for standard errors. ** for $p < .01$ and * for $p < .05$.

pain (Figure 4.7A, top 3 rows) with higher ipsilateral ERP amplitude shown in the topological maps, whilst higher contralateral ERP amplitude was found for world-model pain (Figure 4.7A, bottom). Specifically, ERPs for body-model pain showed significantly higher amplitude than world-model pain during early latency at CP6 ($[112, 556]ms : \Delta = 5.45, t(22) = 2.42, p = .024$), whilst during late latency at T8 ($[576, 716]ms : \Delta = 7.18, t(22) = 2.18, p = .040$) and CP2 ($[548, 1000]ms : \Delta = 8.92, t(22) = 2.81, p = .010$) (Figure 4.7A, top 3 rows). Meanwhile, world-model pain ERPs showed significantly higher amplitude than body-model pain from mid to late latency at FC1 ($[452, 920]ms : \Delta = -10.43, t(22) = -2.85, p = .009$) (Figure 4.7A, bottom).

In response to back bending, body-model pain elicited stronger ERPs at the parietal electrode, whilst world-model pain elicited stronger ERPs at the frontal electrode. Topological maps showed medial-centralised ERP amplitude for body-model pain, whilst the ipsilaterally higher ERP amplitude for world-model pain after back-bending, which was apart from the contralaterally higher ERP amplitude after crystal-touching. ERPs for body-model pain showed significantly higher amplitude than world-model pain during mid latency at Pz ($[292, 432]ms : \Delta = 4.93, t(22) = 2.19, p = .039; [440, 568]ms : \Delta = 5.43, t(22) = 2.17, p = .041$) (Figure 4.7B, top 2 rows), whilst world-model pain ERPs showed significantly higher ERP amplitude than body-model pain at F4 ($[88, 156]ms : \Delta = -5.56, t(22) = -2.20, p = .039$) (Figure 4.7B, bottom).

In summary, body-model pain elicited stronger ERPs in right parietal and central regions (Pz, CP2, CP6, T8, C4, FC6), particularly following shock onset, crystal touching, and back bending. In contrast, world-model pain was associated with enhanced ERP responses at frontal and midline central electrodes (Cz, F4, FC1), with distinct spatial patterns emerging across different task events. The shock onset events implied the bilateral-medial potential discrepancy between body-model and world-model pain during mid to late latency, whilst the action events implied the asymmetric ERP amplitude differing between body-model and world-model pain. These results indicated an overall posterior-anterior difference of EEG-electrode ERPs between body-model and world-model pain, in accordance with our MRI study showing world-model pain error-based learning preference in centrofrontal regions (Figure 3.10A).

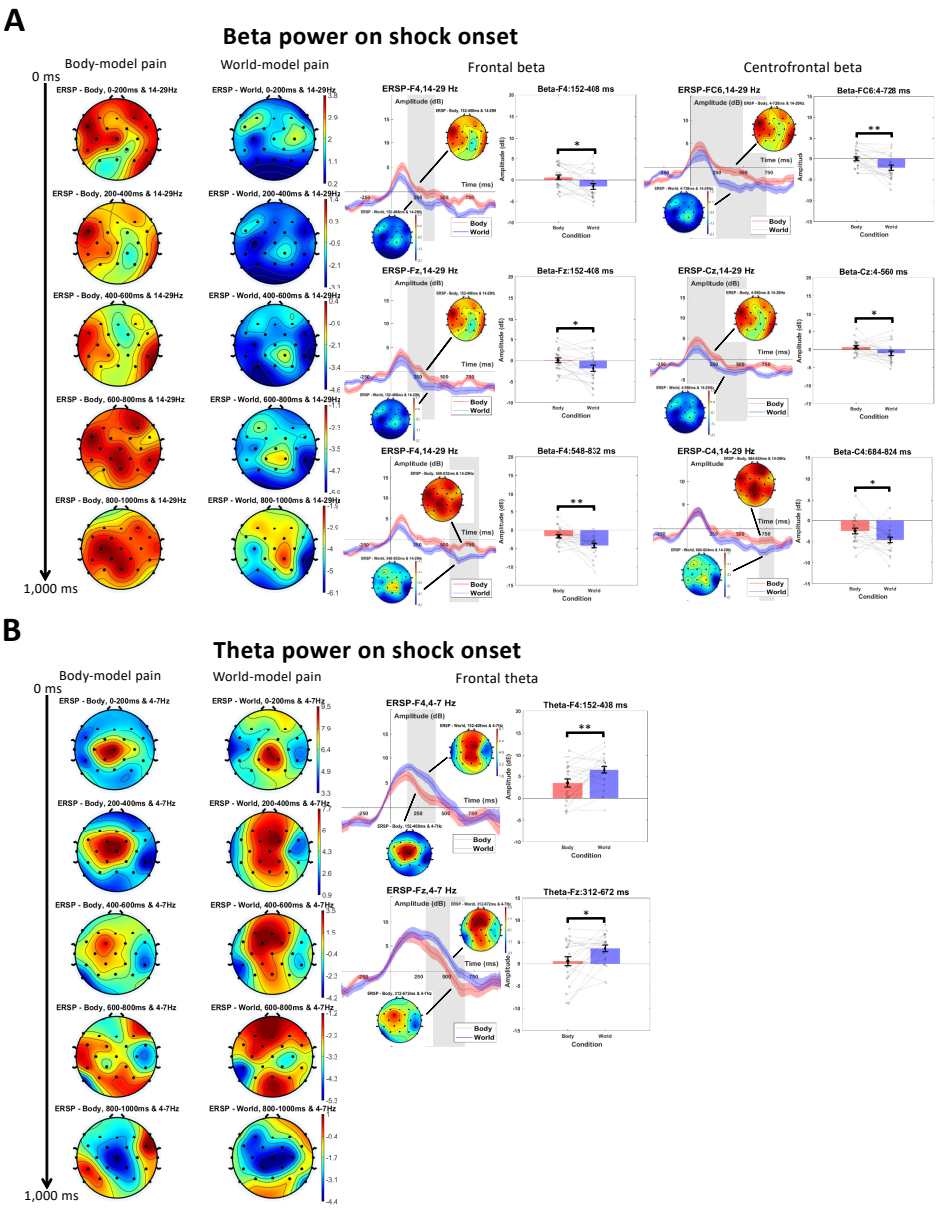

Figure 4.8: **TFR analysis showed body-model pain tracked by beta power activity and world-model pain tracked by frontal theta synchronisation.**

(A) Stronger beta activity was found for body-model pain against world-model pain across the central-frontal electrodes from early to late latency, with an overall beta power enhancement in medial electrodes during late latency for body-model pain shown in topological maps. **(B)** Frontal electrodes showed stronger theta synchronisation of world-model pain against body-model pain from early to mid latency after shock onset, with increasing medial theta synchronisation from early to mid latency for world-model pain shown in topological maps. Topological maps of whole-brain TFR were sorted in the latency order from 0 to 1,000 milliseconds. For each channel TFR plot, local topographic maps subject to pain conditions were given at each clustered time interval (highlighted as grey zones; see Section 4.2 for its achievement method) with significant mean frequency power difference shown in the corresponding bar plot. Error bars stand for standard errors. ** for $p < .01$ and * for $p < .05$.

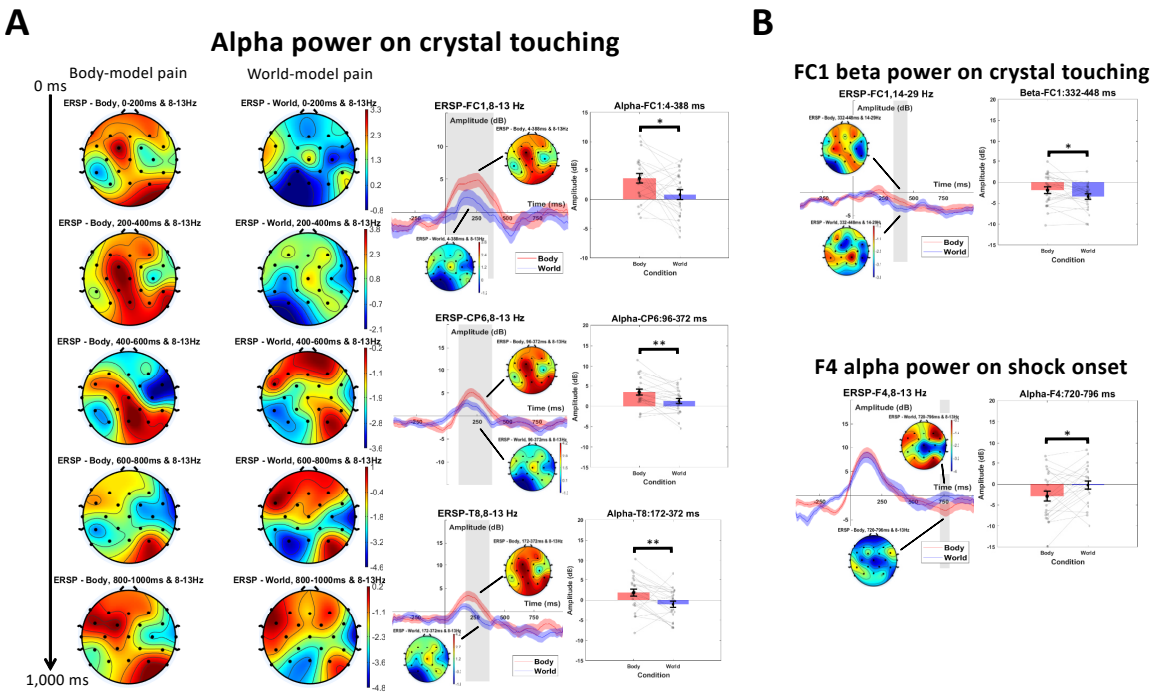

Figure 4.9: **TFR analysis showed stronger central alpha synchronisation for body-model pain during early to mid latency, whilst stronger frontal alpha synchronisation for world-model pain during late latency.**

(A) Stronger alpha synchronisation was found for body-model pain against world-model pain, ranging from temporal to centroparietal and centrofrontal electrodes. **(B)** FC1 also showed higher beta power activity of body-model pain against world-model pain during the mid latency after crystal touching event, whilst F4 showed higher alpha activation in world-model pain during late latency after shock onset. Topological maps of whole-brain TFR were sorted in the latency order from 0 to 1,000 milliseconds, given the specific ERP event and corresponding frequency band. For each channel TFR plot, local topographic maps subject to pain conditions were given at each clustered time interval (highlighted as grey zones; see Section 4.2 for its achievement method) with significant mean frequency power difference shown in the corresponding bar plot. Error bars stand for standard errors. ** for $p < .01$ and * for $p < .05$.

4.3.3 Beta and theta oscillations differentially tracked body-model and world-model pain processing

To further characterise the neural distinction between body-model and world-model pain, we conducted a time–frequency representation (TFR) analysis on electrodes that previously showed significant ERP differences. Electrode Fz was additionally included to specifically assess frontal theta activity linked to cognitive control and evaluative processing.

During shock onset, body-model pain was associated with significantly higher beta power across central-frontal electrodes from early to late latency periods (Figures 4.8A & 4.10A). Particularly, an overall beta power enhancement was found in medial electrodes during late latency for body-model pain shown in topological maps. Specifically, we found higher beta power for body-model pain during early latency at F4 ($[152, 408]ms : \Delta = 2.10, t(22) = 2.80, p = .010$) and Fz ($[152, 408]ms : \Delta = 1.93, t(22) = 2.69, p = .013$), early to mid latency at FC6 ($[4, 728]ms : \Delta = 2.16, t(22) = 3.66, p = .001$) and Cz ($[4, 560]ms : \Delta = 1.71, t(22) = 2.43, p = .023$), and late latency at F4 ($[548, 832]ms : \Delta = 2.58, t(22) = 3.16, p = .005$), Fz ($[652, 920]ms : \Delta = 1.54, t(22) = 2.68, p = .014$), Cz ($[664, 844]ms : \Delta = 2.28, t(22) = 3.13, p = .005$), and C4 ($[684, 824]ms : \Delta = 2.12, t(22) = 2.22, p = .037$). Besides, FC1 also showed higher beta power activity for body-model pain during the mid latency after crystal touching ($[332, 448]ms : \Delta = 1.53, t(22) = 2.18, p = .040$; Figure 4.9B, top & Figure 4.10A).

In comparison, the frontal electrodes showed overall stronger theta power synchronisation of world-model pain against body-model pain from early to mid latency (Figures 4.8B & 4.10A). An overall increasing medial theta synchronisation was found from early to mid latency for world-model pain shown in topological maps. Specifically, we found higher theta power for world-model pain during early latency at F4 ($[152, 408]ms : \Delta = -3.00, t(22) = -3.28, p = .003$) and mid latency at Fz ($[312, 672]ms : \Delta = -2.86, t(22) = -2.72, p = .012$). These suggested the distinct exteroceptive pain representations, as the centrofrontal beta power correlated more to sensorimotor-driven body-model pain chronically after shock onset, whilst frontal theta power correlated more to cognition-driven world-model pain from early to mid latency.

In response to crystal touching, we found an overall higher alpha power synchronisation of body-model pain against world-model pain across right centroparietal and centrofrontal electrodes (Figure 4.9A & 4.10B), with stronger medial alpha synchronisation for body-model pain from early to mid latency shown in topological maps. Specifically,

we found stronger alpha synchronisation during early latency at FC1 ($[4, 388]ms : \Delta = 2.78, t(22) = 2.66, p = .014$), CP6 ($[96, 372]ms : \Delta = 2.17, t(22) = 3.54, p = .002$), and T8 ($[172, 372]ms : \Delta = 2.87, t(22) = 3.00, p = .007$). On the other hand, we noticed a tendency of increasing frontal alpha synchronisation for world-model pain during late latency, relating to the higher F4 alpha activation in world-model pain during late latency after shock onset ($[720, 796]ms : \Delta = -2.66, t(22) = -2.18, p = .041$; Figure 4.9B, bottom & Figure 4.10A). In summary, body-model pain was associated with increased beta and alpha power at central and parietal sites, while world-model pain showed greater frontal theta and late alpha power, with effects varying by event and latency.

4.3.4 Exteroceptive pain context shaped EEG–perception coupling in illusory and sensory nociception

To further discover how the exteroceptive context of pain modulated nociception, we correlated ERPs and TFRs with subjective perception ratings after PCA (Figures 4.4 & 4.5). During shock onset, the ERP representations of the illusory nociception (Q-PC1) could be manipulated by body-model and world-model pain during mid latency. This was shown as the condition-ERP interaction effect during mid latency significantly at Cz ($[360, 676]ms : \beta = .395, t(43) = 2.617, p = .012, p_{FDR} = .025$) (Figure 4.11, left) and near-significantly at F4 ($[276, 372]ms : \beta = .197, t(43) = 2.257, p = .029, p_{FDR} = .059$) (Figure 4.11, right). Specifically, F4 and Cz showed higher ERP amplitude in the world-model pain, which tended to shift ERP representations (from body-model pain) on illusory nociception from negative to positive correlation (Cz $[360, 676]ms$: Body $k = -.147, r = -.322, p = .134$, World $k = .248, r = .429, p = .041, p_{FDR} = .082$; F4 $[276, 372]ms$: Body $k = -.121, r = -.498, p = .016, p_{FDR} = .031$, World $k = .076, r = .216, p = .322$), facilitating increasing frontal ERPs encoding enhanced illusory pain after shock onset. This implied a potential centro-frontal ERPs encoding illusory nociception facilitated by world-model pain (compared to body-model pain), which originated from the external world independent of bodily movement.

To discover how the exteroceptive modulation on nociception was encoded in brain signal oscillations, we further correlated TFRs with subjective perception ratings after shock onset (Figure 4.12A) and crystal touching (Figure 4.12B). For representing the illusory nociception, we found the discrepancy between body-model and world-model pain modulating beta and theta synchronisation/desynchronisation to represent illusory nociception. Specifically, we found central beta encoding modulated by exteroceptive pain at C4 (that itself favoured body-model pain in ERP and beta

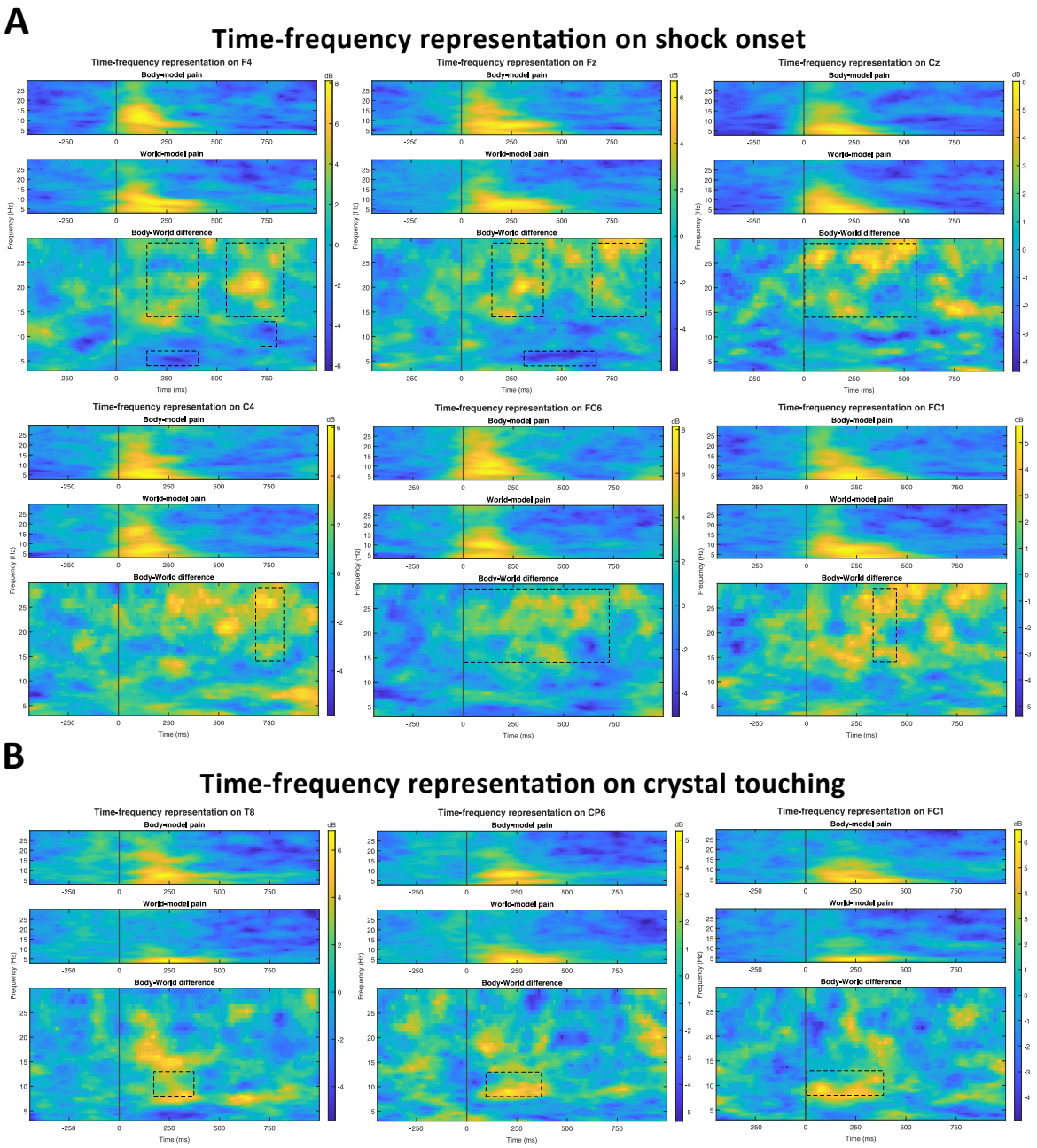

Figure 4.10: **Centrofrontal beta and theta showed significant difference between two exteroceptive pain conditions after shock onset, whilst alpha synchronisation was significantly stronger for body-model pain after touching crystals.**

Time-frequency representation (TFR) maps subject to experimental events ((**A**) for shock onset and (**B**) for crystal touching) were shown across all three frequency bands and the whole epoched time window. Each TRF map on an individual channel showed power signals for body-model pain (top), world-model pain (middle), and the heatmap of their difference (bottom) highlighted by dashed squares as the significant body-model vs world-model power difference used for comparisons in Figure 4.8, with brighter yellow as higher activity for body-model pain, whilst darker blue as higher activity for world-model pain.

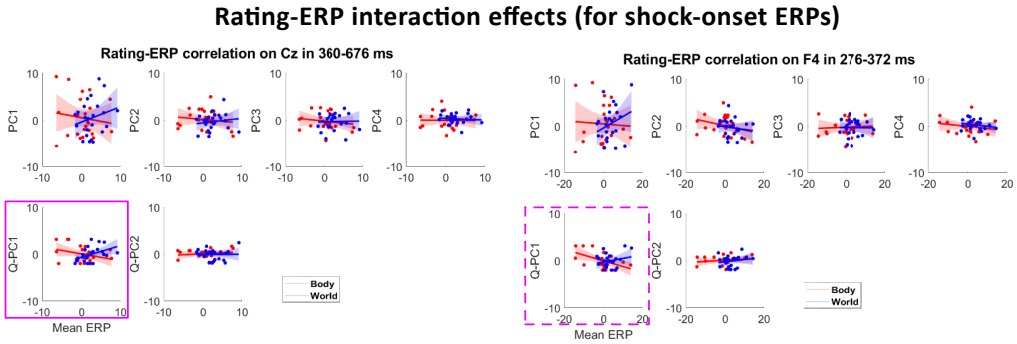

Figure 4.11: **Interaction effects between exteroceptive conditions and EEG potentials over subjective nociceptive ratings suggested exteroceptive pain manipulating illusory and sensory nociception with central and frontal potentials.**

Interaction effects found in Cz and F4 after shock onset, where the world-model pain tended to shift ERP representations (from body-model pain) on Q-PC1 from negative to positive correlation. Subjective ratings after PCA were respectively described as perception (except coldness; PC1), hot tonic pain (PC2), cold pain (PC3), tonic-phasic nociception discrepency (PC4), illusory pain (Q-PC1), and tonic-phasic illusory pain discrepency (Q-PC2) from Figures 4.4 & 4.5. Purple squares indicated significant interaction effects between the exteroceptive pain conditions and EEG signals (mean ERP or frequency power) on subjective PC ratings, with solid purple as $p_{FDR} < .05$ and dashed purple as $p_{FDR} < .1$. Filled shadows stand for 95% CI.

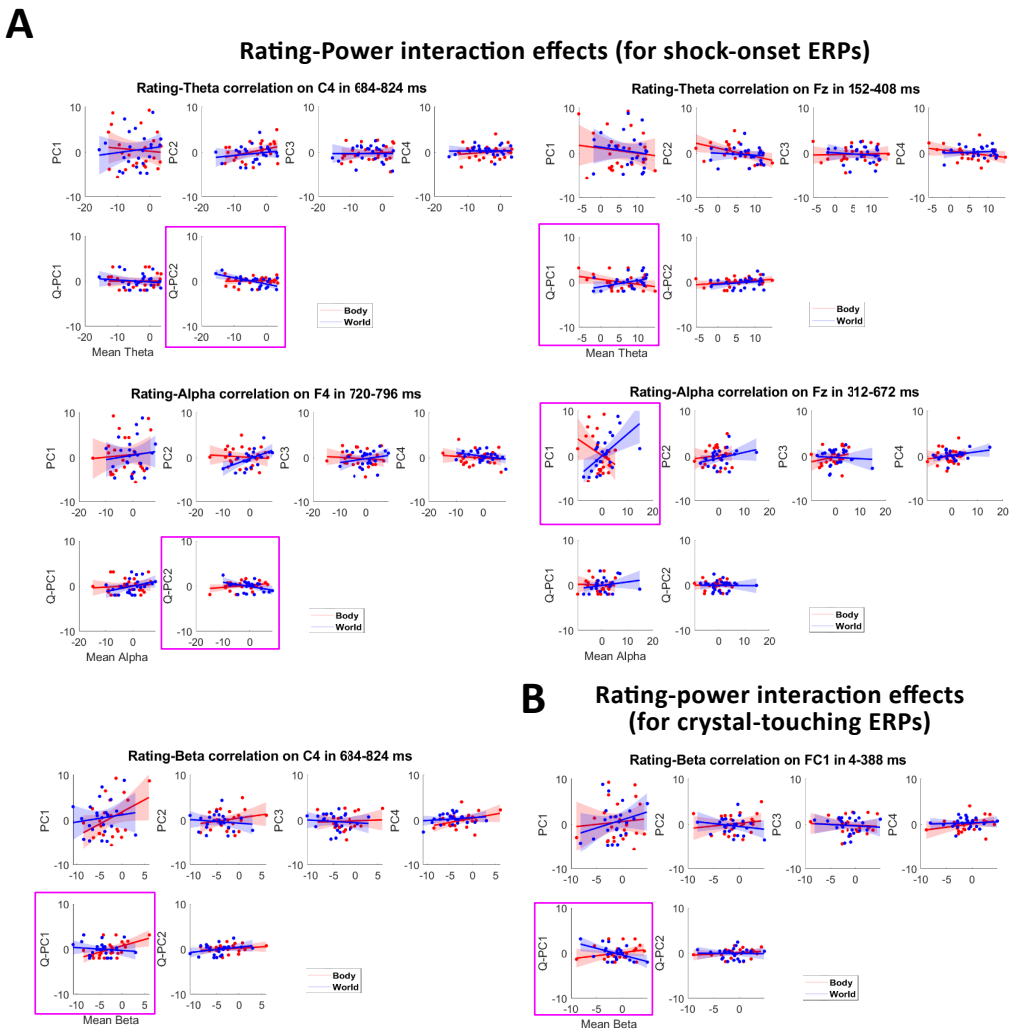

Figure 4.12: **Interaction effects between exteroceptive conditions and EEG frequency power bands over subjective nociceptive ratings suggested exteroceptive pain manipulating illusory and sensory nociception with central and frontal synchronisation.**

Interaction effects found in C4 theta and F4 alpha after shock onset implied the exteroceptive context of world-model pain tending to suppress theta and alpha representation on Q-PC2. Interaction effects found in C4 beta, FC1 beta, and Fz theta implied an exteroceptive discrepancy of pain on beta and theta synchronisation/desynchronisation representing Q-PC1. Interaction effects found in Fz alpha showed that the exteroceptive pain also modulated the sensory level of nociception with alpha synchronisation. Subjective ratings after PCA were respectively described as perception (except coldness; PC1), hot tonic pain (PC2), cold pain (PC3), tonic-phasic nociception discrepency (PC4), illusory pain (Q-PC1), and tonic-phasic illusory pain discrepency (Q-PC2) from Figures 4.4 & 4.5. Rating-power interaction effect scatterplots were sorted in the frequency band order, from theta to alpha and beta as the rows from top to bottom. Purple squares indicated significant interaction effects between the exteroceptive pain conditions and EEG signals (mean ERP or frequency power) on subjective PC ratings, with solid purple as $p_{FDR} < .05$. Filled shadows stand for 95% CI.

activation) during late latency after shock onset ($[684, 824]ms : \beta = -.345, t(43) = -2.394, p = .021, p_{FDR} = .042$; Figure 4.12A, top-right), driven by the body-model pain facilitating beta synchronisation to encode higher illusory pain (Body: $k = .282, r = .537, p = .008, p_{FDR} = .017$, World: $k = -.064, r = -.128, p = .559$). Similar results were also found at FC1 during early latency after touching crystal ($[4, 388]ms : \beta = -.434, t(43) = -2.768, p = .008, p_{FDR} = .017$; Figure 4.12B), driven by the world-model pain with beta desynchronisation encoding higher illusory pain (Body: $k = .138, r = .257, p = .237$, World: $k = -.296, r = -.524, p = .010, p_{FDR} = .021$). These were different from Fz theta in the mid latency $(152, 408]ms : \beta = .252, t(43) = 2.626, p = .012, p_{FDR} = .024$; Figure 4.12A, bottom-left), where body-model pain tended to suppress, whilst world-model pain tended to facilitate, theta synchronisation encoding higher illusory pain (Body: $k = -.011, r = .351, p = .101, p_{FDR} = .101$, World: $k = .141, r = .404, p = .056, p_{FDR} = .101$). These findings suggested that the body-model/world-model pain evoked central beta synchronisation/desynchronisation, whilst suppressing/facilitating theta oscillations, for encoding high illusory nociception.

For the tonic-phasic illusory pain discrepancy, we found the world-model pain tended to suppress theta and alpha oscillations for encoding illusory tonic pain against phasic shocks. Specifically, C4 (that favoured body-model pain in ERP and beta power) during late latency showed significant condition-theta effect on tonic-phasic illusory pain discrepancy ($[684, 824]ms : \beta = -.155, t(43) = -2.787, p = .008, p_{FDR} = .016$; Figure 4.12A, top-left), driven by the world-model pain inhibiting C4 theta power representing tonic menthol pain (Body: $k = .012, r = .071, p = .747$, World: $k = -.143, r = -.605, p = .002, p_{FDR} = .004$). Similar results were also found at F4 (that favoured world-model pain in ERP and theta power) during late latency, showing significant condition-alpha effect on tonic-phasic illusory pain discrepancy ($[720, 796]ms : \beta = -.144, t(43) = -2.386, p = .022, p_{FDR} = .043$; Figure 4.12A, top-middle), implying the exteroceptive context of pain manipulating frontal alpha to represent tonic menthol pain rather than phasic shocks (Body: $k = .500, r = .300, p = .164$, World: $k = -.095, r = -.375, p = .078, , p_{FDR} = .156$). These results implied that the exteroceptive context of world-model pain tended to suppress central theta and frontal alpha representation on illusory tonic vs phasic pain during late latency.

The exteroceptive context of pain also manipulated illusory pain at the sensory level. Fz in mid latency (that favoured world-model pain in theta power and body-model pain in beta power) showed significant condition-alpha interaction effect on sensory pain perception during mid latency ($[312, 672]ms : \beta = .862, t(43) = 3.190, p = .003, p_{FDR} = .011$; Figure 4.12A, bottom-middle), driven by the world-model pain

facilitating frontal alpha synchronisation representing nociceptive sensation (Body: $k = -.382, r = -.327, p = .128$, World: $k = .479, r = .496, p = .003, p_{FDR} = .005$). This implied the world-model pain modulating not only the illusory but also the sensory level of higher nociception with stronger frontal alpha synchronisation.

4.4 Discussion

We aimed to specify the connection between nociception and cognition underlying a fundamental distinction (Section 1.4). This referred to nociceptive problems from the external world (as exteroception) and in the internal body (as interoception) for cognitive modulation on nociception and sensation: How does exteroceptive pain learning provide an endogenous control on the persistent sensation and pain (which itself is not predictable for pain avoidance but yields a chronic sensation state appropriate to injury), so that to keep the cognitive functionality of danger avoidance? We assume that these two different cognitive essences of exteroceptive pain (i.e. body-model and world-model pain) could modulate interoceptive nociception or sensation, represented by certain brain activities. We designed the experiment to discover the interaction of persistent nociception and sensation modulated by these two pain cognition processes within an electroencephalography (EEG) study via an immersive Virtual Reality (iVR) game, allowing free-operant moving in the open arena. We showed discrepant representations of body-model and world-model exteroceptive pain in EEG as posterior-anterior ERPs, and as body-model beta and world-model frontal theta power. We further discovered the exteroceptive pain manipulating sensory and illusory nociception with centrofrontal synchronisation, showing exteroceptive context of world-model pain tending to surppress theta and alpha representation on illusory tonic vs phasic pain, an exteroceptive discrepancy of pain on beta and theta synchronisation/desynchronisation representing general illusory nociception, and exteroceptive pain modulating the sensory level of nociception with alpha synchronisation. These supported our hypothesis on the exteroception-modulated sensory and illusory nociception with electroencephalographic support.

4.4.1 Menthol cream mimicking interoceptive pain with rich nociceptive sensations

In the immersive Virtual Reality task, we used menthol cream to create a permanent tonic pain on participants' backs. This provided an imitated feeling of chronic pain, which was itself not learnable for avoiding external dangers with phasic pain, whereas

still apart from the real interoceptive pain as the sense of bodily internal states. On the other hand, the menthol cream could also provide various sorts of sensations (not merely pain), so that we could also find out how exteroceptive pain might modulate sensation with broader meanings, as designed in the questionnaire. If focusing on interoceptive pain modulation per se, one of the future research directions would be a clinical study on chronic pain patients with real interoceptive pain modulated (ideally suppressed) by manipulating the exteroceptive essence of phasic pain learning tasks as a potential chronic pain treatment.

Apart from menthol cream, other experimental methods like the cuff-pressure conditioning [Tong et al., 2025] and thermal/heat stimulators [Zhang et al., 2018] may also induce prolonged tonic pain mimicking interoception. Whereas these might not provide a nociceptive sensation as rich and ambiguous as menthol cream. When the static contact cooling occurred after the existing contact between menthol and human skin, the menthol cream might induce various sorts of nociceptive sensations known as the innocuous cold nociception (ICN) via the menthol-sensitive channel TRPM8 [Green and Jennifer, 2003, Green and Kate, 2007], which could also induce hyperalgesia [Namer et al., 2005]. As the menthol cream itself was not useful for learning shock avoidance during the iVR game, we could find out how exteroceptive pain might modulate nociceptive sensation and illusion with much broader meanings.

4.4.2 Exteroceptive context as cognitive 'control' on nociception

We have shown a centrofrontal region of EEG electrodes with potentials favouring cognition-oriented world-model than sensorimotor-oriented body-model pain, of which the frontal theta synchronisation correlated more to world-model pain and facilitated illusory pain, whilst the central theta desynchronisation suppressing illusory tonic pain against phasic pain. This was similar to previous cognitive control tasks showing high frontal potentials [Kopp et al., 1996, Brass et al., 2005, Hilgard et al., 2014, Kopp et al., 2020] and mid-frontal theta activity [Cavanagh and Frank, 2014, Zavala et al., 2018, Eisma et al., 2021, Tan et al., 2024], whilst our results suggested a cognitive exteroceptive 'control' also on sensory and illusory nociception, not only in frontal theta avtivity but also in centrofrontal alpha and beta synchronisation due to the bodily movement-focused body-model pain introduced as one of the exteroceptive contexts of pain. Other cognitive task studies linked central alpha and beta oscillation to somatosensory decision-making and sensorimotor processing [Witham et al., 2007,

Haegens et al., 2011, Quandt et al., 2013, Brickwedde et al., 2019], akin to the sensorimotor exteroception of our body-model pain from a 'body map' [Brecht, 2017].

TRF analysis also suggested a possible functional dissociation between frequency bands in the context of body-model and world-model pain. Beta activity in centrofrontal regions may be more closely associated with sensorimotor engagement and action-contingent pain processing, as seen in the body-model pain condition. In contrast, frontal theta activity may reflect cognitive evaluation and contextual processing, particularly during world-model pain. The observed patterns in alpha power further suggest a complex dynamic, potentially reflecting attentional modulation or arbitrative processing (Section 2.2.2) between two exteroceptive pain contexts. However, these interpretations remain speculative and warrant further investigation, particularly to disentangle whether these oscillatory differences were causal markers of cognitive modulation or downstream correlates of distinct attentional and affective states.

4.4.3 Disentangling exteroceptive context from contingency learning

The immersive Virtual Reality (iVR) task we used only required participants to keep the exteroceptive context of shocks in mind, i.e. either receiving the body-model painful shocks related to bodily movement (as back-bending) or the world-model painful shocks related to external spatial state information (with a visible red beacon indicating dangerous regions). Even though participants clearly knew how the danger arose, they could not choose alternative movement choices to actively avoid more shocks. This was different from the desktop-based VR task we used for our previous MRI study (Chapter 3), where participants learnt exteroceptive pain in the context of error-driven learning with multi-choice action space (Chapter 2). When introducing such error-driven contingency learning into exteroceptive pain in our iVR task design, there could be a question about whether the exteroceptive modulation on nociception would be aroused from either the contextual meaning of exteroception (body-model or world-model pain) or the error-driven learning per se. To exclude this potential learning effect for simplicity, we let participants know how the danger arose before the task so that only the different contextual meaning of exteroception was kept between the two experimental days. Further studies could focus on how the exteroceptive modulation within this error-driven contingency learning context might modulate sensory and illusory nociception with various and variant environmental contingency setups akin to the real world with dynamic dangers.

Chapter 5

General discussion

5.1 The conceptual basis of pain learning

We acquired new knowledge on the conceptual basis of pain, along with raising a new computational model on it, with the connection between nociception and cognition. This underlies two fundamental distinctions: i) nociceptive problems from the external world (as exteroception) and in the internal body (as interoception) for cognitive modulation on nociception, and ii) cognitive problems located within the external world (as world-model pain) and induced by the bodily actions (as body-model pain) from nociception for pain avoidance learning.

In Chapter 2, we first suggested a computational framework with RL incorporating a multidimensional representation of pain for learning pain avoidance during navigation in complex environments (Section 2.2). The model consisted of both model-based and model-free RL for body-model and world-model exteroceptive pain, as well as the arbitration between them for flexible learning (Section 2.2.2). **In Chapter 2 we showed the corresponding simulation results**, including its effective predictive error-driven avoidance learning over time, akin to human behaviour in different pain conditions, as well as flexible learning in varying pain conditions. We also showed the necessity of input pain perception as a multisensory vector embedded in a body map for effective pain avoidance, along with partially dysfunctionalised pain avoidance learning following the extent of perceptive bodily impairment. In this way, we suggested a single arbitrative learning system for human cognitive pain representation learning both external threatening environments and internal bodily integrity.

In Chapter 3, we then provided a large-scale neural architecture of dichotomic cognitive learning on pain for understanding the cognitive pain within complex environments: How and where a cognitive representation of

pain is encoded in the brain, and what function it serves in pain-oriented learning behaviour. To address this, **we designed a desktop-based virtual-navigation learning task in the MRI scanner with electric shocks delivered to both arms under different pain conditions that required the corresponding learning strategies (Section 3.2).** Participants showed significant aversive learning on pain avoidance performance (PAP) in both pain conditions. We then fitted our arbitrative RL model to human pain avoidance behaviours and the corresponding neuroimaging data during navigation. Model comparisons showed that model-free learning best fitted the body-model pain avoidance, whilst model-based learning best fitted the world-model pain avoidance. This aligned with our model assumptions that the corresponding pain learning models could best describe the behavioural data within the respective environmental conditions for pain avoidance learning. Model variables showed predictive TD-driven learning with arbitrative reliabilities. Neuroimaging results further imply a temporal-parietal difference in favour of different pain learning pathways, a dorsolateral-medial prefrontal difference in predictive and arbitrative pain learning, sensory-motor and insula activations with various pain-predictive signals, and an anterior cingulate circuit correlating with reliability-based arbitration. Our findings implied a cortical-subcortical activity around the prefrontal, cingulate, and insula cortex correlating with cognitive pain learning arbitration.

In Chapter 4, we further specified the connection between nociception and cognition underlying a fundamental distinction: Nociceptive problems from the external world (as exteroception) and in the internal body (as interoception) for cognitive modulation on nociception and sensation. We assume that these two different cognitive essences of exteroceptive pain could modulate interoceptive nociception or sensation, represented by certain brain activities. **We designed the experiment to discover the interaction of persistent nociception and sensation modulated by these two pain cognition processes within an electroencephalography (EEG) study via an immersive Virtual Reality (iVR) game allowing free-operant moving in the open arena (Section 4.2).** We showed discrepant representations of body-model and world-model exteroceptive pain in EEG as posterior-anterior ERPs, and as body-model beta and world-model frontal theta power. We further discovered the exteroceptive pain manipulating sensory and illusory nociception with centrofrontal synchronisation, showing exteroceptive context of world-model pain tending to suppress theta and alpha representation on illusory tonic vs phasic pain, an exteroceptive discrepancy of pain on beta and theta

synchronisation/desynchronisation representing general illusory nociception, and exteroceptive pain modulating the sensory level of nociception with alpha synchronisation. These supported our hypothesis on the exteroception-modulated sensory and illusory nociception with electroencephalographic support.

5.2 Novelty of the study

Our findings might have important implications for understanding how people learn from pain at a high cognitive level, as well as how the low-level perception could be manipulated by high-level cognitive modulation. Several main novel points of this study were worth mentioning. First, by linking nociception with cognitive learning, we raised a new perspective on the conceptual basis of pain focusing on both nociceptive and cognitive problems. Second, by linking high-level reinforcement learning algorithms with cognitive pain learning, we suggested a single arbitrative learning model for human cognitive pain representation learning both external threatening environments and internal bodily integrity. The hypothesis and the computational model obtained both neuroimaging support for dichotomic exteroceptive pain learning and electroencephalographic support for the exteroception-modulated sensory and illusory nociception. Therefore, it could provide enlightenment on designing artificial intelligence algorithms incorporating human cognitive abilities in pain cognition and empathy. Last but not the least, the iVR technique we used enabled an immersive pain learning environment for human participants to conduct natural bodily movement for learning realistic pain avoidance behaviour. VR has already been recently used for phasic and tonic pain valuation [Hewitt et al., 2025, Tong et al., 2025] as well as chronic pain management and treatment [Li et al., 2011, Birckhead et al., 2021, Goudman et al., 2022, Teh et al., 2024]. Along with our new perspective on the conceptual pain and its treatment, this study would greatly benefit translational research for patients suffering pain, e.g. by helping them relieve chronic pain or other interoceptive noxious sensations given the appropriate high-level cognitive learning task in the immersive Virtual Reality.

5.3 Exteroceptive & interoceptive pain interaction

Apart from exteroceptive phasic pain encoding potential dangers of the external world, the pain information of 'how does the brain realise that the body is injured' could also be inferred and affected from an interoceptive persistent (or chronic) pain that itself

is not learnable [Di Lernia et al., 2016]. These two pathways might share bidirectional interactions with each other. The persistent interoceptive pain drives an endogenous control on the sensation of co-occurring exteroceptive pain (e.g. in the conditioned pain modulation [Kennedy et al., 2016]), whilst learning the exteroceptive pain for avoidance learning requires an inhibition on interoceptive persistent pain that is abnormal within chronic pain conditions [Staud, 2012]. The additional interceptive persistent pain in the body may sensitise/inhibit these exteroceptive pain cognition processes within the relative body part to increase/decrease the corresponding pain avoidance behaviour [Moriarty et al., 2011], cf. conditioned pain modulation on nociception [Kennedy et al., 2016]. On the other hand, efficient exteroceptive pain learning and avoidance needs a certain level of endogenous control on interoceptive persistent pain to suppress the unlearnable 'background noise', cf. cognitive dysfunctionality when suffering chronic pain [Moriarty et al., 2011, Staud, 2012]. Our findings showed a cognitive distinction in the exteroceptive pain with their manipulation on perceptive nociception and sensation under the corresponding pain conditions relative to different cognitive essence of phasic pain, becoming a potential research direction for treating specific cognitive function performance impairment specific to a certain pain condition under chronic pain [Moriarty et al., 2011].

5.4 Conditioned pain modulation not restricted to the reported nociception

In both of the experiments, no reported subjective pain intensity showed a significant difference under experoceptive context modulation (i.e. body-model vs world-model pain learning) due to the recalibration at the start of each session. That is, the classical conditioned pain modulation (CPM) effect [Kennedy et al., 2016] was prevented from inhibiting phasic pain perception. Neither did the menthol cream (as tonic pain) show a significant CPM effect on inhibiting electrical shock perception (as phasic pain) in the iVR task. We rather showed the behavioural changes subject to the learnt experoceptive context, as well as exteroception-modulated sensory and illusory nociception encoded in the electroencephalographic signals. These implied a potential endogenous control not manifested in the subjective report on pain perception, but rather in learning behaviours and brain signal encoding. Further research may discover if a classical CPM would inhibit phasic shock perception (e.g. by not doing recalibration during the whole task), whether these exteroceptive modulations would be influenced by or still be steady against the descending phasic pain perception. This could then

indicate whether the classical CPM and exteroceptive modulations share overlapped descending pathways for the endogenous control.

5.5 Tradeoff between pain avoidance learning and task complexity

Our computational model aimed to support exteroceptive pain avoidance learning within complex and dynamic environments, facilitated by the MF-MB arbitration. Whereas in our two experiments, participants were all informed of the pain condition they were about to suffer before the task block. The pain condition was also kept universal within a single task block, rather than dynamically changing during navigation, akin to 2-stage tasks on MF-MB tradeoff learning without pain [Lee et al., 2014]. These were so decided due to the first pilot experiments (not included in this dissertation), where participants failed to show efficient pain avoidance learning if not informed of the corresponding pain condition. It might be because of the shocks providing frequent and uncertain (note the non-dangerous actions might still give rise to shocks as a contingency design) noxious feelings, as many pilot participants complained after the task, potentially resulting in a learnt helplessness [Seligman, 1972] much faster than other cognitive learning tasks without pain. Thus, the experiments might not directly incorporate an arbitrative mechanism between two exteroceptive pains within one-time navigation, which we only showed from data fitting to the arbitrative learning model (Chapter 3). This could be a general challenge in cognitive pain avoidance learning studies with more complex and dynamic dangers. We hoped learning pain within an immersed environment (e.g. with immersive Virtual Reality) would provide a more realistic pain suffering scenario with natural body movement as a potential future direction on pain study design [Tong et al., 2025]. Though our current iVR task only aimed to show a primary exteroception-interoception connection, and to prevent the nociception from chronic menthol existence from interfering with uncorrespondent phasic exteroceptive pain, we even separated pain conditions in different experimental days.

5.6 Limitations and future directions

Our computational model aimed to support exteroceptive pain avoidance learning within complex and dynamic environments. Whereas the computational model we built up for efficient pain avoidance behaviour learning accepted shocks at most on 6

discrete bodily parts simultaneously. The simulation results implied the bodily parts as a necessary multidimensional input for pain prediction rather than the single scalar input in traditional reward-oriented tasks. Though the discrete 6-electrode input in our simulation setup was still too simplified to consider a topological somatosensory pain perception in a 2-dimensional surface space covering the whole body, and one of the future directions would be to incorporate a continuous bodily topological space as a continuous 'body map' for a more realistic pain avoidance learning [Mancini et al., 2012, 2014].

The successor representation we used for model-based learning addressed a factorisation between spatial-state predictive coding and pain-state evaluation, which enabled flexible learning across multiple environments by explicitly encoding and storing predictive relationships among spatial states as a 'world map' [Tolman, 1948]. It was also suggested as a hippocampal mechanism for between-state predictive learning without explicit planning [Stachenfeld et al., 2017, Geerts et al., 2020], where the hippocampus had been reported showing neuroimaging activations in various reward-oriented model-based learning tasks [Bornstein and Daw, 2012, Sebold et al., 2017]. Future research would consider more complex environmental setups with multiple scenarios, so that to discover pain avoidance learning under large-scale navigation across sensory-rich environments.

A core idea of using an arbitrator over multiple learning pathways in a single system was to handle the varying environmental conditions for learning [Lee et al., 2014], aligned with the human learning capability on flexible shifting between 'automatic' habitual behaviour and 'deliberative' goal-directed behaviour for evaluating actions [Balleine and Dickinson, 1998, Dayan and Daw, 2008, Drummond and Niv, 2020]. Whereas our virtual-navigation tasks did not implement such environmental variations, as raising the task difficulty might suppress their learning efficiency within the limited time due to the potential quick learnt helplessness. Participants were always informed of the current pain condition to control the task difficulty, though this resulted in worse data fitting results of the arbitrative model compared to the pure MF or pure MB model with explicit prior knowledge in model selection in our MRI study. We hoped learning pain within immersive Virtual Reality (iVR) would provide a more realistic pain suffering scenario with natural body movement as a potential future direction on pain study design [Tong et al., 2025], though our current iVR task only aimed to show a primary exteroception-interoception connection thus yet to incorporate dynamic dangers (that our computational model first aim to incorporate) at this stage, and to prevent the nociception on chronic menthol existence from interfering

with uncorrespondent phasic exteroceptive pain we even separated pain conditions in different experimental days. Even though the tradeoff between pain learning and task difficulty could be a general challenge for pain learning studies, future research would focus on the model simulation and experimental design with variant pain conditions in a single block, so that to show the necessity of the arbitrative mechanism for pain avoidance learning on complex and varying dangers.

The immersive Virtual Reality (iVR) task we used only required participants to keep the exteroceptive context of shocks in mind, i.e. either receiving the body-model painful shocks related to bodily movement (as back-bending) or the world-model painful shocks related to external spatial state information (with a visible red beacon indicating dangerous regions). Even though participants knew how the danger arose, they could not choose alternative movement choices to actively avoid more shocks. This was different from the desktop-based VR task we used for our MRI study, where participants learnt exteroceptive pain in the context of error-driven learning with a multi-choice action space. When introducing such error-driven contingency learning into exteroceptive pain in our iVR task design, there could be a question about whether the exteroceptive modulation on nociception would be aroused from either the contextual meaning of exteroception (body-model or world-model pain) or the error-driven learning per se. To exclude this potential learning effect for simplicity, we let participants know how the danger arose before the task so that only the different contextual meaning of exteroception was kept between the two experimental days. Further studies could focus on how the exteroceptive modulation within this error-driven contingency learning context might modulate sensory and illusory nociception with various and variant environmental contingency setups akin to the real world with dynamic dangers, meanwhile as a potential direction for treating specific cognitive function performance impairment specific to a certain pain condition under chronic pain [Moriarty et al., 2011].

In the iVR task, we used menthol cream to create a permanent tonic pain on participants' backs. This provided an imitated feeling of chronic pain, which was itself not learnable for avoiding external dangers with phasic pain, whereas still apart from the real interoceptive pain as the sense of bodily internal states. On the other hand, the menthol cream could also provide various sorts of sensations (not merely pain), so that we could also find out how exteroceptive pain might modulate sensation with broader meanings, as designed in the questionnaire. If focusing on interoceptive pain modulation per se, one of the future research directions would be a clinical study on chronic pain patients with real interoceptive pain modulated (ideally suppressed) by

manipulating the exteroceptive essence of phasic pain learning tasks as a potential chronic pain treatment.

Conclusions

We acquired new knowledge on the conceptual basis of pain, linking nociception with cognition. We suggested a single arbitrative learning model for human cognitive pain representation learning both external threatening environments and internal bodily integrity, with neuroimaging support to dichotomic exteroceptive pain learning and electroencephalographic support to the exteroception-modulated sensory and illusory nociception. Our findings might have important implications for understanding how people learn from pain at a high cognitive level, as well as how the low-level perception could be manipulated by high-level cognitive modulation. It also provided potential enlightenment on designing artificial intelligence with human cognitive abilities on pain cognition, e.g. to provide efficient self-protection and interaction among agents under rich and complex environments. In addition, such a novel approach could provide a new perspective on pain treatment, benefiting translational research for patients suffering pain, e.g. by helping them relieve chronic pain or other interoceptive noxious sensations, given the appropriate high-level cognitive learning task in immersive Virtual Reality.

Acknowledgements

This work is funded by the Wellcome Trust (214251/Z/18/Z). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript. The author has declared that no competing interests exist.

This dissertation comprised 3 main studies in my four and a half years' DPhil life, all starting from 2020-2021 as the first year during COVID. Very little progress could be made under the epidemic at that time, except conceiving research ideas and taking online courses, along with the limited socialisation as a verily hard time. Stepping into the second year after COVID, most of the time was spent in modifying experimental designs and setting up experimental kits for the MRI study, with the design never done by others before, which seemed to be much more complex than many of the other classic pain-related experiments and beyond my expectation of the timeline. Meanwhile, many and many incidents happened, which are not necessitated to mention here, piece after piece and time after time in my life. Even when the study finally seemed to go smoothly from the third year, other trivial matters in my life still kept occurring and haunting me, blurring my memory of the day when I got off the bus from Heathrow to Oxford for the first time.

Time passed by and many things have changed since then, not only my memories but also my worldview and probably even my personality. One might call it 'growing up', which nonetheless always incorporates irreversibility, followed by memory ambiguity. It has taken much time and effort for me to reconcile with myself, but until then did I realise that this dissertation seemed to be the only words I could read to trace my memory back to these wavering years. There might be so many things I have frozen and dropped away from my heart, whereas this dissertation was accomplished with the help of so many generous, tremendous, and prosperous people who are indeed worth memorising.

I would first like to express my gratitude to my first supervisor, Prof. Ben Seymour. The first time we met was in Japan, and coincidentally, just a few days before the start of COVID. Many thanks to Ben for providing me the opportunity to conduct this study in Oxford, applying neuro-computational modelling on human pain. As an expert in pain cognition, he was the one leading me to an interest in pain as a unique, complex, and charming concept of understanding human cognition in a novel way. He was also verily supportive and helpful to my study, providing various and tremendous help across my DPhil years.

Then it comes to my gratitude to my second supervisor, Dr. Laurence Hunt. As an expert in decision-making cognition, he was verily supportive and kind to me, giving me abundant precious suggestions on my MRI study and on my scientific paper writing. I am also grateful to Laurence for having me in his journal club, discussing papers and projects with so many brilliant people in his lab. I would also look forward to his further precious comments and suggestions on my paper manuscripts for potential journal submissions.

In Ben's lab I have many fabulous people as my lab colleagues, and I would like to show my gratitude to some of them also providing tremendous help to my study. Dr. Shaungyi Tong and Pranav Mahajan have given many fabulous thoughts and suggestions on my computational modelling, and beyond that we keep discussing computational models and other AI-related hi-tech topics during leisure time. Dr. Shaungyi Tong has also given massive technical help in the VR study with EEG. Dr. Danielle Hewitt has provided enormous help in my experimental design, experiment performance, and paper writing. The EEG study could not be finalised as now without her expertise and generosity. Dr. Suyi Zhang has helped in conceiving research ideas at the very early stage of the study, and also in providing shock-related experimental kits from her previous studies. Other former and current lab members including Dr. Rachel Crockett, Dr. Wako Yoshida, Ruohan Liu, and Sarah Schreiber have provided many useful suggestions on my study as well.

I would also like to spare my gratitude to other scholars who gave precious suggestions on the study or helped in the experimental setup. Prof. Jill O'Reilly and Dr. Katja Wiech have given many supportive suggestions

during the transfer of status and confirmation of status viva. Particular gratitude is also given to Prof. Susanne Becker and Dr. Miriam Klein-Flügge for their constructive suggestions and comments during the DPhil viva. Prof. Sang Wan Lee has orally given encouraging comments on my study, with his neuoimaging study on MB-MF arbitrative modelling having enlightened my computational modelling ideas. Dr. Andrew Segerdahl helped a lot with the capsaicin usage, though it was finally replaced with menthol cream in the EEG study design. Dr. Andrew Rudgewick-Brown provided many practical helps in MRI experimental kits organisation and manufacturing, with Dr. Jon Campbell, Dr. Michael Sanders, and Dr. Sebastian Rieger as radiographers offering great help for conducting my MRI experiment in practice. Many other scholars could receive my gratitude for shedding enlightening thoughts and comments on my study posters during the symposiums I have attended.

Besides scholars, I would like to thank all participants who came to attend my study, suffering painful shocks and menthol just for their curiosity about the cognitive pain experiment and paid game-playing (particularly with Virtual Reality) experience. I would also like to mention general appreciation to New College, University of Oxford, Nuffield Department of Clinical Neurosciences, Wellcome Centre for Integrative Neuroimaging, and Institute of Biological Engineering, providing me the place to conduct my DPhil study and experiments in practice.

Last but not the least, I would like to specify my profound gratitude to my family and friends who indeed have helped or even supported me more or less during my DPhil years. Particularly I am grateful to my father Dr. Ahorn Yan for partially sponsoring my DPhil study and life expenses with both financial and spiritual support, as well as my brother Tom-Clarence Yan for providing me with much remote encouragement and approval.

Time passed by and many things would still keep changing in the next few years, and probably more would have been frozen and dropped away from my heart. Whereas the memorable time spent on this study and with fabulous people mentioned above would never fade from my memories, accompanying me to my next stage of the wavering life.

Bibliography

[1] H. Akaike. A new look at the statistical model identification. *IEEE Transactions on Automatic Control*, 19(6):716–723, 1974. doi:10.1109/TAC.1974.1100705.

[2] D. Anchisi and M. Zanon. A bayesian perspective on sensory and cognitive integration in pain perception and placebo analgesia. *PLOS ONE*, 10(2):e0117270, 2015. doi:10.1371/journal.pone.0117270.

[3] A. V. Apkarian and C. J. Hodge. A dorsolateral spinothalamic tract in macaque monkey. *Pain*, 37(3):323–333, 1989. doi:10.1016/0304-3959(89)90198-X.

[4] B. W. Balleine and A. Dickinson. Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. *Neuropharmacology*, 37(4): 407–419, 1998. doi:10.1016/S0028-3908(98)00033-1.

[5] M. T. Banich, P. Dukes, and D. Caccamise. *Generalization of knowledge: Multidisciplinary perspectives.* Psychology Press, Hove, 2010.

[6] A. Banino, C. Barry, B. Uria, C. Blundell, T. Lillicrap, P. Mirowski, A. Pritzel, M. J. Chadwick, T. Degris, J. Modayil, G. Wayne, H. Soyer, F. Viola, B. Zhang, R. Goroshin, N. Rabinowitz, R. Pascanu, C. Beattie, S. Petersen, A. Sadik, S. Gaffney, H. King, K. Kavukcuoglu, D. Hassabis, R. Hadsell, and D. Kumaran. Vector-based navigation using grid-like representations in artificial agents. *Nature*, 557(7705):429–433, 2018. doi:10.1038/s41586-018-0102-6.

[7] T. E. J. Behrens, T. H. Muller, J. C. R. Whittington, S. Mark, A. B. Baram, K. L. Stachenfeld, and Z. Kurth-Nelson. What is a cognitive map? organizing knowledge for flexible behavior. *Neuron*, 100(2):490–509, 2018. doi:10.1016/j.neuron.2018.10.002.

[8] Y. Benjamini and Y. Hochberg. Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society. Series B (Methodological)*, 57(1):289–300, 1995. doi:10.2307/2346101.

[9] B. Birckhead, S. Eberlein, G. Alvarez, R. Gale, T. Dupuy, K. Makaroff, G. Fuller, X. Liu, K.-S. Yu, J. T. Black, M. Ishimori, S. Venuturupalli, J. Tu, T. Norris, M. Tighiouart, L. Ross, K. McKelvey, M. Vrahas, I. Danovitch, and B. Spiegel. Home-based virtual reality for chronic pain: protocol for an nih-supported randomised-controlled trial. *BMJ Open*, 11(6), 2021. doi:10.1136/bmjopen-2021-050545.

[10] A. M. Bornstein and N. D. Daw. Dissociating hippocampal and striatal contributions to sequential prediction learning. *European Journal of Neuroscience*, 35 (7):1011–1023, 2012. doi:10.1111/j.1460-9568.2011.07920.x.

[11] M. Botvinick, J. X. Wang, W. Dabney, K. J. Miller, and Z. Kurth-Nelson. Deep reinforcement learning and its neuroscientific implications. *Neuron*, 107 (4):603–616, 2020. doi:10.1016/j.neuron.2020.06.014.

[12] M. Brass, M. Ullsperger, T. Knösche, D. Cramon, and N. Phillips. Who comes first? the role of the prefrontal and parietal cortex in cognitive control. *Journal of Cognitive Neuroscience*, 17(9):1367–1375, 2005. doi:10.1162/0898929054985400.

[13] M. Brecht. The body model theory of somatosensory cortex. *Neuron*, 94(5): 985–992, 2017. doi:10.1016/j.neuron.2017.05.018.

[14] M. Brickwedde, M. Krüger, and H. Dinse. Somatosensory alpha oscillations gate perceptual learning efficiency. *Nature Communications*, 10(1):263, 2019. doi:10.1038/s41467-018-08012-0.

[15] C. A. Brown, B. Seymour, W. El-Deredy, and A. K. Jones. Confidence in beliefs about pain predicts expectancy effects on pain perception and anticipatory processing in right anterior insula. *PAIN*, 139(2):324–332, 2008. doi:10.1016/j.pain.2008.04.028.

[16] S. Carey. Précis of the origin of concepts. *Behavioral and Brain Sciences*, 34(3): 113–124, 2011. doi:10.1017/s0140525x10000919.

[17] J. F. Cavanagh and M. J. Frank. Frontal theta as a mechanism for cognitive control. *Trends in Cognitive Sciences*, 18(8):414–421, 2014. doi:10.1016/j.tics.2014.04.012.

[18] A. D. Craig. Distribution of brainstem projections from spinal lamina i neurons in the cat and the monkey. *J Comp Neurol*, 361(2):225–248, 1995. doi:10.1002/cne.903610204.

[19] A. D. Craig. How do you feel? interoception: the sense of the physiological condition of the body. *Nature Reviews Neuroscience*, 3(8):655–666, 2002. doi:10.1038/nrn894.

[20] A. D. Craig. Interoception: the sense of the physiological condition of the body. *Current Opinion in Neurobiology*, 13(4):500–505, 2003. doi:10.1016/s0959-4388(03)00090-4.

[21] A. D. Craig. Retrograde analyses of spinothalamic projections in the macaque monkey: Input to the ventral lateral nucleus. *The Journal of Comparative Neurology*, 508(2):315–328, 2008. doi:10.1002/cne.21672.

[22] A. D. Craig, M. C. Bushnell, E. T. Zhang, and A. Blomqvist. A thalamic nucleus specific for pain and temperature sensation. *Nature*, 372(6508):770–773, 1994. doi:10.1038/372770a0.

[23] N. D. Daw, Y. Niv, and P. Dayan. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nature Neuroscience*, 8(12):1704–1711, 2005. doi:10.1038/nn1560.

[24] P. Dayan. Improving generalization for temporal difference learning: The successor representation. *Neural Computation*, 5(4):613–624, 1993. doi:10.1162/neco.1993.5.4.613.

[25] P. Dayan and N. D. Daw. Decision theory, reinforcement learning, and the brain. *Cognitive, Affective, & Behavioral Neuroscience*, 8(4):429–453, 2008. doi:10.3758/CABN.8.4.429.

[26] A. Delorme and S. Makeig. Eeglab: an open-source toolbox for analysis of eeg dynamics. *Journal of neuroscience methods*, 134(1):9–21, 2004. doi:10.1016/j.jneumeth.2003.10.009.

[27] D. Di Lernia, S. Serino, P. Cipresso, and G. Riva. Ghosts in the machine. interoceptive modeling for chronic pain treatment. *Frontiers in Neuroscience*, 10:314, 2016. doi:10.3389/fnins.2016.00314.

[28] D. Di Lernia, S. Serino, and G. Riva. Pain in the body. altered interoception in chronic pain conditions: A systematic review. *Neuroscience & Biobehavioral Reviews*, 71:328–341, 2016. doi:10.1016/j.neubiorev.2016.09.015.

[29] J. O. Dostrovsky and A. D. Craig. Ascending projection systems. In S. B. McMahon and M. Koltzenburg, editors, *Wall and Melzack's Textbook of Pain*, page 187–203. Elsevier Health Sciences, Churchill Livingstone, 2006.

[30] K. Doya, K. Samejima, K. Katagiri, and M. Kawato. Multiple model-based reinforcement learning. *Neural Computation*, 14(6):1347–1369, 2002. doi:10.1162/089976602753712972.

[31] N. Drummond and Y. Niv. Model-based decision making and model-free learning. *Current Biology*, 30(15):R860–R865, 2020. doi:10.1016/j.cub.2020.06.051.

[32] B. Ehinger and O. Dimigen. Unfold: An integrated toolbox for overlap correction, non-linear modeling, and regression-based eeg analysis. *PeerJ*, 7:e7838, 2019. doi:10.7717/peerj.7838.

[33] J. Eisma, E. Rawls, S. Long, R. Mach, and C. Lamm. Frontal midline theta differentiates separate cognitive control strategies while still generalizing the need for cognitive control. *Scientific Reports*, 11(1):14641, 2021. doi:10.1038/s41598-021-94162-z.

[34] F. Fardo, R. Auksztulewicz, M. Allen, M. J. Dietz, A. Roepstorff, and K. J. Friston. Expectation violation and attention to pain jointly modulate neural gain in somatosensory cortex. *NeuroImage*, 153:109–121, 2017. doi:10.1016/j.neuroimage.2017.03.041.

[35] K. Farkas. Constructing a world for the senses. In U. Kriegel, editor, *Phenomenal Intentionality*, pages 99–115. Oxford University Press, Oxford, 2013. doi:10.1093/acprof:oso/9780199764297.

[36] S. Fazeli and C. Büchel. Pain-related expectation and prediction error signals in the anterior insula are not related to aversiveness. *Journal of Neuroscience*, 38 (29):6461–6474, 2018. doi:10.1523/JNEUROSCI.0671-18.2018.

[37] J. A. Fodor and Z. W. Pylyshyn. Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28(1):3–71, 1988. doi:10.1016/0010-0277(88)90031-5.

[38] K. Friston. A theory of cortical responses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 360(1456):815–836, 2005. doi:10.1098/rstb.2005.1622.

[39] M. Fumagalli and R. Ferrario. Representation of concepts in ai: Towards a teleological explanation. In *Joint Ontology Workshops*, 2019.

[40] S. Gallagher and D. Zahavi. *The phenomenological mind*. Routledge, London, 2012.

[41] J. García and F. Fernández. A comprehensive survey on safe reinforcement learning. *The Journal of Machine Learning Research*, 16(1):1437–1480, 2015. doi:10.5555/2789272.2886795.

[42] J. P. Geerts, F. Chersi, K. L. Stachenfeld, and N. Burgess. A general model of hippocampal and dorsal striatal learning and decision making. *Proceedings of the National Academy of Sciences*, 117(49):31427–31437, 2020. doi:10.1073/pnas.2007981117.

[43] F. A. Geldard. *The Human Senses*. John Wiley and Sons, Inc., New York, NY, 1972. doi:10.2307/1421863.

[44] S. J. Gershman. The successor representation: Its computational logic and neural substrates. *Journal of Neuroscience*, 38(33):7193–7200, 2018. doi:10.1523/JNEUROSCI.0151-18.2018.

[45] S. Geuter, S. Boll, F. Eippert, and C. Büchel. Functional dissociation of stimulus intensity encoding and predictive coding of pain in the insula. *eLife*, 6, 2017. doi:10.7554/elife.24770.

[46] J. Gläscher, N. Daw, P. Dayan, and J. P. O'Doherty. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, 66(4):585–595, 2010. doi:10.1016/j.neuron.2010.04.016.

[47] I. Goodfellow, Y. Bengio, and A. Courville. *Deep Learning*. MIT Press, Cambridge, MA, 2016.

[48] L. Goudman, J. Jansen, M. Billot, N. Vets, A. De Smedt, M. Roulaud, P. Rigoard, and M. Moens. Virtual reality applications in chronic pain management: Systematic review and meta-analysis. *JMIR Serious Games*, 10(2):e34402, 2022. doi:10.2196/34402.

[49] B. G. Green and V. P. Jennifer. Innocuous cooling can produce nociceptive sensations that are inhibited during dynamic mechanical contact. *Experimental brain research*, 148:290–299, 2003. doi:10.1007/s00221-002-1280-9.

[50] B. G. Green and L. S. Kate. Thermal and nociceptive sensations from menthol and their suppression by dynamic contact. *Behavioural brain research*, 176: 284–291, 2007. doi:10.1016/j.bbr.2006.10.013.

[51] R. L. Gregory. Perceptions as hypotheses. *Philosophical Transactions of the Royal Society of London. B, Biological Sciences*, 290(1038):181–197, 1980. doi:10.1098/rstb.1980.0090.

[52] S. Haegens, V. Nácher, A. Hernández, R. Luna, O. Jensen, and R. Romo. Beta oscillations in the monkey sensorimotor network reflect somatosensory decision making. *Proceedings of the National Academy of Sciences*, 108(26):10708–10713, 2011. doi:10.1073/pnas.1107297108.

[53] T. Hafting, M. Fyhn, S. Molden, M.-B. Moser, and E. I. Moser. Microstructure of a spatial map in the entorhinal cortex. *Nature*, 436(7052):801–806, 2005. doi:10.1038/nature03721.

[54] G. C. Hatfield. *The natural and the normative: Theories of spatial perception from Kant to Helmholtz*. MIT Press, Cambridge, MA, 1990.

[55] D. Hewitt, S. Tong, S. Schreiber, and B. Seymour. Tonic pain revalues associative memories of phasic pain. *bioRxiv*, 2025. doi:10.1101/2025.03.04.641525.

[56] J. Hilgard, B. Bartholow, C. Dickter, and H. Blanton. Characterizing switching and congruency effects in the implicit association test as reactive and proactive cognitive control. *Social cognitive and affective neuroscience*, 10(3):381–388, 2014. doi:10.1093/scan/nsu060.

[57] B. Hjorth. An on-line transformation of eeg scalp potentials into orthogonal source derivations. *Electroencephalography and clinical neurophysiology*, 39(5): 526–530, 1975. doi:10.1016/0013-4694(75)90056-5.

[58] R. Hoskin and D. Talmi. Adaptive coding of pain prediction error in the anterior insula. *European Journal of Pain*, 27(6):766–778, 2023. doi:10.1002/ejp.2093.

[59] Y. Huang, Z. A. Yaple, and R. Yu. Goal-oriented and habitual decisions: Neural signatures of model-based and model-free learning. *NeuroImage*, 215:116834, 2020. doi:10.1016/j.neuroimage.2020.116834.

[60] H. Huynh and L. S. Feldt. Estimation of the box correction for degrees of freedom from sample data in randomized block and split-plot designs. *Journal of Educational Statistics*, 1(1):69–82, 1976. doi:10.3102/10769986001001069.

[61] H. H. Jasper. Report of the committee on methods of clinical examination in electroencephalography: 1957. *Electroencephalography and Clinical Neurophysiology*, 10(2):370–375, 1958. doi:10.1016/0013-4694(58)90053-1.

[62] R. E. Kass and A. E. Raftery. Bayes factors. *Journal of the American Statistical Association*, 90(430):773–795, 1995. doi:10.1080/01621459.1995.10476572.

[63] C. Kemp and J. B. Tenenbaum. The discovery of structural form. *Proceedings of the National Academy of Sciences*, 105(31):10687–10692, 2008. doi:10.1073/pnas.0802631105.

[64] D. L. Kennedy, H. I. Kemp, D. Ridout, D. Yarnitsky, and A. S. C. Rice. Reliability of conditioned pain modulation: a systematic review. *Pain*, 157(11): 2410–2419, 2016. doi:10.1097/j.pain.0000000000000689.

[65] R. Klatzky. Allocentric and egocentric spatial representations: Definitions, distinctions, and interconnections. In C. Freksa, C. Habel, and K. F. Wender, editors, *Spatial Congition: An interdisciplinary approach to representing and processing spatial knowledge*, pages 1–17. Springer-Verlag, Berlin, Heidelberg, 1998. doi:10.1007/3-540-69342-4_1.

[66] B. Kopp, U. Mattler, R. Goertz, and F. Rist. N2, p3 and the lateralized readiness potential in a nogo task involving selective response priming. *Electroencephalography and Clinical Neurophysiology*, 99(1):19–27, 1996. doi:10.1016/0921-884X(96)95617-9.

[67] B. Kopp, A. Steinke, and A. Visalli. Cognitive flexibility and n2/p3 event-related brain potentials. *Scientific Reports*, 10(1):9859, 2020. doi:10.1038/s41598-020-66781-5.

[68] L. Koppel, G. Novembre, R. Kämpe, M. Savallampi, and I. Morrison. Prediction and action in cortical pain processing. *Cerebral Cortex*, 33(3):794–810, 2022. doi:10.1093/cercor/bhac102.

[69] L. K. Krugel, G. Biele, P. N. C. Mohr, S. C. Li, and H. R. Heekeren. Genetic variation in dopaminergic neuromodulation influences the ability to rapidly and flexibly adapt decisions. *Proceedings of the National Academy of Sciences*, 106 (42):17951–17956, 2009. doi:10.1073/pnas.0905191106.

[70] K. P. Körding, U. Beierholm, W. J. Ma, S. Quartz, J. B. Tenenbaum, and L. Shams. Causal inference in multisensory perception. *PLoS ONE*, 2(9):e943, 2007. doi:10.1371/journal.pone.0000943.

[71] M. E. Le Pelley. The role of associative history in models of associative learning: a selective review and a hybrid model. *Quarterly Journal of Experimental Psychology*, 57(3):193–243, 2004. doi:10.1080/02724990344000141.

[72] M. C. Lee, A. Mouraux, and G. D. Iannetti. Characterizing the cortical activity through which pain emerges from nociception. *Journal of Neuroscience*, 29(24): 7909–7916, 2009. doi:10.1523/JNEUROSCI.0014-09.2009.

[73] S. W. Lee, S. Shimojo, and P. O'Doherty, John. Neural computations underlying arbitration between model-based and model-free learning. *Neuron*, 81(3):687–699, 2014. doi:10.1016/j.neuron.2013.11.028.

[74] A. Li, Z. Montaño, V. J. Chen, and J. I. Gold. Virtual reality and pain management: Current trends and future directions. *Pain Management*, 1(2): 147–157, 2011. doi:10.2217/pmt.10.15.

[75] J. Li, D. Schiller, G. Schoenbaum, E. A. Phelps, and N. D. Daw. Differential roles of human striatum and amygdala in associative learning. *Nature Neuroscience*, 14:1250–1252, 2011. doi:10.1038/nn.2904.

[76] J. Lötsch and A. Ultsch. Machine learning in pain research. *Pain*, 159(4): 623–630, 2018. doi:10.1097/j.pain.0000000000001118.

[77] R. Mackel, A. Iriki, and E. E. Brink. Spinal input to thalamic vl neurons: evidence for direct spinothalamic effects. *J Neurophysiol*, 67(1):132–44, 1992. doi:10.1152/jn.1992.67.1.132.

[78] P. Mahajan, S. Tong, S. W. Lee, and B. Seymour. Balancing safety and efficiency in human decision making. *eLife*, 2024. doi:10.7554/elife.101371.1.

[79] S. Makeig. Auditory event-related dynamics of the eeg spectrum and effects of exposure to tones. *Electroencephalography and Clinical Neurophysiology*, 86(4): 283–293, 1993. doi:10.1016/0013-4694(93)90110-H.

[80] F. Mancini, P. Haggard, G. D. Iannetti, M. R. Longo, and M. I. Sereno. Fine-grained nociceptive maps in primary somatosensory cortex. *Journal of Neuroscience*, 32(48):17155–17162, 2012. doi:10.1523/jneurosci.3059-12.2012.

[81] F. Mancini, A. Bauleo, J. Cole, F. Lui, C. A. Porro, P. Haggard, and G. D. Iannetti. Whole-body mapping of spatial acuity for pain and touch. *Annals of Neurology*, 75(6):917–924, 2014. doi:10.1002/ana.24179.

[82] E. Maris and R. Oostenveld. Nonparametric statistical testing of eeg- and meg-data. *Journal of Neuroscience Methods*, 164:177–190, 2007. doi:10.1016/j.jneumeth.2007.03.024.

[83] J. W. Mauchly. Significance test for sphericity of a normal n-variate distribution. *The Annals of Mathematical Statistics*, 11(2):204–209, 1940.

[84] R. Melzack and K. L. Casey. Sensory, motivational, and central control determinants of pain: a new conceptual model. In D. R. Kenshalo, editor, *The Skin Senses*, pages 423–439. CC Thomas, Springfield, 1968.

[85] R. Melzack and P. D. Wall. Pain mechanisms: a new theory. *Science*, 150(3699): 971–979, 1965. doi:10.1126/science.150.3699.971.

[86] D. Mobbs, D. B. Headley, W. Ding, and P. Dayan. Space, time, and fear: Survival computations along defensive circuits. *Trends in Cognitive Sciences*, 24 (3):228–241, 2020. doi:https://doi.org/10.1016/j.tics.2019.12.016.

[87] S. Mohamed and D. Ott. Pain and machine learning. In *Workshop on Biological and Artificial Reinforcement Learning*, NeuRIPS 2020, 2020.

[88] O. Moriarty, B. E. McGuire, and D. P. Finn. The effect of pain on cognitive function: A review of clinical and preclinical research. *Progress in Neurobiology*, 93(3):385–404, 2011. doi:10.1016/j.pneurobio.2011.01.002.

[89] G. Murphy. *The big book of concepts*. MIT press, Cambridge, MA, 2002.

[90] B. Namer, F. Seifert, H. O. Handwerker, and C. Maihöfner. Trpa1 and trpm8 activation in humans: effects of cinnamaldehyde and menthol. *Neuroreport*, 16: 955–959, 2005. doi:10.1097/00001756-200506210-00015.

[91] V. Neugebauer. Amygdala pain mechanisms. In H. G. Schaible, editor, *Pain Control*, pages 261–284. Springer Berlin Heidelberg, 2015. doi:10.1007/978-3-662-46450-2_13.

[92] J. O'Keefe and L. Nadel. *The Hippocampus as a Cognitive Map*. Oxford: Clarendon Press, 1978.

[93] R. Oostenveld, P. Fries, E. Maris, and J.-M. Schoffelen. Fieldtrip: Open source software for advanced analysis of meg, eeg, and invasive electrophysiological data. *Computational intelligence and neuroscience*, 2011:156869, 2011. doi:10.1155/2011/156869.

[94] R. A. Poldrack. Region of interest analysis for fmri. *Social Cognitive and Affective Neuroscience*, 2(1):67–70, 2007. doi:10.1093/scan/nsm006.

[95] D. D. Price, J. D. Greenspan, and R. Dubner. Neurons involved in the exteroceptive function of pain. *Pain*, 106(3):215–219, 2003. doi:10.1016/j.pain.2003.10.016.

[96] L. Quandt, P. Marshall, and C. Bouquet. Somatosensory experiences with action modulate alpha and beta power during subsequent action observation. *Brain research*, 1534:55–65, 2013. doi:10.1016/j.brainres.2013.08.043.

[97] S. N. Raja, D. B. Carr, M. Cohen, N. B. Finnerup, H. Flor, S. Gibson, F. J. Keefe, J. S. Mogil, M. Ringkamp, K. A. Sluka, X.-J. Song, B. Stevens, M. D. Sullivan, P. R. Tutelman, T. Ushida, and K. Vader. The revised international association for the study of pain definition of pain: concepts, challenges, and compromises. *PAIN*, 161(9):1976–1982, 2020. doi:10.1097/j.pain.0000000000001939.

[98] G. E. Schwarz. Estimating the dimension of a model. *The Annals of Statistics*, 6(2):461–464, 1978. doi:10.1214/aos/1176344136.

[99] M. Sebold, S. Nebe, M. Garbusow, M. Guggenmos, D. J. Schad, A. Beck, S. Kuitunen-Paul, C. Sommer, R. Frank, P. Neu, et al. When habits are dangerous: alcohol expectancies and habitual decision making predict relapse in alcohol dependence. *Biological psychiatry*, 82(11):847–856, 2017. doi:10.1016/j.biopsych.2017.04.019.

[100] M. E. P. Seligman. Learned helplessness. *Annual Review of Medicine*, 23: 407–412, 1972. doi:10.1146/annurev.me.23.020172.002203.

[101] B. Seymour. Pain: A precision signal for reinforcement learning and control. *Neuron*, 101(6):1029–1041, 2019. doi:10.1016/j.neuron.2019.01.055.

[102] B. Seymour and F. Mancini. Hierarchical models of pain: Inference, information-seeking, and adaptive control. *NeuroImage*, 222:117212, 2020. doi:10.1016/j.neuroimage.2020.117212.

[103] C. S. Sherrington. *The integrative action of the nervous system.* Archibald Constable and Co., Ltd., London, 1906.

[104] K. L. Stachenfeld, M. M. Botvinick, and S. J. Gershman. The hippocampus as a predictive map. *Nature Neuroscience*, 20(11):1643–1653, 2017. doi:10.1038/nn.4650.

[105] R. Staud. Abnormal endogenous pain modulation is a shared characteristic of many chronic pain conditions. *Expert Review of Neurotherapeutics*, 12(5): 577–585, 2012. doi:10.1586/ern.12.41.

[106] K. E. Stephan, W. D. Penny, J. Daunizeau, R. J. Moran, and K. J. Friston. Bayesian model selection for group studies. *NeuroImage*, 46(4):1004–1017, 2009. doi:10.1016/j.neuroimage.2009.03.025.

[107] R. T. Stevens, S. M. London, and A. Vania Apkarian. Spinothalamocortical projections to the secondary somatosensory cortex (sii) in squirrel monkey. *Brain Research*, 631(2):241–246, 1993. doi:10.1016/0006-8993(93)91541-Y.

[108] I. A. Strigo and A. D. Craig. Interoception, homeostatic emotions and sympathovagal balance. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 371(1708):20160010, 2016. doi:10.1098/rstb.2016.0010.

[109] R. S. Sutton. Adapting bias by gradient descent: an incremental version of delta-bar-delta. In *Proceedings of the Tenth National Conference on Artificial Intelligence*, AAAI'92, page 171–176, 1992. doi:10.5555/1867135.1867162.

[110] R. S. Sutton and A. G. Barto. Toward a modern theory of adaptive networks: expectation and prediction. *Psychological Review*, 88(2):135–170, 1981. doi:10.1037/0033-295X.88.2.135.

[111] R. S. Sutton and A. G. Barto. *Reinforcement learning: An introduction.* MIT press, Cambridge, MA, 1998.

[112] A. Tabor and C. Burr. Bayesian learning models of pain: A call to action. *Current Opinion in Behavioral Sciences*, 26:54–61, 2019. doi:10.1016/j.cobeha.2018.10.006.

[113] A. Tabor, M. A. Thacker, G. L. Moseley, and K. P. Körding. Pain: A statistical account. *PLOS Computational Biology*, 13(1):e1005142, 2017. doi:10.1371/journal.pcbi.1005142.

[114] E. Tan, S. V. Troller-Renfree, S. Morales, G. A. Buzzell, M. McSweeney, M. Antúnez, and N. A. Fox. Theta activity and cognitive functioning: Integrating evidence from resting-state and task-related developmental electroencephalography (eeg) research. *Developmental Cognitive Neuroscience*, 67:101404, 2024. doi:doi.org/10.1016/j.dcn.2024.101404.

[115] J. J. Teh, D. J. Pascoe, S. Hafeji, R. Parchure, A. Koczoski, M. P. Rimmer, K. S. Khan, and B. H. Al Wattar. Efficacy of virtual reality for pain relief in medical procedures: a systematic review and meta-analysis. *BMC Med*, 22(1):64, 2024. doi:10.1186/s12916-024-03266-6.

[116] E. C. Tolman. Cognitive maps in rats and men. *Psychol Rev*, 55(4):189–208, 1948. doi:10.1037/h0061626.

[117] S. Tong, T. Denison, D. Hewitt, S. W. Lee, and B. Seymour. Phasic and tonic pain serve distinct functions during adaptive behaviour. *bioRxiv*, 2025. doi:10.1101/2025.02.10.637253.

[118] C. Tsallis. Possible generalization of boltzmann-gibbs statistics. *Journal of Statistical Physics*, 52:479–487, 1988. doi:10.1007/BF01016429.

[119] C. Tsallis and D. A. Stariolo. Generalized simulated annealing. *Physica A: Statistical Mechanics and its Applications*, 233(1):395–406, 1996. doi:10.1016/S0378-4371(96)00271-3.

[120] S. D. Vann, J. P. Aggleton, and E. A. Maguire. What does the retrosplenial cortex do? *Nat Rev Neurosci*, 10(11):792–802, 2009. doi:10.1038/nrn2733.

[121] D. Vernon, G. Metta, and G. Sandini. A survey of artificial cognitive systems: Implications for the autonomous development of mental capabilities in computational agents. *IEEE Transactions on Evolutionary Computation*, 11(2):151–180, 2007. doi:10.1109/TEVC.2006.890274.

[122] P. D. Wall. On the relation of injury to pain the john j. bonica lecture. *PAIN*, 6 (3):253–264, 1979. doi:10.1016/0304-3959(79)90047-2.

[123] O. Wang, S. W. Lee, J. O'Doherty, B. Seymour, and W. Yoshida. Model-based and model-free pain avoidance learning. *Brain and Neuroscience Advances*, 2: 239821281877296, 2018. doi:10.1177/2398212818772964.

[124] J. C. R. Whittington, T. H. Muller, S. Mark, G. Chen, C. Barry, N. Burgess, and T. E. J. Behrens. The tolman-eichenbaum machine: Unifying space and relational memory through generalization in the hippocampal formation. *Cell*, 183(5):1249–1263.e23, 2020. doi:10.1016/j.cell.2020.10.024.

[125] F. Wilcoxon. Individual comparisons by ranking methods. *Biometrics Bulletin*, 1(6):80–83, 1945. doi:doi.org/10.2307/3001968.

[126] C. Witham, M. Wang, and S. Baker. Cells in somatosensory areas show synchrony with beta oscillations in monkey motor cortex. *The European journal of neuroscience*, 26(9):2677–2686, 2007. doi:10.1111/j.1460-9568.2007.05890.x.

[127] T. Wolbers and J. M. Wiener. Challenges for identifying the neural mechanisms that support spatial navigation: the impact of spatial scale. *Frontiers in Human Neuroscience*, 8:571, 2014. doi:10.3389/fnhum.2014.00571.

[128] C. W. Woo, A. Krishnan, and T. D. Wager. Cluster-extent based thresholding in fmri analyses: Pitfalls and recommendations. *NeuroImage*, 91:412–419, 2014. doi:10.1016/j.neuroimage.2013.12.058.

[129] A. Wunderlich, R. Klug, G. Stuber, B. Landwehrmeyer, F. Weber, and W. Freund. Caudate nucleus and insular activation during a pain suppression paradigm comparing thermal and electrical stimulation. *The open neuroimaging journal*, 5:1–8, 2011. doi:10.2174/1874440001105010001.

[130] K. Wunderlich, P. Dayan, and R. J. Dolan. Mapping value based planning and extensively trained choice in the human brain. *Nature Neuroscience*, 15(5): 786–791, 2012. doi:10.1038/nn.3068.

[131] Y. Xiang, D. Sun, W. Fan, and X. G. Gong. Generalized simulated annealing algorithm and its application to the thomson model. *Physics Letters A*, 233: 216–220, 1997. doi:10.1016/S0375-9601(97)00474-X.

[132] W. Yoshida, B. Seymour, M. Koltzenburg, and R. J. Dolan. Uncertainty increases pain: Evidence for a novel mechanism of pain modulation involving the periaqueductal gray. *Journal of Neuroscience*, 33(13):5638–5646, 2013. doi:10.1523/jneurosci.4984-12.2013.

[133] B. Zavala, A. Jang, M. Trotta, C. I. Lungu, P. Brown, and K. A. Zaghloul. Cognitive control involves theta power within trials and beta power across trials in the prefrontal-subthalamic network. *Brain*, 141(12):3361–3376, 2018. doi:10.1093/brain/awy266.

[134] S. Zhang, H. Mano, M. Lee, W. Yoshida, M. Kawato, T. W. Robbins, and B. Seymour. The control of tonic pain by active relief learning. *eLife*, 7, 2018. doi:10.7554/elife.31949.