



Solving the exploration-exploitation dilemma in a dynamic pain-related context: The benefits of exploratory behaviors

Maryna Alves^{a,b,*}, Angelos-Miltiadis Kryptos^{a,c}, Geert Crombez^b, Johan W.S. Vlaeyen^{a,d}

^a Research Group Health Psychology, Psychological Institute KU Leuven, Tiensestraat 102, Leuven, 3000, Belgium

^b Department of Experimental-Clinical and Health Psychology, Ghent University, Henri Dunantstraat, Gent, 59000, Belgium

^c Department of Clinical Psychology, Utrecht University, Heidelberglaan 1, Utrecht, 3584 CS, the Netherlands

^d Experimental Health Psychology, Maastricht University, P.O.Box 616, Maastricht, 6200 MD, the Netherlands

ARTICLE INFO

Keywords:

Exploration-exploitation dilemma

Chronic pain

Computational modeling

ABSTRACT

Daily decision-making involves navigating the explore-exploit dilemma (EED), balancing novel information gathering against maintaining current strategies. Individuals with chronic pain often exhibit exploitation responses, choosing pain avoidance over exploring potentially beneficial activities. This study investigated decision-making during the EED in individuals with chronic pain and pain-free individuals, examining whether self-initiated exploration predicts improved outcomes when environmental contingencies change. Participants completed a multi-armed bandit task where contingency changes occurred exclusively for unchosen options, requiring active exploration for discovery. We assessed behavioral performance, computational modeling parameters, and individual difference variables including fear of pain, intolerance of uncertainty, and behavioral activation/inhibition. Initial high exploration rates predicted superior task performance across all groups, despite few participants discovering the contingency change. Contrary to our original hypotheses, no consistent differences emerged between the pain-free and pain groups. However, differences between the pain groups suggested that greater pain severity may be associated with less stable and less exploitative decision-making. These findings highlight the importance of an adaptive balance between exploration and exploitation for discovering environmental changes and achieving optimal long-term outcomes, supporting its role in avoidance extinction relevant to chronic pain treatment.

1. Introduction

Daily life consists of a chain of decisions. In a dynamic and often uncertain environment, decision making is characterized by a conflict between several, often opposing, demands. One such conflict is the explore/exploit dilemma (EED) that represents a choice between two behavioral options each having benefits and costs (Cohen et al., 2007; Sutton & Barto, 2018). Exploration is gathering novel information with the opportunity to improve the current situation, but with also the risk to worsen it. Exploitation is maintaining the current status but with the risk of missing out on opportunities for improvement. To function adaptively, one must find the best trade-off between exploring an uncertain environment and exploiting existing knowledge (Cohen et al., 2007; Sutton & Barto, 2018). Within an operant framework, exploration and exploitation can be understood as patterns of behavior shaped by reinforcement histories and by the environmental contingencies.

Accordingly, the EED is used here as a functional description of how individuals adapt their behavior to changing environmental demands. The modeling and understanding of solving the EED have been a topic of interest in computer science (Charmanas et al., 2023; Wang et al., 2018), neuroscience (Cohen et al., 2007), organizational and health economics (Azoulay-Schwartz et al., 2004; Uotila et al., 2008). Nevertheless, research on EED in the context of pain remains limited (Kryptos et al., 2024).

Pain is a unpleasant experience and often results in a conflict between keeping pain at a manageable level and engaging in valued activities (Roux et al., 2022). Individuals in pain often show an 'exploitation' response by repeatedly choosing the action leading to the least pain (pain avoidance), instead of exploring other ways to continue activities, and may even help keep pain at bay. To experimentally study behavioral decision making in pain, previous studies have used 4-armed bandit tasks in which participants freely selected one option amongst

* Corresponding author. Tiensestraat 102, 3000, Leuven, Belgium.

E-mail address: maryna-alves@outlook.com (M. Alves).

<https://doi.org/10.1016/j.brat.2026.105142>

Received 22 February 2026; Received in revised form 15 July 2026; Accepted 10 August 2026

Available online 11 August 2026

0005-7967/© 2026 Published by Elsevier Ltd.

four at each trial, with each option having a specific probability of a reward and/or pain-related stimulus (e.g., Alves, 2024; Krypotos et al., 2022; Seymour et al., 2012). Across studies, the results indicate that participants readily learned the changes in choice-outcome contingencies during the experiments and behaved accordingly. Specifically, after initial exploration, participants most often exploited the safest option, and weighted rewards more heavily than receiving pain. Additionally, after receiving a painful outcome, participants were more inclined to explore (Alves, 2024; Krypotos et al., 2022; Seymour et al., 2012). However, in these studies the choice-outcome contingencies changed for all options simultaneously. As a result, the decision to explore might have been caused by the contingency changes (e.g., a rewarding option stops being rewarding). This may therefore hamper the systematic study of self-initiated exploration in pain-related decision-making, as participants may be prompted to adapt in similar ways by externally imposed contingency changes.

In the present study, we investigated decision making during the EED with individuals reporting chronic pain and pain-free individuals. We created a context where only self-initiated exploration could lead to the discovery of the choice-outcome contingency change that transformed the less favorable former options (i.e., low reward and high pain) into the most favorable ones (i.e., low pain and high reward). We hypothesized that higher levels of exploration are associated with better task performance in all groups. In addition, we hypothesized that compared to the pain-free individuals, individuals with chronic pain would exhibit more exploitative behavior along with higher learning and sensitivity rates for punishments, indicating a heightened response to pain-related stimuli. This hypothesis is in line with a meta-analysis showing an attentional bias towards pain-related information, including words and pictures, in individuals with chronic pain (Crombez et al., 2013; Todd et al., 2018). We also explored whether the individual differences in fear of pain, intolerance of uncertainty, behavioral activation, and behavioral inhibition affect the EED outcomes.

2. Methods

2.1. Participants

Participants were recruited via the Prolific online platform (<https://www.prolific.co/>). These individuals were divided into groups based on the report of pain at the time of the study: *the pain-free group* (i.e., individuals experiencing no chronic pain) and *the pain group* (i.e., individuals with self-reported chronic pain). The first step in the recruitment was to pre-screen specific individuals' characteristics on Prolific, tailoring the study visibility to individuals who met the specified criteria. The inclusion criteria were: (1) age range of 18 to 65 years; (2) a proficient command of the English language; (3) presence of pain lasting for at least 3 months for the pain group. Exclusion criteria were: (1) participation in our previous studies; (2) presence of pain lasting for at least 3 months for the pain-free group. All participants completed the Von Korff Pain Grading Scale to estimate the intensity and interference of their pain (Smith et al., 1997). The ethical committee of KU Leuven approved this study (Reg. #G-2023-6568-R4). Participants were compensated with a monetary reward of £7 for their participation. The present experiment was pre-registered before data collection (<https://osf.io/bfm53>).

Participants' data were included in the analyses only if they met the following criteria: (1) completing the entire experiment, (2) not missing more than 10% of the trials (i.e., late responses), (3) a minimum score of 5 out of 7 in the attention checks (as recommended by Prolific; see Supplementary Note 1 for the attention-check instructions).

Two hundred participants completed this experiment. Three participants, 2 from the pain group and 1 from the pain-free group, were excluded from the data analyses because they missed more than 10% of the trials. The pain-free group was initially composed of 99 participants, birth sex females = 48 (48.5%), Mean age = 30.06 (SD = 9.47). The pain

group comprised 98 volunteers, birth sex females = 48 (49%), Mean age = 28.5 (SD = 7.29). However, several individuals in the pain-free group had a Grade I or higher score on the Von Korff Pain Grading Scale, indicating the presence of low pain and low disability. To avoid conflating individuals with and without pain within the same comparison group, participants were separated into three groups based on the Von Korff Pain Grading Scale¹: Grade 0 (no pain), Grade I (low-intensity and low-disability pain), and Grades II–IV (Grade II = high-intensity and low-disability pain; Grade III = high-disability and moderately limiting pain; Grade IV = high-disability and severely limiting pain). Grades II–IV were combined to avoid unstable estimates due to small subgroup sizes and to distinguish low-severity pain from more severe pain conditions. The final distribution across groups was as follows: The pain-free group or Grade 0: N = 31, birth sex females = 17 (54.8%), Mean age = 28.81 (SD = 9.23). The low-pain group or Grade I: N = 94, birth sex females = 40 (42.6%), Mean age = 30.57 (SD = 9.31). Finally, the moderate-to-severe pain group or Grade II–IV: N = 72, birth sex females = 39 (54.2%), Mean age = 27.81 (SD = 6.64). See [Supplementary Material 1](#) for the frequency of scores on Von Korff Pain Grading and [Supplementary Material 2](#) for the score of pain experienced at the time of the experiment.

2.2. Questionnaires

To explore the potential impact of individual differences on the EED in a pain-related context, data on the following questionnaires were collected: the Intolerance of Uncertainty Scale (IUS; internal consistency: $\alpha = .94$; Buhr & Dugas, 2002), the Fear of Pain Questionnaire III (FPQ-III; internal consistency: $\alpha = .92$; McNeil & Rainwater, 1998), the Behavioral Inhibition System/Behavioral Activation System Scales (BIS/BAS; reliability BIS internal consistency: $\alpha = .74$; BAS internal consistency: $\alpha = .66-.76$; Carver & White, 1994), and the Von Korff Pain Grading Scale (internal consistency: $\alpha = .90$; Smith et al., 1997). These questionnaires were selected because the individual differences they assess have previously been associated with altered pain perception and avoidance, which could influence performance during the EED in the presence of pain-related outcomes (Donthula et al., 2020; Markfelder & Pauli, 2020; Serrano-Ibáñez et al., 2019).

2.3. Stimuli

The present study was conducted online, which made the use of actual physical pain stimulation impossible. Consequently, twenty pain-related pictures from the Experimental Pain Picture System (EPPS; Alves et al., 2023) were used as pain-related stimuli (See Fig. S1 for examples of pictures). Previous studies have shown that images depicting pain-related content can elicit discomfort and distress, as well as physiological arousal (Fusaro et al., 2016; Preis & Kroener-Herwig, 2012).

A gold coin served as the representation of the reward. Participants were informed about the opportunity to earn an additional £50 if they achieved the highest number of rewards among all participants in the task.

2.4. Ratings

Before starting the experiment and after the termination of the task, we asked participants to rate the painfulness and the arousal of the pain-related pictures using 10-point scales (Alves, 2024), with the painfulness scale ranging from 1 (= not pain at all) to 10 (= worst pain possible) and the arousal scale ranging from 1 (= totally calm) to 10 (= highest level of arousal). At the end of the experiment, participants rated the

¹ The three-group classification and the corresponding re-analyses were conducted in response to a suggestion by an anonymous reviewer, whom we thank for the recommendation.

pleasantness and the appreciation of the pain-related picture and the reward, respectively ranging between 1 (= very pleasant) to 10 (= very unpleasant) and 1 (= very much appreciated) to 10 (= not at all appreciated). All instructions for the evaluation of the stimuli are provided in Supplementary Note 2.

2.5. Procedure

This experiment was programmed with Psychopy (version 2022.2.3; Peirce et al., 2019) and run online via the Pavlovia platform (<https://pavlovia.org/>). The questionnaires and demographical information were filled out on Qualtrics (<https://www.qualtrics.com/>). Participants started the experiment by completing the informed consent, the Von Korff Pain Grading Scale (Smith et al., 1997) and the following demographical information (self-report): age, gender, birth sex, education level, income class, ethnicity, country of residence, presence of chronic pain, pain interference, presence of neurological disorder, and level of fatigue at the moment of completing the measure (see Supplementary Material 3 for the results). Next, we presented 20 pictures of different types of pain from the EPPS (Alves et al., 2023) and asked participants to watch each picture carefully and select the picture that would provoke the worst pain if they were in that situation. The selected pain picture was used as a punishment throughout the experiment.

For the experimental paradigm, we used an online version of the 4-armed bandit task (Alves, 2024; Kryptos et al., 2022). Four squares were displayed on the corners of the screen, and participants were asked to select one of them with a mouse click. After selecting the square, the selected square changed color, and the painful picture and/or the reward were displayed for 500 ms depending on the probability associated with the selected square (see Fig. 1 for an example of a trial). The inter-trial interval varied between 1500 and 3000 ms. The experiment started with 10 practice trials, followed by two consecutive blocks of 80 trials, without a break or clear distinction between blocks. In the first block, the probabilities were set as follows: (1) 20% chance of pain and 50% chance of reward; (2) 30% chance of pain and 50% chance of reward; (3) 80% chance of pain and 20% chance of reward; (4) 70% chance of pain and 30% chance of reward. In the second block, the choice-outcome contingencies of the *less favorable options* changed (i.e., options 3 and 4), having a 20% chance of pain and 80% chance of reward for option 3 and a 30% chance of pain and 70% chance of reward for option 4 (Fig. 1). Finally, after the termination of the experimental

task, participants completed the remaining questionnaires (i.e., IUS, FPQ-III, and BIS/BAS) and reported their level of fatigue once more. The mean duration of the full task, including the questionnaires, was 35 min. All the instructions given during the experimental task are reported in Supplementary Note 3.

2.6. Statistical analyses

We conducted all the statistical analyses using the R (version R4.5.1; R Core Team, 2020). For the behavioral analyses, we used both frequentist and Bayesian approaches to provide a more comprehensive and robust assessment of the evidence. Where appropriate, we used non-parametric tests and controlled for multiple comparisons using the Holm correction. Significant overall effects were followed by pairwise comparisons with Holm-adjusted p-values.

We used computational models to characterize the behavioral data collected during the task (e.g., choices and outcomes) into mathematical parameters that can be interpreted as model-based estimates of cognitive latent processes (e.g., learning rates – how fast participants learn – or sensitivity – reactivity to outcomes), without assuming that these patterns necessarily reflect deliberate or explicit cognitive computations. To do so, models were fitted using the hBayesDM package in R (Ahn et al., 2017). We fitted these five models for each group and each block separately: (1) Bandit 4arm 4par, (2) Bandit 4arm 2par Lapse, (3) Bandit 4arm Lapse, (4) Bandit4arm Lapse Decay, and (5) Bandit4arm SingleA Lapse (see Table 1 for model specification). Each model was estimated using 3 chains for each model with 1000 warm-ups and 5000 iterations per chain. Next, we selected the best model using the leave-one-out information criterion (LOOIC; Vehtari et al., 2016); the model having the lowest LOOIC was selected as the best-fitting model. Convergence of the selected model was assessed by examining the Markov Chain Monte Carlo (MCMC), with well-mixed chains fluctuating around a stable mean (“hairy caterpillars”) taken as evidence of good convergence. Differences in relevant model parameters across blocks and groups were evaluated by calculating the posterior distributions of the corresponding contrasts and comparing their 95% highest density interval (HDI) to the region of practical equivalence (ROPE). A difference was considered meaningful when the HDI fell entirely outside the ROPE, practically equivalent to zero when it fell entirely within it, and inconclusive when the two overlapped only partially (Kruschke, 2011).

Outputs from the selected computational model were used to test our hypotheses and answer our research questions. We modified the selected

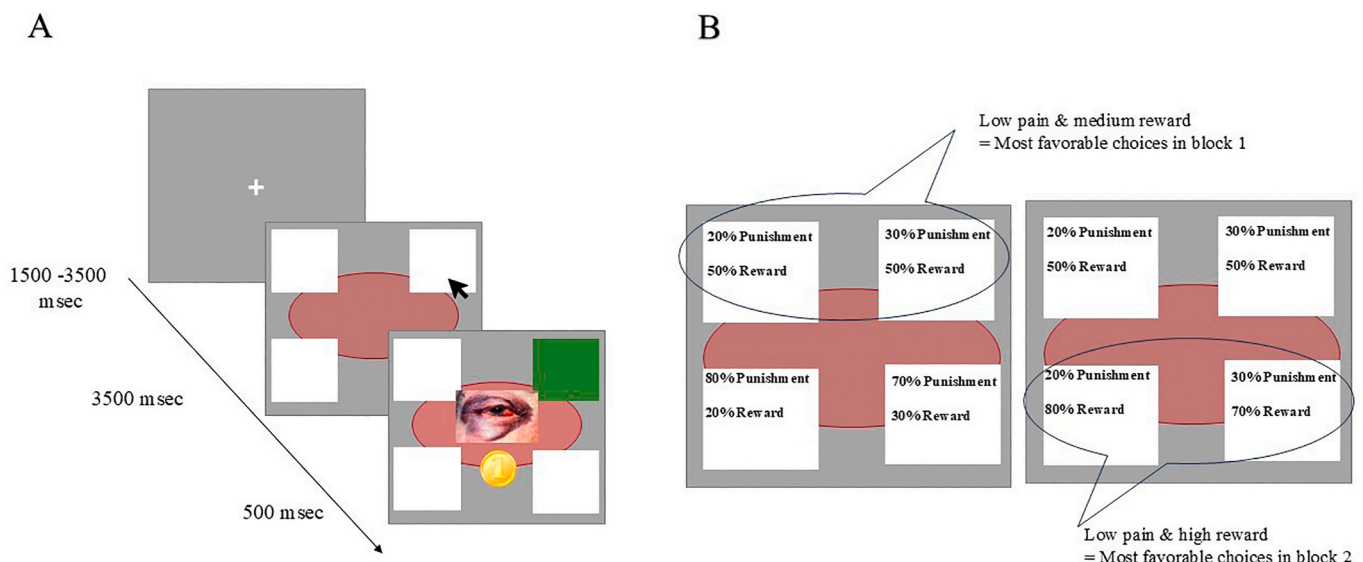


Fig. 1. (A) Example of a trial. (B) Probabilities for blocks 1 and 2.

Table 1

Model specification.

Model	Parameters
Bandit 4arm 4par	Reward learning rate ^a ; Punishment learning rate ^b ; Reward sensitivity ^c ; Punishment sensitivity ^d
Bandit 4arm 2par Lapse	Reward learning rate ^a ; Punishment learning rate ^b ; Lapse rate ^e
Bandit 4arm Lapse	Reward learning rate ^a ; Punishment learning rate ^b ; Reward sensitivity ^c ; Punishment sensitivity ^d ; Lapse rate ^e
Bandit 4arm Lapse Decay	Reward learning rate ^a ; Punishment learning rate ^b ; Reward sensitivity ^c ; Punishment sensitivity ^d ; Lapse rate ^e ; Decay rate ^f
Bandit 4arm SingleA Lapse	Learning rate ^g ; Reward sensitivity ^c ; Punishment sensitivity ^d ; Lapse rate ^e

Note. ^a The reward learning rate ($0 < A_{rew} < 1$) and ^b the punishment learning rate ($0 < A_{pun} < 1$), indicate how individuals update their estimated value for a given option after receiving a reward or a punishment, respectively. ^c The reward sensitivity parameter shows the reactivity of participants to the reward ($0 < R < 30$), in other words, how much they appreciated and anticipated the reward. ^d The punishment sensitivity points out participants' reactivity to punishment and how much they disliked and anticipated it ($0 < P < 30$). ^e The lapse rate demonstrates the randomness in participants' choices ($0 < \xi < 1$). ^f The decay rate reflects the forgetfulness for unchosen options ($0 < d < 1$). ^g Finally, the learning rate represents the learning rates for reward and punishment together ($0 < A < 1$).

model to extract the action value (i.e., the value for a given option based on the interaction with the task) on every trial for each participant, enabling the estimation of both exploration and exploitation at each trial. To estimate the action value of each option at each trial, the computational model considered both reward and pain-related outcomes received on previous trials, as well as model parameters for each participant (e.g., reward sensitivity). In this experiment, exploitation did not necessarily correspond to selecting the option associated with the highest reward, but rather to selecting the option with the highest subjective action value for that individual at trial t (Cohen et al., 2007; Mehlhorn et al., 2015). For example, an option associated with lower probability of reward and lower probability of pain-related outcome may be considered the exploitative choice if, based on the individual's learning history and model parameters, it has the highest estimated action value. Exploration was defined as selecting any alternative option.

The action value Q for bandit i at trial t was derived from the computational model as follows (Ahn et al., 2017; Aylward et al., 2019):

$$Q_{t(i)} = Q_{t(i)}^{\text{reward}} + Q_{t(i)}^{\text{punishment}} \quad (1)$$

The action value for reward and punishment are estimated separately.

$$\begin{aligned} Q_{t(i)}^{\text{reward}} &= Q_{t(i)}^{\text{reward}} + A_{\text{reward}} * \text{Prediction Error}_{t(i)}^{\text{reward}} Q_{t(i)}^{\text{punishment}} \\ &= Q_{t(i)}^{\text{punishment}} + A_{\text{punishment}} * \text{Prediction Error}_{t(i)}^{\text{punishment}} \end{aligned} \quad (2)$$

A_{reward} and $A_{\text{punishment}}$ are the learning rates for reward and punishment, respectively.

$$\begin{aligned} \text{If } i = \text{chosen} \\ \text{Prediction Error}_{t(i)}^{\text{reward}} &= R * \text{Reward}_t - Q_{t-1(i)}^{\text{reward}} \\ \text{Prediction Error}_{t(i)}^{\text{punishment}} &= P * \text{Punishment}_t - Q_{t-1(i)}^{\text{punishment}} \\ \text{If } i = \text{unchosen} \\ \text{Prediction Error}_{t(i)}^{\text{reward}} &= -Q_{t-1(i)}^{\text{reward}} \\ \text{Prediction Error}_{t(i)}^{\text{punishment}} &= -Q_{t-1(i)}^{\text{punishment}} \end{aligned} \quad (3)$$

R represents the reward sensitivity and P the punishment sensitivity. $\text{Reward}_t = 1$ if a reward was delivered at trial t , and 0 otherwise. The Punishment_t is estimated similarly.

The choice probability is estimated by using the SoftMax learning

rule and the lapse ξ (i.e., noise) to account for choice randomness.

$$p(\text{choice} = i) = \text{softmax}(Q_{t(i)}) * (1 - \xi) + \frac{\xi}{4} \quad (4)$$

$$\text{softmax}(Q_{t(i)}) = \frac{\exp(Q_{t(i)}^{\text{reward}} + Q_{t(i)}^{\text{punishment}})}{\sum_j \exp(Q_{t(j)}^{\text{reward}} + Q_{t(j)}^{\text{punishment}})}$$

Here, j represents all the bandits.

For the unchosen options, the action value is updated using the decay rate d .

$$Q_{t(1)} = (1 - d) * Q_{t(1)} \text{ if } i = \text{unchosen} \quad (5)$$

To assess whether participants learned the contingency in each block, we calculated a difference score between the number of times they selected the two most favorable options and the two least favorable options. A difference score equal to or below zero indicated that participants either failed to learn the contingency or responded randomly.

Testing the first hypothesis (i.e., higher levels of exploration during the first block are associated with better performance during block 2), we conducted a multiple linear regression with the exploration frequency in block 1 and group as predictors and the number of rewards in the second block as the outcome variable.

To test our second hypothesis (i.e., individuals with chronic pain would exhibit more exploitative behavior and higher learning and sensitivity rates for punishments) and test the differences between groups, we conducted a series of between-group analyses. We examined the behavioral performance using pairwise independent-samples t-tests comparing the three groups. More specifically, we compared the following variables: the number of rewards and number of punishments received, the proportion of switches after receiving a reward, the proportion of switches after a punishment, the proportion of switches after a punishment and a reward, and the proportion of switches after no outcome.

We conducted the switch-no-switch analyses using a mixed ANOVA with the switch-no-switch (i.e., switch was selecting a different option than on the previous trial and no-switch was selecting the same option as on the previous trial) and the block as within-subjects factors and group as between-subjects factor.

We performed the exploration-exploitation analyses using a mixed ANOVA with exploration-exploitation factor and block as within-subjects factors, and group as between-subjects factor.

We also conducted a mixed-effects logistic regression testing the influence of reward, punishment, choice perseveration (i.e., the tendency to select the same option independently of the outcome), and group on the decision to explore or exploit at a trial-by-trial level. A participant-specific random intercept was included to account for the repeated-measures structure of the data (following Seymour et al. (2012); see Supplementary Note 4 for more details). The punishment and reward bias represents the probability of selecting an option again in the next trial if participants encountered punishment and reward when selecting it in the recent past (i.e., five preceding trials).

To investigate associations between individual difference (i.e., intolerance of uncertainty, fear of pain, and behavioral activation/inhibition systems) and model parameters, and whether these associations differed across groups, we conducted exploratory regression analyses. Each regression included group, a questionnaire score, and their interaction as predictors. To avoid multicollinearity from simultaneously including total and subscale scores from the same questionnaire, we fitted two alternative models for each outcome-questionnaire combination. One included the total score and its interaction with group, the other included the subscale scores and their interactions with group. These alternative models were compared using the Akaike Information Criterion (AIC), and the model with the lower AIC was selected.

We also conducted a multiple linear regression analysis using a stepwise forward and backward variable selection strategy to identify which factors impacted the optimal-choice performance (i.e., the

number of times participants selected the most favorable option during the second block). We included the following predictors: cognitive processes derived from the model parameters (i.e., reward learning rate, punishment learning rate, reward sensitivity, punishment sensitivity, decay rate, and lapse rate), grade on the Von Korff Pain Grading Scale, exploration frequency in block 1, group, and the pain score at the moment.

A note of caution is warranted. These analyses were largely exploratory, and the relatively small sample size in the pain-free group means the results should be interpreted with care, despite the use of conservative analytical choices (e.g., non-parametric and correction for multiple comparisons).

3. Results

3.1. Computational analyses

The model comparison analyses reveal that the best model fit of the data in all the groups and blocks was the Bandit4arm Lapse Decay (Table S6 in the Supplementary Material). This model has been the winning model in our previous studies as well (Alves, 2024; Kryptos et al., 2022). The Bandit4arm Lapse Decay model has six parameters. First, the reward learning rate ($0 < A_{rew} < 1$) indicates how individuals update their behavior after receiving a reward; high values imply that more importance is given to recent outcomes, and small values indicate that more importance is given to past outcomes. Second, the punishment learning rate ($0 < A_{pun} < 1$) reflects how individuals adjust their behavior following a punishment. Third, the reward sensitivity parameter shows the reactivity and appreciation of participants to the reward ($0 < R < 30$). Fourth, the punishment sensitivity illustrates participants' reactivity and aversion to punishment ($0 < P < 30$). Fifth, the decay rate reports the tendency to forget unchosen options ($0 < d < 1$). Finally, the lapse rate demonstrates the randomness in participants' choices ($0 < \xi < 1$).

Visual inspection of the MCMC indicated satisfactory convergence (Fig. S2–S4 in the Supplementary Material). However, for some groups, the estimates of the lapse rate and/or punishment learning rate were concentrated near the boundaries of their permitted ranges, which may suggest restricted variability of these parameters.

In general, the results for all groups show low reward and high punishment learning rates (Table 2, Figs. 2–4). The sensitivity for reward is low but higher than the sensitivity for the punishment (Table 2, Figs. 2–4). Finally, the results indicate low lapse rates (Table 2, Figs. 2–4). The results indicated either no practically meaningful differences or inconclusive evidence for differences in model parameters between blocks across all groups (Fig. S5–S7 in Supplementary Material 4). Contrary to our expectations, the results did not provide practically meaningful evidence that participants in the pain groups showed greater sensitivity to punishment than participants in the pain-free group (Fig. S8–S13 in Supplementary Material 4). However, in block 1, the moderate-to-severe pain group showed a higher decay rate than the low-pain group, which may suggest faster forgetting of the values associated with unchosen options. All remaining between-group comparisons were

inconclusive or not practically meaningful (Fig. S5–S13).

3.2. Descriptive analyses

All groups consistently rated the pictures as painful and arousing at the beginning and end of the experiment. Participants from the pain-free and pain groups expressed that the presented pictures were unpleasant and that they did not appreciate encountering them during the task. The reward appreciation and pleasantness were moderate in all groups (Table 3).

Four participants failed to learn the choice-outcome contingency during block 1 but did learn it in block 2. Eleven participants did not learn the choice-outcome contingency at all (neither in block 1 nor in block 2). Few participants learned the choice-outcome contingency change during block 2 (58 out of 197; 29.44%). Among those who learned the contingency in block 2, 11 belonged to the pain-free group (35.48% of the pain-free group) and 47 belonged to the pain groups (28.31% of the pain groups). Of these, 22 were in the moderate-to-severe pain group (30.56% of that group) and 25 were in the low-pain group (26.60% of that group).

3.3. Predictive performance analyses

The results show that the overall regression analysis is statistically significant ($F(3,193) = 3.70, p = .01$) and explains a relatively small proportion of variance in rewards obtained ($R^2_{adj} = .04$), suggesting that other unmeasured factors likely play a substantial role. A higher exploration frequency during block 1 significantly predicted a higher number of rewards during block 2 ($b = 0.07, p_{adj} = .004, 95\% \text{ CI } [0.03, 0.11], BF_{10} = 21.4$). There was no overall evidence for the effect of the group in predicting the number of rewards, $F(2, 193) = 0.90, p = .408, BF_{10} = 0.14$. Neither the moderate-to-severe group ($b = -0.14, p_{adj} = .872, 95\% \text{ CI } [-1.83, 1.55]$) nor the low-pain group ($b = 0.68, p_{adj} = .829, 95\% \text{ CI } [-0.96, 2.31]$) differed significantly from the pain-free group. Collectively, these findings suggest that more exploratory behavior predicts better subsequent performance in terms of rewards obtained, whereas pain severity did not influence task performance.

3.4. Behavioral performance analyses

The number of rewards received differed across the three groups (Kruskal-Wallis $\chi^2(2) = 11.37, p = .003$). Dunn post-hoc comparisons showed that this effect was driven by a difference between the low-pain group ($M = 83.71, SD = 6.30$) and the moderate-to-severe pain group ($M = 80.85, SD = 5.53$), $z = -3.19, p_{adj} = .004, BF_{10} = 11.93$, with the low-pain group obtaining more rewards. The pain-free group ($M = 81.94, SD = 5.50$) did not differ significantly from either the low-pain group, $z = 2.09, p_{adj} = .073, BF_{10} = 0.517$, or the moderate-to-severe pain group, $z = -0.31, p_{adj} = .758, BF_{10} = 0.325$. The number of punishments also differed across groups ($F(2, 194) = 3.61, p = .029$), again driven by the contrast between the low-pain group ($M = 49.45, SD = 9.11$) and the moderate-to-severe pain group ($M = 53.04, SD = 8.06$), $z = -3.60, p_{adj} = .033, BF_{10} = 4.13$, the moderate-to-severe

Table 2
Parameter estimates (means and standard deviations).

	Low-pain group		Moderate-to-severe pain group		Pain-free group	
	Block 1	Block 2	Block 1	Block 2	Block 1	Block 2
$0 < A_{rew} < 1$	0.17 ± 0.04	0.16 ± 0.04	0.24 ± 0.06	0.25 ± 0.05	0.18 ± 0.05	0.29 ± 0.08
$0 < A_{pun} < 1$	0.80 ± 0.07	0.90 ± 0.06	0.74 ± 0.10	0.96 ± 0.03	0.45 ± 0.10	0.69 ± 0.18
$0 < R < 30$	4.74 ± 0.69	6.53 ± 0.87	3.87 ± 0.72	4.69 ± 0.51	4.12 ± 0.71	4.20 ± 1.07
$0 < P < 30$	0.98 ± 0.19	0.74 ± 0.19	1.31 ± 0.31	0.77 ± 0.21	1.73 ± 0.57	0.53 ± 0.29
$0 < d < 1$	0.21 ± 0.04	0.43 ± 0.08	0.50 ± 0.06	0.67 ± 0.07	0.22 ± 0.06	0.54 ± 0.11
$0 < \xi < 1$	0.003 ± 0.003	0.005 ± 0.003	0.10 ± 0.03	0.03 ± 0.02	0.02 ± 0.02	0.01 ± 0.01

Note. Arew: reward learning rate; Apun: punishment learning rate; R: reward sensitivity; P: punishment sensitivity; d: decay rate; ξ : lapse rate.

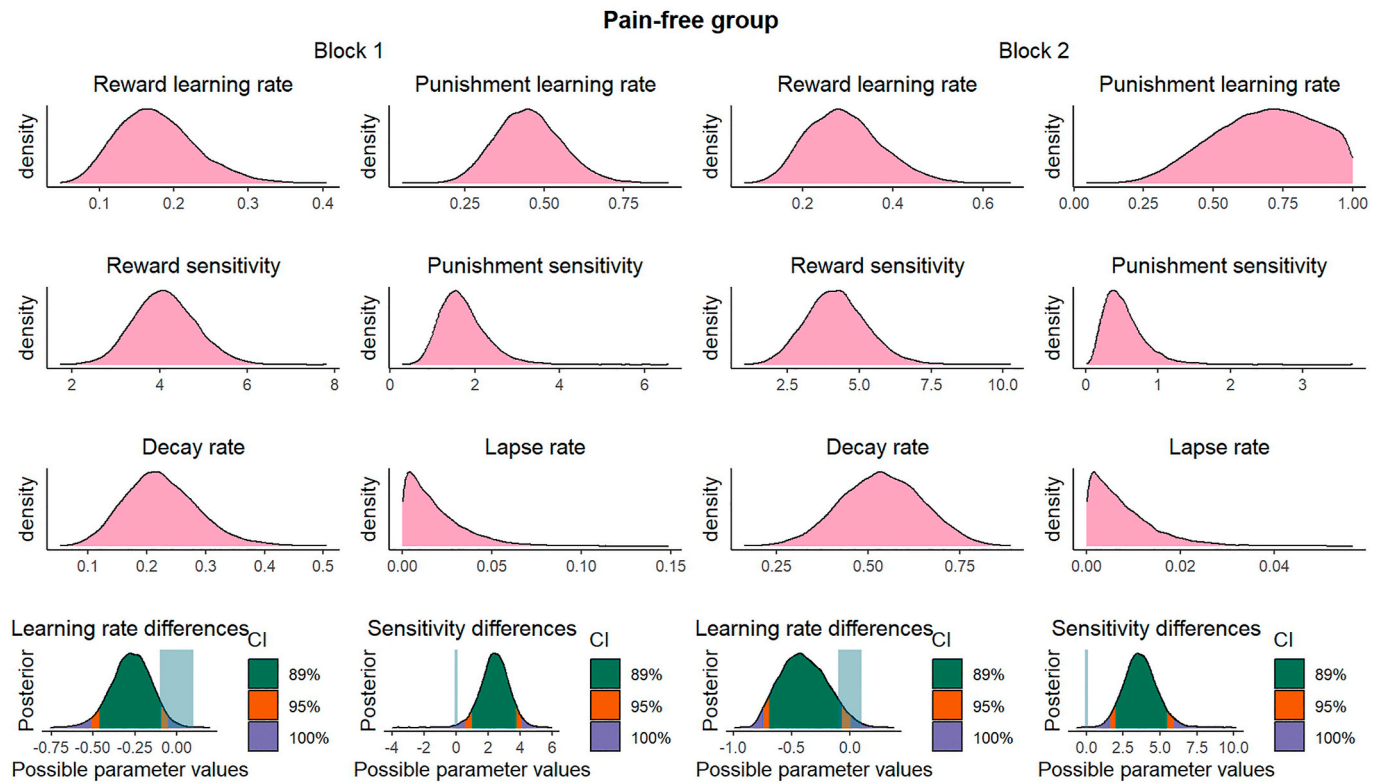


Fig. 2. Posterior distribution plots for the model parameters for the pain-free group.

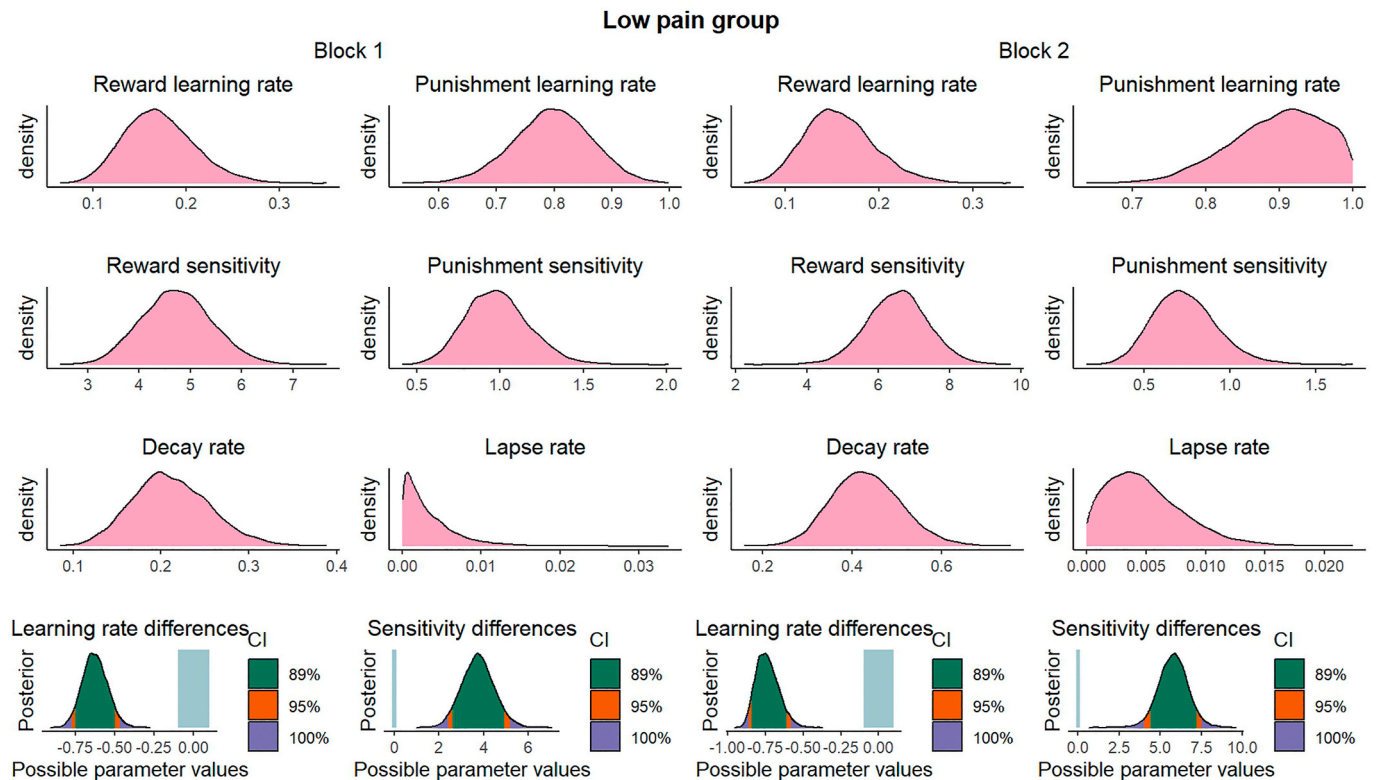


Fig. 3. Posterior distribution plots for the model parameters for the low-pain group.

pain group receiving more punishments. The pain-free group ($M = 49.65$, $SD = 10.20$) did not differ from the low-pain group, $z = 0.198$, $p_{adj.} = .915$, $BF_{10} = 0.218$, or the moderate-to-severe pain

group, $z = -3.397$, $p_{adj.} = .156$, $BF_{10} = .928$. Together, these results indicate that behavioral performance varied with pain severity: the low-pain group performed best (i.e., most rewards and fewest punishments)

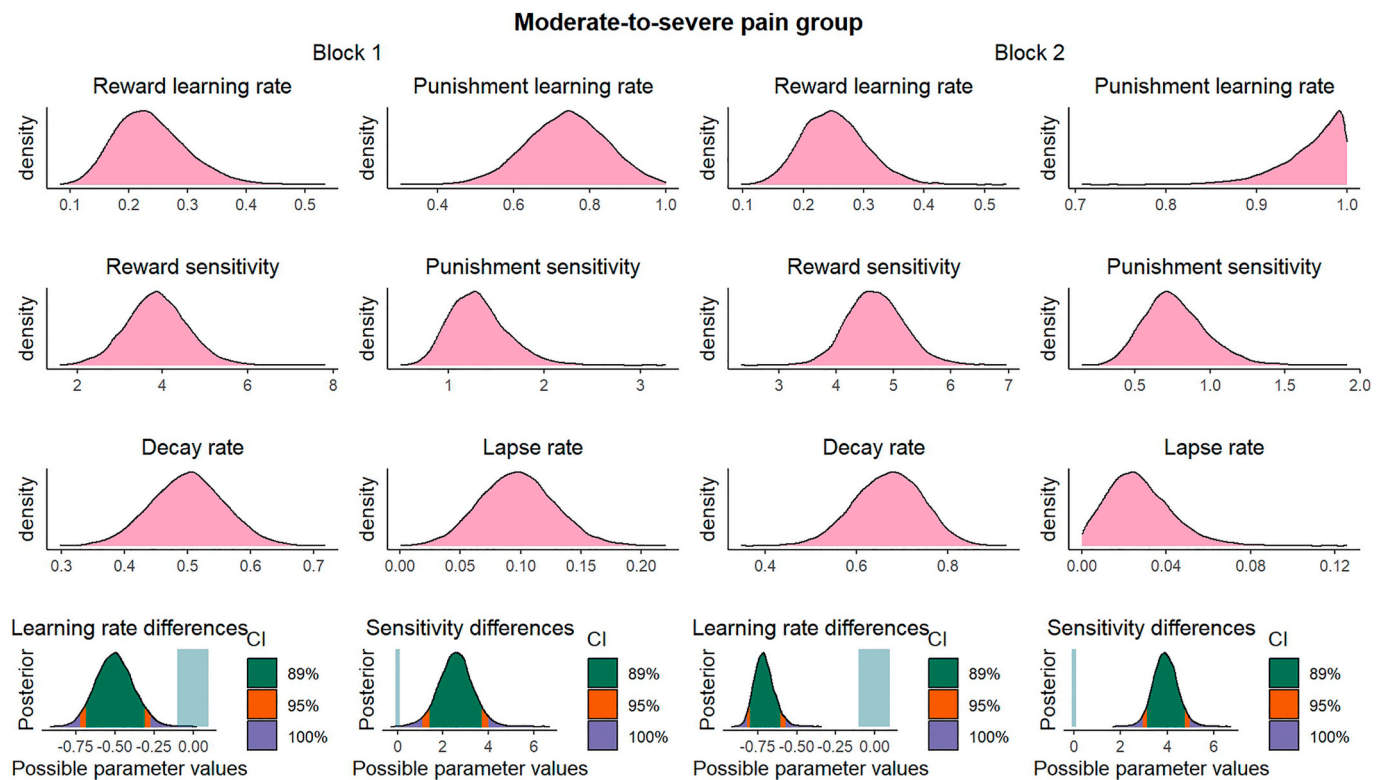


Fig. 4. Posterior distribution plots for the model parameters for the moderate-to-severe pain group.

Table 3

Evaluation of the stimuli.

	Pain-free group		Low-pain group		Moderate-to-severe pain group	
	Before	After	Before	After	Before	After
Painfulness of the picture	9.04 ± 1.39	9.09 ± 1.12	8.37 ± 2.04	8.15 ± 2.32	8.75 ± 1.47	8.46 ± 2.14
Arousal of the picture	8.65 ± 1.93	8.61 ± 1.66	7.62 ± 2.59	7.55 ± 2.69	7.91 ± 2.36	7.50 ± 2.96
Pleasantness of the picture		2.49 ± 3.12		2.83 ± 3.17		2.84 ± 3.29
Appreciation of the picture		3.06 ± 2.72		2.40 ± 1.80		2.63 ± 2.25
Pleasantness of the reward		4.69 ± 3.97		5.73 ± 3.35		5.55 ± 3.61
Appreciation of the reward		4.19 ± 4.60		3.85 ± 4.40		4.06 ± 4.50

compared to the moderate-to-severe pain group, with the pain-free group falling in between.

The proportion of switches also varied with pain severity across all four outcome types. Switches were more frequent in one group than another after no outcome ($F(2, 194) = 4.35, p = .014$), after punishment ($F(2, 194) = 3.90, p = .022$), after simultaneous punishment and reward ($F(2, 194) = 3.43, p = .034$), and after reward ($F(2, 194) = 5.63, p = .004$). In each case the effect was driven by the difference between low and moderate-to-severe pain groups, with the moderate-to-severe pain group switching more; the only contrast reaching significance after correction was switches after reward ($p_{adj.} = .042, BF_{10} = 29.166$). All the remaining contrasts were non-significant (all $p_{adj.} \geq .147, BF_{10}$ between 0.218 and 6.735; Table S7).

In general, participants from all groups switched more after receiving the punishment compared to the rewards (switches after punishment: $M = 0.78, SD = 0.21$; switches after a reward: $M = 0.47 (SD = 0.34), t(333) = -11.0, p_{adj.} < .001, BF_{10} > 1000$; the reward and punishment together (switches after reward and punishment: $M = 0.59, SD = 0.31, t(351) = 6.92, p_{adj.} < .001, BF_{10} > 1000$; or no outcome at all (switches after no outcome: $M = 0.64, SD = 0.26, t(378) = 5.62, p_{adj.} < .001, BF_{10} > 1000$).

Taken together, these analyses show that pain severity was associated with task performance: the low-pain group obtained more rewards,

fewer punishments, and switched less frequently, compared to the moderate-to-severe pain group. The pain-free group consistently fell between the two without differing significantly from either. Generally, participants tended to select the same option more often after receiving a reward than after receiving a punishment alone or with the reward, suggesting the urge to avoid pain-related pictures.

3.5. Switch and no-switch analyses

The mixed ANOVA revealed a significant main effect of the switch-no-switch factor, $F(1, 194) = 140.26, p < .001$, and a significant main effect of block, $F(1, 194) = 38.98, p < .001$, whereas the main effect of group was not significant, $F(2, 194) = 2.11, p = .124$. The block $\times$ switch interaction was significant, $F(1, 194) = 4.66, p = .032$: post-hoc comparisons showed that participants switched more than they stayed in both blocks, but this gap was larger in block 1 ($b = -25.60, t(776) = -16.36, p_{adj.} < .001, BF_{10} = 8.74e+31$) than in block 2 ($b = -21.44, t(776) = -13.71, p_{adj.} < .001, BF_{10} = 7.35e+13$).

The group $\times$ switch-no-switch interaction was also significant, $F(2, 194) = 5.31, p = .006$. Post-hoc comparisons indicated that this effect was driven by the low-pain versus moderate-to-severe pain group contrast: the two groups differed both in their number of no-switches ($b = 6.56, t(776) = 4.26, p_{adj.} < .001, BF_{10} = 30.325$) and switches

($b = -6.13$, $t(776) = -3.97$, $p_{adj.} < .001$, $BF_{10} = 16.630$), whereas pain-free group did not differ significantly from either group (all $p_{adj.} > .07$; Table S8).

The remaining interactions did not reach significance: the group $\times$ block interaction, $F(2, 194) = 1.30$, $p = .276$, the group $\times$ block $\times$ switch-no-switch interaction, $F(2, 194) = 0.57$, $p = .564$.

To summarize, our findings indicate that participants initially switched more between options than they selected the same option in the previous trial. During the second block, the number of switches decreased. Switching behavior also varied with pain severity: the low-pain and moderate-to-severe pain groups differed in both their switch and no-switch counts, whereas the pain-free group did not differ significantly from either.

3.6. Exploration-exploitation analyses

We analyzed the exploration rate using an Aligned Rank Transform (ART) factorial ANOVA, with group and block as factors and a by-participant random intercept. There was a significant main effect of block, $F(1, 194) = 126.92$, $p < .001$. The main effect of group was not significant, $F(2, 194) = 1.36$, $p = .258$. The group $\times$ block interaction did not reach significance, $F(2, 194) = 2.92$, $p = .056$, indicating only a trend toward group differences in how exploration changed across blocks.

The aligned-rank-transform interaction comparisons showed that exploration declined significantly from block 1 to block 2 within every group (low-pain: $t(194) = 6.76$, $p < .001$; moderate-to-severe pain: $t(194) = 9.99$, $p < .001$; pain-free: $t(194) = 5.26$, $p < .001$). Between-group differences were present in the first block only: the low-pain group explored significantly less than the moderate-to-severe group in block 1 ($t(284) = -2.97$, $p_{adj.} = .023$, $BF_{10} = 16.144$), whereas no group differences remained in block 2 (all $p_{adj.} = 1.00$; Table S9). The pain-free group did not differ significantly from either pain group within a given block (all $p_{adj.} > 0.24$; Table S9).

These results suggest no difference between the amount of exploration and exploitation during the first block, indicating no difference between the number of times participants selected the option with the highest action value and the number of times they explored other options (Fig. 5). During the second block, however, participants mainly selected the best option based on their interaction with the task (i.e., exploitation) but they maintained a certain level of exploration (Fig. 3). Furthermore, the results indicated only a difference between the low-pain group and the moderate-to-severe pain group during the first block.

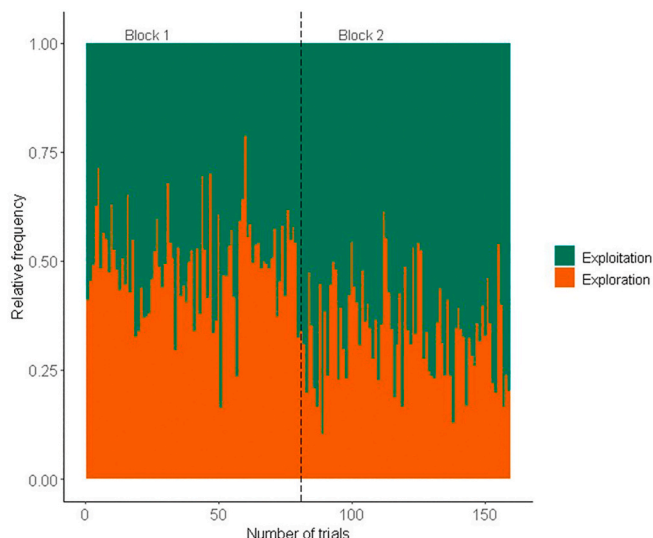


Fig. 5. Frequency of exploration and exploitation per trial.

3.7. Logistic regression

Our results indicate that when keeping all other predictors constant, there was a 2.9% reduction in the odds of exploration for one increased unit of the reward weight, implying that individuals who exhibit a stronger bias towards rewards are more likely to exploit forthcoming trials if the previous trials have been rewarding ($b = -0.030$, $p_{adj.} < .001$, 95% CI [-0.036, -0.024]). After applying the Holm correction to the p-value, results show no impact of the punishment weight on the odds of exploitation ($b = -0.002$, $p_{adj.} = .951$). The odds of exploration were decreased by 62.5 % for each unit increase of the choice perseveration (OR = 0.375). Participants who tended to choose the same option regardless of the outcomes were more likely to exploit in the subsequent trial ($b = -0.982$, $p_{adj.} < .001$, 95% CI [-1.027, -0.937]). Group as a whole impacted the balance between exploration and exploitation at a trial-by-trial level ($\chi^2(2) = 14.570$, $p = .0006$). Holm-corrected pairwise comparisons showed that this effect was driven by a difference between the two pain groups: the odds of exploration were decreased by 28.97% in the low-pain group compared to the moderate-to-severe pain group ($b = -0.342$, $p_{adj.} = .0005$). The pain-free group did not differ from the low-pain group ($b = 0.086$, $p_{adj.} = .475$), nor from the moderate-to-severe pain group in terms of odds of exploration ($b = -0.256$, $p_{adj.} = .077$).

In sum, at a trial-by-trial level, the weight for reward increased the odds for exploitation, while the weight for punishment did not impact the odds of exploitation after the Holm correction. Nevertheless, participants who tended to select the same option independently of the outcomes are more likely to exploit. With respect to group differences, results show a decrease in the odds of exploitation for the moderate-to-severe pain group compared to the low-pain group, but no significant difference between either pain group and the pain-free group.

3.8. Associations between model parameters and individual differences by group

Full results for the association between model parameters (i.e., reward learning rate, punishment learning rate, reward sensitivity, punishment sensitivity, decay rate and lapse rate), individual differences (i.e., intolerance of uncertainty, fear of pain, and behavioral activation/inhibition systems) and group (i.e., pain-free, low-pain and moderate-to-severe pain groups) are reported in the Supplementary Material 5 (Table S10–S13). After applying the Holm correction for multiple comparisons, the analyses revealed the following significant effects.

For punishment learning rate, pairwise comparisons showed a graded pattern across all models: lowest in the pain-free group ($M = 0.568$, $SD = 0.01$), higher in the low-pain group ($M = 0.842$, $SD = 0.004$), and highest in the moderate-to-severe pain group ($M = 0.846$, $SD = 0.004$; all $p_{adj.} < .001$ see Table S11–S13). In the behavioral activation/inhibition systems model, there was also a significant main effect of behavioral activation drive score on punishment learning rate, $F(1, 182) = 9.10$, $p_{adj.} = .020$, and a significant group $\times$ behavioral activation drive score interaction, $F(2, 182) = 7.41$, $p_{adj.} = .006$. Higher drive was associated with lower punishment learning rate in the pain-free group only (vs. low-pain: $b = -0.002$, $p_{adj.} = .001$, 95% CI [-0.004, -0.001]; vs. moderate-to-severe pain: $b = -0.002$, $p_{adj.} = .002$, 95% CI [-0.004, -0.001]), with no association in the other groups. As behavioral activation drive reflects persistent goal pursuit, motivation, and ambition, these findings suggest that individuals without pain who report higher goal-directed drive may show reduced learning from punishment.

For reward sensitivity, the low-pain group ($M = 5.635$, $SD = 1.195$) scored higher than both the pain-free group ($M = 4.164$, $SD = 0.908$, all $p_{adj.} \leq .022$) and the moderate-to-severe pain group ($M = 4.280$, $SD = 0.746$, $p_{adj.} < .001$) across all three models; the pain-free and moderate-to-severe pain groups did not differ.

For decay rate, the moderate-to-severe pain group ($M = 0.585$,

SD = 0.108) scored higher than both the pain-free group ($M = 0.381$, $SD = 0.180$, $p_{adj.} \leq .008$) and the low-pain group ($M = 0.322$, $SD = 0.127$, $p_{adj.} < .001$) across all three models; the pain-free and low-pain groups did not differ.

For lapse rate, all three groups differed significantly from one another across models (all $p_{adj.} < .001$), lowest in the low-pain group ($MD = 0.004$, $SD = 0.003$), intermediate in the pain-free group ($MD = 0.013$, $SD = 0.013$), and highest in the moderate-to-severe pain group ($MD = 0.062$, $SD = 0.043$). In the behavioral activation/inhibition systems model, a significant group $\times$ behavioral activation total score interaction also emerged, $F(2, 188) = 4.73$, $p_{adj.} = .040$. Pairwise comparisons showed that higher behavioral activation predicted lower lapse rate only in the moderate-to-severe pain group suggesting greater approach motivation may relate to fewer inconsistent choices specifically in this group.

Overall, these exploratory analyses showed consistent group differences in punishment learning rate, reward sensitivity, decay rate, and lapse rate, suggesting that pain severity was associated with differences in how participants learned from negative outcomes, responded to rewards, updated unchosen options, and made inconsistent choices. In addition, group moderated specific associations between behavioral activation and model parameters. More specifically, higher behavioral activation drive was associated with lower punishment learning rate only in participants without pain, whereas higher behavioral activation total score was associated with lower lapse rate only in the moderate-to-severe pain group. All other associations between individual differences and model parameters were non-significant after correction for multiple comparisons.

3.9. Multilinear regression analysis

The stepwise regression shows that the most optimal model to predict behavioral performance during block 2 includes the following factors: exploration frequency during block 1 ($b = 0.43$, $p_{adj.} = .019$, 95% CI [0.16, 0.71]), the behavioral activation score ($b = -0.48$, $p_{adj.} = .121$, 95% CI [-0.87, -0.08]), group (moderate-to-severe pain vs. pain-free group: $b = 24.88$, $p_{adj.} = .287$, 95% CI [-10.77, 60.54]; low-pain vs. pain-free group: $b = -9.67$, $p_{adj.} = .187$, 95% CI [-18.46, -0.89]), the reward learning rate ($b = 10.57$, $p_{adj.} = .220$, 95% CI [-0.23, 21.37]), reward sensitivity ($b = 1.04$, $p_{adj.} = .287$, 95% CI [-0.19, 2.27]), the decay rate ($b = 19.96$, $p_{adj.} = .010$, 95% CI [8.07, 31.86]), and the lapse rate ($b = -527.71$, $p_{adj.} = .287$, 95% CI [-1171.84, 116.43]).

The overall model was significant and explained a modest proportion of the variance in performance ($R^2_{adj.} = .16$, $F(8, 188) = 5.82$, $p < .001$), meaning that each parameter composing the model presented here contributed to obtaining the most optimal fit. However, not all the predictors significantly impacted behavioral performance once the Holm correction was applied. Two predictors remained significant: exploration frequency during the first block and the decay rate. Indeed, participants exploring more during the first block tended to perform better during the second block and hence selected more options that were most beneficial. Likewise, a higher decay rate was associated with better performance, suggesting that participants who more readily forgot the value of unchosen options achieved better outcomes.

4. Discussion

In this experiment, we aimed to investigate the impact of exploration during the EED in a pain-free group and groups of individuals reporting chronic pain. To do so, we designed the experiment such that choice-outcome contingency change could only be discovered after self-initiated exploration. The results can be readily summarized as follows: First and as predicted, initial high exploration was a predictor for better behavioral performance later on, despite the low number of participants who discovered the contingency change. This suggests that early exploratory behavior may have facilitated later adaptation to the

changed environment.

Second, we found some group differences in behavioral performance. The low-pain group obtained more rewards, received fewer punishments, and switched less frequently between options compared to the moderate-to-severe pain group. The low-pain group also explored less than the moderate-to-severe pain group during the first block. In addition, group differences emerged at the trial-by-trial level, with the moderate-to-severe pain group showing lower odds of exploitation than the low-pain group. In terms of computational modeling, the ROPE analyses revealed a difference in decay rate between the low-pain and moderate-to-severe pain groups for block 1 only. However, the exploratory regression analyses showed that groups differed in punishment learning rate, reward sensitivity, decay rate, and lapse rate. Punishment learning rate and lapse rate differed across all three groups, reward sensitivity was highest in the low-pain group, and decay rate was highest in the moderate-to-severe pain group. Finally, individual differences showed selective associations with model parameters. Behavioral activation moderated specific associations with punishment learning rate and lapse rate. Higher behavioral activation drive was associated with lower punishment learning rate only among participants without pain, whereas higher behavioral activation total score was associated with lower lapse rate only among participants in the moderate-to-severe pain group.

As expected, high initial exploration predicted better task performance in terms of selecting options that led to the highest chances of receiving a reward. In addition to the exploration frequency, the decay rate also predicted better task performance afterward. In other words, participants forgetting more or at a faster pace about unreinforced options would tend to approach these options later on. These findings suggest that exploration and a certain level of forgetfulness of unchosen options are required to discover environmental changes (e.g., reduction of aversive outcomes or increase of attractive outcomes) and hence for the extinction of avoidance behaviors that are commonly observed in individuals with chronic pain. Avoidance can be adaptive in the presence of genuine threats but becomes dysfunctional when it prevents engagement in non-harmful activities (Vlaeyen & Crombez, 2020). However, exploration alone may not be sufficient for optimal performance. Individuals with moderate-to-severe pain explored more than those with low-pain during the first block, yet they obtained fewer rewards and encountered more pain-related pictures. This suggests that adaptive behavior depends on a flexible balance between exploration and exploitation: individuals need to explore enough to detect changes in the environment, but also to exploit beneficial options once they have been identified. Therefore, encouraging exploration may lead to avoidance extinction by enabling individuals to learn that previously avoided situations are not always harmful and may even become rewarding, but this process likely requires flexible adjustment between exploration and exploitation.

Exploratory behavior represents a strategic investment in information acquisition despite uncertain immediate returns, as it guarantees knowledge acquisition while offering no assurance regarding outcome quality. Exploration becomes increasingly advantageous under extended time horizons that permit leveraging newly acquired information, heightened environmental uncertainty regarding available options, and the introduction of novel alternatives requiring assessment (Addicott et al., 2017; Cohen et al., 2007; Kembro et al., 2019). Conversely, exploitation strategies prove superior when decisions carry high immediate stakes or when information can be obtained independently of choice behavior (Cohen et al., 2007), thereby diminishing the instrumental value of exploratory sampling for future decision-making, such as in acute pain episodes. Therefore, the optimal balance between exploration and exploitation is context dependent.

Our findings are partially consistent with previous computational modeling studies of decision-making in chronic pain without pain-related cues. In line with Zhang et al. (2022), who reported no difference in the overall performance between a chronic pain group and a

control group, the pain-free group in the present study did not consistently differ from either pain-severity group. However, differences emerged between the low-pain and moderate-to-severe pain groups. Participants with moderate-to-severe pain obtained fewer rewards, received more pain-related pictures, switched more frequently, and showed lower odds of exploitation than participants with low pain. Although the moderate-to-severe pain group showed more exploratory responses, this pattern did not translate into better performance. Rather, it may reflect a less efficient balance between exploration and exploitation, together with greater randomness in choice behavior, as indicated by the elevated lapse rate observed in the moderate-to-severe pain group.

This interpretation is consistent with findings from Zhang et al. (2022) and Hess et al. (2014), who found that individuals with chronic pain showed less strategic decision-making, often selecting less optimal options. These studies also reported higher reward and punishment learning rates, together with greater reward sensitivity, in individuals with chronic pain compared with healthy controls. The present results showed that groups differed in punishment learning rate, reward sensitivity, decay rate, and lapse rate. In particular, punishment learning rate and lapse rate increased with pain severity, while decay rate was highest in the moderate-to-severe pain group. This suggests that individuals with higher pain severity may update more strongly from negative outcomes, discount unchosen options more rapidly, and show greater inconsistency in their choices. This pattern may suggest that moderate-to-severe pain is associated with unstable or inefficient exploration: participants may continue sampling alternatives, but without effectively using the information gained to maximize rewards and avoid pain-related outcomes. Nevertheless, approach motivation appeared to play a protective role in the moderate-to-severe pain group, as higher behavioral activation was associated with more consistent choice behavior.

Our results reveal a dissociation between motivational salience and learning efficiency, wherein participants demonstrated heightened reward sensitivity coupled with accelerated learning rates for punishment. This asymmetry suggests that while positive reinforcement drives behavioral motivation, aversive experiences produce more rapid belief updating, a pattern consistent with evolutionary theories emphasizing the adaptive value of rapid threat learning over gradual reward optimization (Lin et al., 2020; Rasmussen & Newland, 2008).

Despite these strengths, this study also has limitations. The first limitation is that the experiment was conducted online, reducing control over the conditions in which participants completed the task. The second limitation is the sampling bias that might have occurred by recruiting participants via Prolific. Even though Prolific provides a diverse worldwide sample, the participants are often young and technologically literate, which might not be representative of the population of individuals with chronic disabling pain, which tends to be older and less educated (Dahlhamer et al., 2018). Furthermore, most participants had low-pain and low disability scores; future studies should recruit a more diverse sample in terms of pain severity. Lastly, we used pain-related pictures to study pain-related responses instead of noxious stimuli.

The present findings may have several practical implications for future research and, more broadly, for the assessment of pain-related avoidance. First, the reinforcement-learning approach used here allows exploration, exploitation, and avoidance-related behavior to be examined as dynamic patterns of choices shaped by previous outcomes, rather than as the mere presence or absence of a predefined avoidance response. This approach may help identify when individuals persistently select options associated with lower perceived threat or lower expected pain, even when alternative options may be safe or potentially more adaptive. In this sense, the task may contribute to characterizing individual differences in rigid versus flexible avoidance behavior.

Second, the findings suggest that encouraging exploration may be useful in chronic pain interventions, but that exploration alone may not

be sufficient. Adaptive functioning may require helping individuals flexibly balance exploration with exploitation, so that they can test whether avoided situations are safe while also learning to rely on beneficial actions once these have been identified. This distinction is important because excessive or poorly regulated exploration may not necessarily lead to better outcomes.

Third, although the present study was not designed to test intervention strategies, future research could examine whether approaches such as graded exposure, reductions in perceived threat or uncertainty, or affective manipulations promote more flexible choice behavior in pain-related contexts. Positive affect may be a particularly promising avenue for increasing adaptive exploration (Fredrickson, 2001; Rowe et al., 2007; Saragosa-Harris et al., 2022). Fredrickson's broaden-and-build theory (Fredrickson & Branigan, 2005) holds that positive affect broadens the behavioral and cognitive repertoire, loosening the pull toward habitual, narrow responding — a process that maps directly onto increased response variability and reduced exploitation of familiar options. Positive mood states also appear to heighten reward sensitivity and lower the threshold for approaching uncertain alternatives (Craske et al., 2023), both of which are conducive to exploration. Conversely, chronic pain is strongly associated with negative affect, which tends to promote behavioral rigidity and a preference for the familiar over the uncertain (Zaman et al., 2021). Even modest affective shifts could therefore indirectly widen the behavioral repertoire and facilitate exploration.

CRedit authorship contribution statement

Maryna Alves: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Investigation, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Angelos-Miltiadis Kryptos:** Writing – review & editing, Supervision, Project administration, Conceptualization. **Geert Crombez:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Johan W.S. Vlaeyen:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization.

Declaration of competing interests

The authors declare that there are no known competing financial or personal interests that could have influenced the work reported in this paper.

Acknowledgements

The present work was supported by an FWO grant (Reg. #G071118N) awarded to J.W.S. Vlaeyen and G. Crombez. J.W.S. Vlaeyen is also supported by the Asthenes long-term structural funding–Methusalem grant from the Flemish Government, Belgium (METH/15/011). The data and code supporting the findings of this study are available from the corresponding author upon request.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.brat.2026.105142>.

Data availability

Data will be made available on request.

References

- Addicott, M. A., Pearson, J. M., Sweitzer, M. M., Barack, D. L., & Platt, M. L. (2017). A primer on foraging and the explore/exploit trade-off for psychiatry research.

- Neuropsychopharmacology, 42(10), 1931–1939. <https://doi.org/10.1038/npp.2017.108>. <https://www.nature.com/articles/npp2017108>. (Accessed 29 May 2017)
- Ahn, W. Y., Haines, N., & Zhang, L. (2017). Revealing neurocomputational mechanisms of reinforcement learning and decision-making with the hBayesDM package. *Computational psychiatry (Cambridge, Mass.)*, 1, 24–57. https://doi.org/10.1162/CPSY_a.00002
- Alves, M. (2024). *Should I stay or should I go? The exploration-exploitation dilemma when in pain (doctoral dissertation)*. Belgium: KU Leuven & Ghent University.
- Alves, M., Krypotos, A. M., Crombez, G., & Vlaeyen, J. W. S. (2023). Experimental pain picture system (EPPS): Development and validation. *The Journal of Pain*, 24(11), 2052–2062. <https://doi.org/10.1016/j.jpain.2023.06.014>
- Aylward, J., Valton, V., Ahn, W. Y., Bond, R. L., Dayan, P., Roiser, J. P., & Robinson, O. J. (2019). Altered learning under uncertainty in unmedicated mood and anxiety disorders. *Nature Human Behaviour*, 3(10), 1116–1123. <https://doi.org/10.1038/s41562-019-0628-0>
- Azoulay-Schwartz, R., Kraus, S., & Wilkenfeld, J. (2004). Exploitation vs. exploration: Choosing a supplier in an environment of incomplete information. *Decision Support Systems*, 38(1), 1–18. [https://doi.org/10.1016/S0167-9236\(03\)00061-7](https://doi.org/10.1016/S0167-9236(03)00061-7)
- Buhr, K., & Dugas, M. J. (2002). The intolerance of uncertainty scale: Psychometric properties of the English version. *Behaviour Research and Therapy*, 40(8), 931–945. [https://doi.org/10.1016/S0005-7967\(01\)00092-4](https://doi.org/10.1016/S0005-7967(01)00092-4)
- Carver, C. S., & White, T. L. (1994). Behavioral inhibition, behavioral activation, and affective responses to impending reward and punishment: The BIS/BAS scales. *Journal of Personality and Social Psychology*, 67(2), 319–333. <https://doi.org/10.1037/0022-3514.67.2.319>
- Charmanas, K., Mittas, N., & Angelis, L. (2023). Exploitation of vulnerabilities: A topic-based machine learning framework for explaining and predicting exploitation. *Information*, 14(7), 403. <https://doi.org/10.3390/info14070403>
- Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society of London - Series B: Biological Sciences*, 362(1481), 933–942. <https://doi.org/10.1098/rstb.2007.2098>
- Craske, M. G., Meuret, A. E., Echiverri-Cohen, A., Rosenfield, D., & Ritz, T. (2023). Positive affect treatment targets reward sensitivity: A randomized controlled trial. *Journal of Consulting and Clinical Psychology*, 91(6), 350–366. <https://doi.org/10.1037/ccp0000805>
- Crombez, G., Van Ryckeghem, D. M. L., Eccleston, C., & Van Damme, S. (2013). Attentional bias to pain-related information: A meta-analysis. *Pain*, 154(4), 497–510. <https://doi.org/10.1016/j.pain.2012.11.013>
- Dahlhamer, J., Lucas, J., Zelaya, C., Nahin, R., Mackey, S., DeBar, L., Kerns, R., Von Korff, M., Porter, L., & Helmick, C. (2018). Prevalence of chronic pain and high-impact chronic pain among adults - United States, 2016. *MMWR. Morbidity and Mortality Weekly Report*, 67(36), 1001–1006. <https://doi.org/10.15585/mmwr.mm6736a2>
- Donthula, D., Kortlever, J. T. P., Ring, D., Donovan, E., Reichel, L. M., & Vagner, G. A. (2020). Does intolerance of uncertainty affect the magnitude of limitations or pain intensity? *Clinical Orthopaedics and Related Research*, 478(2), 381–388. <https://doi.org/10.1097/CORR.0000000000001042>
- Fredrickson, B. L. (2001). The role of positive emotions in positive psychology. The broaden-and-build theory of positive emotions. *American Psychologist*, 56(3), 218–226. <https://doi.org/10.1037/0003-066x.56.3.218>
- Fredrickson, B. L., & Branigan, C. (2005). Positive emotions broaden the scope of attention and thought-action repertoires. *Cognition & Emotion*, 19(3), 313–332. <https://doi.org/10.1080/02699930441000238>
- Fusaro, M., Tieri, G., & Aglioti, S. M. (2016). Seeing pain and pleasure on self and others: Behavioral and psychophysiological reactivity in immersive virtual reality. *Journal of Neurophysiology*, 116(6), 2656–2662. <https://doi.org/10.1152/jn.00489.2016>
- Hess, L. E., Haimovici, A., Muñoz, M. A., & Montoya, P. (2014). Beyond pain: Modeling decision-making deficits in chronic pain. *Frontiers in Behavioral Neuroscience*, 8, 263. <https://doi.org/10.3389/fnbeh.2014.00263>
- Kembro, Jackelyn M., Lihoreau, Mathieu, Garriga, Joan, Raposo, Ernesto P., & Bartumeus, Frederic (2019). Bumblebees learn foraging routes through exploitation–exploration cycles. *Journal of The Royal Society Interface*, 16(156), 20190103. <https://doi.org/10.1098/rsif.2019.0103>. <https://royalsocietypublishing.org/doi/10.1098/rsif.2019.0103>. (Accessed 10 July 2019)
- Kruschke, J. K. (2011). Bayesian assessment of null values via parameter estimation and model comparison. *Perspectives on Psychological Science: A Journal of the Association for Psychological Science*, 6(3), 299–312. <https://doi.org/10.1177/1745691611406925>
- Krypotos, A. M., Crombez, G., Alves, M., Claes, N., & Vlaeyen, J. W. S. (2022). The exploration-exploitation dilemma in pain: An experimental investigation. *Pain*, 163(2), e215–e233. <https://doi.org/10.1097/j.pain.0000000000002352>
- Krypotos, A. M., Crombez, G., & Vlaeyen, J. W. S. (2024). The dynamics of pain avoidance: The exploration-exploitation dilemma. *Pain*, 165(10), 2145–2149. <https://doi.org/10.1097/j.pain.0000000000003222>
- Lin, Z., Cabrera-Haro, L. E., & Reuter-Lorenz, P. A. (2020). Asymmetrical learning and memory for acquired gain versus loss associations. *Cognition*, 202, Article 104318. <https://doi.org/10.1016/j.cognition.2020.104318>
- Markfelder, T., & Pauli, P. (2020). Fear of pain and pain intensity: Meta-analysis and systematic review. *Psychological Bulletin*, 146(5), 411–450. <https://doi.org/10.1037/bul0000228>
- McNeil, D. W., & Rainwater, A. J., 3rd (1998). Development of the fear of pain Questionnaire–III. *Journal of Behavioral Medicine*, 21(4), 389–410. <https://doi.org/10.1023/a:1018782831217>
- Mehlhorn, S., Newell, B. R., Todd, P. M., Lee, M., Morgan, K., Braithwaite, V., Hausmann, D., Fiedler, K., & González, C. (2015). Unpacking the exploration–exploitation tradeoff: A synthesis of human and animal literatures. *Decision*, 2(3), 191–215. <https://doi.org/10.1037/dec0000033>
- Peirce, J., Gray, J. R., Simpson, S., MacAskill, M., Höchenberger, R., Sogo, H., Kastman, E., & Lindeløv, J. K. (2019). PsychoPy2: Experiments in behavior made easy. *Behavior Research Methods*, 51(1), 195–203. <https://doi.org/10.3758/s13428-018-01193-y>
- Preis, M. A., & Kroener-Herwig, B. (2012). Empathy for pain: The effects of prior experience and sex. *European Journal of Pain*, 16(9), 1311–1319. <https://doi.org/10.1002/j.1532-2149.2012.00119.x>. <https://onlinelibrary.wiley.com/doi/10.1002/j.1532-2149.2012.00119.x>. (Accessed 21 February 2012)
- R Core Team. (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. <https://www.r-project.org/>.
- Rasmussen, E. B., & Newland, M. C. (2008). Asymmetry of reinforcement and punishment in human choice. *Journal of the Experimental Analysis of Behavior*, 89(2), 157–167. <https://doi.org/10.1901/jeab.2008.89-157>
- Roux, L., Gustin, S. M., & Newton-John, T. R. O. (2022). To persist or not to persist? The dilemma of goal adjustment in chronic pain. *Pain*, 163(5), 820–823. <https://doi.org/10.1097/j.pain.0000000000002503>
- Rowe, G., Hirsh, J. B., & Anderson, A. K. (2007). Positive affect increases the breadth of attentional selection. *Proceedings of the National Academy of Sciences of the United States of America*, 104(1), 383–388. <https://doi.org/10.1073/pnas.0605198104>
- Saragosa-Harris, N. M., Cohen, A. O., Reneau, T. R., Villano, W. J., Heller, A. S., & Hartley, C. A. (2022). Real-world exploration increases across adolescence and relates to affect, risk taking, and social connectivity. *Psychological Science*, 33(10), 1664–1679. <https://doi.org/10.1177/09567976221102070>
- Serrano-Ibáñez, E. R., Ramírez-Maestre, C., Esteve, R., & López-Martínez, A. E. (2019). The behavioural inhibition system, behavioural activation system and experiential avoidance as explanatory variables of comorbid chronic pain and posttraumatic stress symptoms. *European Journal of Psychotraumatology*, 10(1), Article 1581013. <https://doi.org/10.1080/20008198.2019.1581013>
- Seymour, B., Daw, N. D., Roiser, J. P., Dayan, P., & Dolan, R. (2012). Serotonin selectively modulates reward value in human decision-making. *Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 32(17), 5833–5842. <https://doi.org/10.1523/JNEUROSCI.0053-12.2012>
- Smith, B. H., Penny, K. L., Purves, A. M., Munro, C., Wilson, B., Grimshaw, J., Chambers, W. A., & Smith, W. C. (1997). The Chronic pain grade questionnaire: Validation and reliability in postal research. *Pain*, 71(2), 141–147. [https://doi.org/10.1016/S0304-3959\(97\)03347-2](https://doi.org/10.1016/S0304-3959(97)03347-2)
- Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction* (2nd ed.).
- Todd, J., van Ryckeghem, D. M. L., Sharpe, L., & Crombez, G. (2018). Attentional bias to pain-related information: A meta-analysis of dot-probe studies. *Health Psychology Review*, 12(4), 419–436. <https://doi.org/10.1080/17437199.2018.1521729>
- Uotila, J., Maula, M., Keil, T., & Zahra, S. A. (2008). Exploration, exploitation, and financial performance: Analysis of S&P 500 corporations. *Strategic Management Journal*, 30(2), 221–231. <https://doi.org/10.1002/smj.738>
- Vehtari, A., Gelman, A., & Gabry, J. (2016). Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5), 1413–1432. <https://doi.org/10.48550/arXiv.1507.04544>
- Vlaeyen, J. W. S., & Crombez, G. (2020). Behavioral conceptualization and treatment of chronic pain. *Annual Review of Clinical Psychology*, 16, 187–212. <https://doi.org/10.1146/annurev-clinpsy-050718-095744>
- Wang, H., Zariwopoulou, T., & Zhou, X. (2018). *Exploration versus exploitation in reinforcement learning: A stochastic control approach*. <https://doi.org/10.48550/arXiv.1812.01552>
- Zaman, J., Van Oudenhove, L., & Vlaeyen, J. W. S. (2021). Uncertainty in a context of pain: Disliked but also more painful? *Pain*, 162(4), 995–998. <https://doi.org/10.1097/j.pain.0000000000002106>
- Zhang, L., Vashisht, H., Nethra, A., Slattey, B., & Ward, T. (2022). Differences in learning and persistency characterizing behavior in chronic pain for the Iowa gambling task: Web-based laboratory-in-the-field study. *Journal of Medical Internet Research*, 24(4), Article e26307. <https://doi.org/10.2196/26307>