Research Paper

PAIN®

The exploration–exploitation dilemma in pain: an experimental investigation

Angelos-Miltiadis Krypotos[a,b,*], Geert Crombez[c], Maryna Alves[a], Nathalie Claes[a], Johan W.S. Vlaeyen[a,d]

Abstract

Daily life consists of a chain of decisions. Typically, individuals may choose to pursue what they already know (*exploitation*) or to search for other options (*exploration*). This *exploration–exploitation dilemma* is a topic of interest across multiple scientific fields. Here we propose that investigating how individuals solve this dilemma may improve our understanding of how individuals make behavioral decisions (eg, avoidance) when facing pain. To this end, we present the data of 3 experiments in which healthy individuals were given the opportunity to choose between 4 different movements, with each movement being associated with different probabilities of receiving a painful outcome only (experiment 1) or pain and/or a reward (experiment 2). We also investigated whether participants stuck to their decisions when the contingencies between each movement and the painful/rewarding outcome changed during the task (experiment 3). The key findings across all experiments are the following: First, after initial exploration, participants most often exploited the safest option. Second, participants weighted rewards more heavily than receiving pain. Finally, after receiving a painful outcome, participants were more inclined to explore than to exploit a rewarding movement. We argue that by focusing more on how individuals in pain solve the exploration–exploitation dilemma is helpful in understanding behavioral decision making in pain.

Keywords: Reinforcement learning, Fear, Avoidance, Positive reinforcement, Negative reinforcement

1. Introduction

Pain motivates escape and drives learning, including avoidance learning (ie, behavior that results in the elimination of anticipated pain).[49] Avoidance may be considered as the result of decision-making processes, where the individual considers different actions before attempting to perform one. To decide which action to perform, individuals typically enter a cycle of *exploration* (ie, actions that lead to the acquisition of new information) and *exploitation* (ie, actions that do not change the current information status).[32] The exploration–exploitation dilemma has been studied in several scientific areas (eg, computer science[47] and computational psychiatry).[19,51]

We propose that studying the exploration–exploitation dilemma is valuable for research into pain, and chronic pain in particular.[49] It has been suggested that individuals with chronic pain may persist in choosing actions that lead to less pain, or less pain exacerbation (eg, by avoidance), at the cost of pursuing valuable goals. Such behavior is an example of exploiting current knowledge. Exploration, on the other hand, might result in the discovery of alternative actions that could lead to valuable outcomes. Investigating the exploration–exploitation dilemma in pain could be useful in studying how avoidance is acquired and extinguished in acute and chronic pain.

With this in mind, the primary goal of this article is explorative in nature and aimed to describe the exploration–exploitation dilemma in the context of pain through an experimental paradigm in healthy volunteers using computational models. In this way behavioral decision making is broken down into the relevant latent parameters (eg, how sensitive participants are to punishment or rewards; refer to Methods section for more information). (Our research does not follow the traditional hypothesis-testing framework, in which a hypothesis is tested using traditional inference frameworks [eg, null hypothesis significance testing or Bayesian hypothesis testing]. Instead, we used various computational models to describe the pattern of responses.) Given that this is the first of its kind, we decided to focus on acute experimental pain; further research could include individuals with chronic pain. To achieve our goal, we integrated 2 tasks: the classic *n*-bandit task[9,20,47] and the pain avoidance task.[36] We subsequently broke down the participant's performance into various parameters, using several mathematical reinforcement models.[2]

The secondary goal was to *explore* the correlation between the parameters of the model and individual characteristics that have been associated with avoidance. For this, we collected self-report data regarding intolerance of uncertainty,[4,18] neuroticism,[12] fear of pain,[48] and behavioral inhibition/activation.[14]

This article describes 3 experiments. In experiment 1, we tested how individuals solve the exploration–exploitation dilemma when moving a joystick in different directions on a computer screen, with each movement being associated with a painful outcome (*positive punishment*) or a nonpainful outcome (*negative reinforcement*). In experiment 2, each movement was associated

Sponsorships or competing interests that may be relevant to content are disclosed at the end of this article.

[a] *Research Group Health Psychology, KU Leuven, Leuven, Belgium,* [b] *Department of Clinical Psychology, Utrecht University,* [c] *Department of Experimental-Clinical and Heath Psychology, Ghent University, Ghent, Belgium,* [d] *Experimental Health Psychology, Maastricht University, Maastricht, the Netherlands*

**Corresponding author. Address: Research Group Health Psychology, KU Leuven, Tiensestraat 102, 3000 Leuven, Belgium. Tel.: +32 16 19 44 25. E-mail address: amkrypotos@gmail.com (A.-M. Krypotos).*

Supplemental digital content is available for this article. Direct URL citations appear in the printed text and are provided in the HTML and PDF versions of this article on the journal's Web site (www.painjournalonline.com).

PAIN 163 (2022) e215–e233

http://dx.doi.org/10.1097/j.pain.0000000000002352

with different probabilities of pain (*positive punishment*) and/or a reward (*positive reinforcement*). In experiment 3, we investigated whether individuals modified their performance when the contingencies between each movement and the outcome changed halfway through the experiment. Given that the specific tasks were being used for the first time, all analyses were preregistered as exploratory.[27]

1.1. Method of experiment 1

1.1.1. Participants

Fifty participants were recruited (sex: 11 males and 39 females, age: M = 20.90, SD = 3.70). The exclusion criteria were (based on self-reports) age outside the range of 18 to 35 years; insufficient command of the Dutch language; inability to move arms, hands, or shoulders; a cardiovascular problem; a neurological disorder; a psychiatric disorder; any other serious medical problem; pregnancy; use of drugs (eg, cannabis); acute or chronic pain in hands, arms, shoulders, or other related areas; recovery from a trauma or surgery; use of medication that could influence the central nervous system; use of an electric implant (eg, pacemaker); hearing problems; and disordered vision that cannot be corrected by means of glasses or contact lenses (eg, color blindness). We did not run a power analysis, given that our main analyses are not based on frequentist inference (see below). Although, to the best of our knowledge, there are no simulation studies for evaluating how many trials and participants are required for valid parameter estimation, we followed the following 3 guides: First, we included multiple trials per participant (ie, 300 trials for experiments 1 and 2 and 600 trials for experiment 3) and a relatively large sample size (ie, 50 participants per experiment). Second, we used a Bayesian version of all computational models, which enables the inclusion of previous information, apart from the data, in all parameter estimations. Finally, we used a hierarchical version of the model, in which the individual parameters were informed by the group parameters and vice versa. In the absence of clear-cut guidelines regarding optimal sample sizes, we believe we took sufficient steps to reach meaningful conclusions. The study was approved by the Ethics Committee of KU Leuven (#G-2018 12 1444). The study's preregistration, all experimental material, and data are available at https://osf.io/32m5p/. The preregistration of all studies was conducted before data collection.

1.1.2. Questionnaires

All participants filled in Dutch translations of the following questionnaires: The Intolerance of Uncertainty Scale (long version),[4,17] the Neuroticism scale of the Eysenck Personality Questionnaire,[12] the Fear of Pain Questionnaire,[54] and the Behavioral Inhibition or Activation scales.[13]

1.1.3. Ratings

At different points in the experiment, participants rated the painfulness of the electrocutaneous stimulus using a 10-point painfulness rating scale (ie, How painful did you find this stimulus? 1: no sensation at all; 10: the worst pain imaginable). The scale was explained to the participants as follows: "1" indicated that they felt nothing; "2" that they first perceived the electrocutaneous stimulus as aversive; "3" that the sensation was starting to become aversive and slightly painful; "8" that the sensation was moderately painful and demanded some effort to tolerate; and "10" that the sensation was the worst pain they could imagine. We aimed for a rating of at least 8 across all participants, and the calibration staircase procedure ended when this level was reached[36] (ie, for details on the calibration procedure refer to the "Stimulus" section). Participants also rated the electrocutaneous stimulus based on its unpleasantness (ie, How unpleasant do you find this stimulus? 1: not unpleasant at all; 10: extremely unpleasant) and their tolerance of it (ie, How difficult was it for you to endure the electrocutaneous stimulus? 1: not difficult at all; 10: extremely difficult).

Participants rated the contingencies between each square or stimulus with each stimulus referring to the movements that the participants had to make and the presentation of a painful electrocutaneous stimulus in terms of expectancy of receiving an electrocutaneous stimulus (ie, *To what extent do you expect to receive an electrocutaneous stimulus after you have selected the blue square?* 1: not at all; 10: very much). They also rated their willingness to select a given square (ie, *To what extent did you want to select the blue square?* 1: not at all; 10: very much) and their fear of each square (ie, *How afraid were you to select the blue square?* 1: not at all; 10: very much).

1.1.4. Stimuli

Four white squares, presented in each of the 4 corners of a computer screen, served as indicators of the target movement location. In the middle of the screen, a purple circle was presented. The electrocutaneous stimulus was applied to the wrist of the nondominant hand using 2.8-mm diameter bar electrodes with 30 mm spacing in between them. Sigma gel was applied between the electrodes and the participant's wrist. Before the beginning of the main experiment, we followed a staircase calibration procedure to individually determine the intensity of the electrocutaneous stimulation. For this, we started from an intensity of 1 mA and increased the intensity until participants judged the stimulus to be "moderately painful and demanded some effort to tolerate" (ie, rating of 8) using the painfulness rating scale presented in the "Ratings" section. The possible intensities of the electrocutaneous stimulus were the following: 1, 2, 4, 6, 8, 11, 14, 17, 20, 24, 28, 32, 36, 40, 44, 48, and 52 mA. Participants could give their responses using a joystick (Logitech Attack 3). The experiment was programmed in PsychoPy.[37]

1.1.5. Procedure

At least one day before the experiment, participants filled in the questionnaires online using LimeSurvey.[30] For this online session, participants read an information brochure about the questionnaires and provided informed consent online. If participants did not meet any of the exclusion criteria, they were invited for the experimental session. We report the data of all participants who participated in the full experiment.

On the day of the experiment, participants read a new information brochure and provided informed consent for the experiment. The electrodes for the electrocutaneous stimulus were attached on the participant's nondominant hand, and the level of the electrocutaneous stimulus was determined. Participants then received instructions about the experiment, both on screen and orally.

The main experimental task consisted of 300 trials, divided into 2 blocks of 150 trials each (refer to the left panel of **Fig. 1**). Before the main task, participants completed a practice block of 20 trials, during which no painful outcome was administrated. Participants were requested to focus on the fixation cross at the center of the screen. In

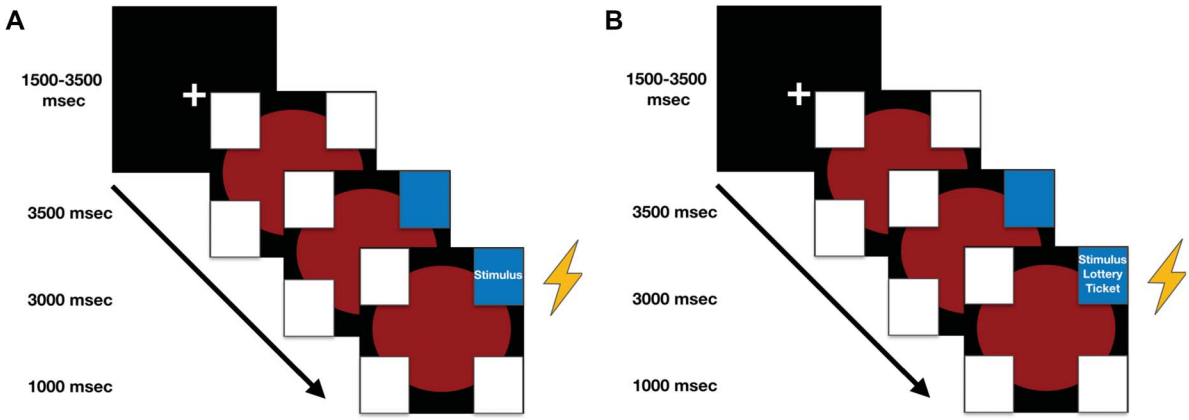

Figure 1. (A) Trial sequence for experiment 1: Participants had 3000 milliseconds to move the joystick to one of the 4 squares. After they made a selection, the square corresponding to the participant's move turned blue for 3500 milliseconds. In case of a painful stimulus, participants read the word "prikkel" (stimulus) on screen for 1000 milliseconds and simultaneously received the painful stimulus. The jittered intertrial intervals ranged from 1500 to 3500 milliseconds where participants saw a black screen and a white fixation cross. (B) Trial sequence for experiments 2 and 3: The trial sequence is identical to experiment 1 except the following: after participants made a choice, and in case of a reward, participants read the words "loterij ticket" (lottery ticket) on screen for 1000 milliseconds.

each trial, after the disappearance of the fixation cross, participants could move the joystick to any of the squares placed in the 4 corners of the screen (max 3000 milliseconds). After they had made their selection, the square in question turned blue and remained on screen for 3000 milliseconds. In the case of pain, the electrocutaneous stimulus was administered and simultaneously the word "prikkel" (Dutch for "stimulus") was displayed on screen for 1000 milliseconds. Each stimulus was associated with a different mean probability of a painful outcome. Specifically, the 4 stimuli were associated with 10%, 30%, 50%, and 70% probability, respectively, of receiving a painful outcome. Which stimulus was associated with each probability was determined randomly for each participant and remained fixed for the whole experiment. The intertrial intervals were jittered, ranging from 1500 to 3500 milliseconds. At the end of each block, participants were asked to evaluate each visual stimulus, as well as the electrocutaneous stimulus, using the visual analogue scales described above. All ratings were provided with references to the block that had just ended.

1.1.6. Statistical analyses

At the end of the experiment, the electrodes were removed, and participants were thanked, debriefed, and compensated with 8 euros or 1 research credit.

For the self-report data, we analyzed all rating scales (see "ratings" subsections) using 4 × 2 repeated measures analyses of variance (ANOVAs), with stimulus and block as the within-subject factors: stimulus (10% vs 30% vs 50% vs 70%) × block (block 1 vs block 2). To test how often participants switched between the different options we ran a 2 × 2 repeated measures ANOVA: switch vs no-switch × block 1 vs block 2. We followed up all ANOVAs with pairwise comparisons with the Holm correction (for the sake of brevity, the exact statistics of all the pairwise comparisons are reported in the supplementary material, available at http://links.lww.com/PAIN/B394; refer to Tables 1-9 for experiment 1, Tables 10-21 for experiment 2, and Tables 22-33 for experiment 3). After our preregistration plan, we did not remove any outliers from our data.

We ran our analyses using null hypothesis significance testing (NHST) in combination with Bayes factors. Null hypothesis significance testing is the standard inferential approach across experimental studies, something that allows to easily compare studies with each other as well as, because of its popularity, makes communication of statistics between researchers easy. We have previously argued for the use of Bayes factors in addition to NHST for experimental data.[23,24,26] Bayes factors are more informative than *P* values because, in comparison to NHST, they can provide (1) relative evidence for 2 different hypotheses (ie, null vs alternative), (2) evidence for the absence of an effect, and (3) provide a continuous, rather than a dichotomous (significant vs nonsignificant), grading of evidence. For a thorough discussion of Bayes factors see Ref. 11,23,25. For computing the Bayes factors, we used the BayesFactor R package,[45] which computes Bayes factors for Bayesian *t* tests and repeated measures ANOVAs. This package requires the user to define the previous distributions of the alternative hypothesis. For this, we used a Cauchy distribution with the scale factor set to the value of $s = 0.707$, the default option of the package and the same option we used in the previous work. However, we also ran the analyses using other scale factors (ie, 1 and 1.4), and the direction of the results remained the same. For the Bayes factors, we treated values above 1 as strong support for the hypothesis that the data correlate with each other and values below 1 as strong support for the reverse hypothesis.[23,26]

For the computational modelling analyses, we fitted 5 models using the hBayesDM[1] package for R,[41] each model having a different set of parameters relevant to n-bandit tasks (ie, punishment/reward sensitivity, punishment/reward learning rate, noise, decay rate, and learning rate). The *sensitivity* parameter indicates how much participants expect to like a reward or dislike a painful outcome. The *learning rate* parameter reflects how quickly information from previous trials is integrated, or in other words, how quickly an individual persists in choosing the same stimulus (ie, square) after the presentation of an aversive outcome or how long the individual chooses the same stimulus after a reward. The *lapse* parameter shows the degree of randomness in the choices of participants and may indicate exploration due to uncertainty. Finally, the *decay* parameter reveals the degree to which participants forget the values of the different options after not choosing them. High decay values may also indicate more exploration, as participants seek additional information. In **Table 1**, we refer to the models as referred to in the hBayesDM R package and the parameters that each model refers to. Please note that similar models have been used in previous studies using the n-bandit task.[2] For all models, *punishment* was defined as the

Table 1

Leave-out-one information criterion values per model for experiment 1.

Model	Parameters	LOOIC
bandit4arm_lapse_decay	Reward learning rate, punishment learning rate, reward sensitivity, punishment sensitivity, noise, and decay rate	22,256.77
bandit4arm_lapse	Reward learning rate, punishment learning rate, reward sensitivity, punishment sensitivity, and noise	22,717.28
bandit4arm_4par	Reward learning rate, punishment learning rate, reward sensitivity, andpunishment sensitivity	22,721.39
bandit4arm_singleA_lapse	Learning rate, reward sensitivity, punishment sensitivity, and noise	23,283.16
bandit4arm_2par_lapse	Reward sensitivity, punishment sensitivity, choice persistence, and noise	32,468.27

The winning model is the one with the lowest LOOIC value.
LOOIC, leave-out-one information criterion.

presence of painful outcome (*positive punishment*) and *reward* the absence of painful outcome (*negative reinforcement*). To evaluate which model fitted the data best, we used the leave-out-one information criterion (LOOIC).[48] Leave-out-one information criterion entails training the winning model on all but one observation in the data set. After the execution of the same procedure for all observations in the data set, the performance of the model is assessed by calculating the sum of all the computed scores. The model that fits the data best (ie, the *winning model*) is the model with the lowest LOOIC value. After choosing the winning model, we assessed whether the Markov chain Monte Carlo (MCMC) methods for all the model parameters converged using the *R–hat* criterion,[15] with values below 1.1 suggesting successful convergence. We also plotted the different MCMC methods, with those looking like "*fat hairy caterpillars*" suggesting successful convergence.[29] Finally, we simulated data according to the winning model parameters and plotted the predicted values against the real data to evaluate whether the winning model parameters predicted the collected data (ie, *postpredictive checks*). The postpredictive checks' plot for experiment 1 is available in Supplementary Fig. 1 (available at http://links.lww.com/PAIN/B394).

We also wish to explore whether participants learn faster after a rewarding outcome or a punishing outcome, as well as how sensitive participants are to rewards vs punishments. To this end, we compared the learning and sensitivity parameters, first computing the differences between the posterior distributions of the learning and sensitivity parameter values and then plotting the *region of practical equivalence* (ROPE) of these distributions.[22] We then evaluated whether the differences in the learning and sensitivity parameters were relevant for the winning model by plotting the posterior distributions of the differences and computing the percentage that fell outside the area of practical relevance—here the 0 point. If 95% of the distribution was outside the ROPE, then the 0 value was rejected and the differences were judged to be relevant. If 95% of the posterior distribution fell within the ROPE, the 0 value was accepted and the differences were judged to be irrelevant.

Finally, to test whether individual differences might relate to any of the parameters of the model, we performed both standard frequentist correlations and Bayesian correlations using the BayesFactor R package.[42] For the NHST analyses, we used an *alpha* level of 0.01 for the individual difference analyses. To reduce the probability of a type 1 error, we selected a lower *alpha* level than the traditional 0.05 value. Similar to the ANOVAs, for the Bayesian correlation we used the default values of the correlationBF function for our priors, again from the BayesFactor package, and ran sensitivity analyses by changing the scale of the *r* parameter of the Cauchy distribution. The direction of the results again did not change considerably, so we only report the results where the default options were used.

1.1.7. Descriptive statistics

1.1.7.1. Results of experiment 1

Figure 2 depicts the mean responses on the rating scales across all blocks of the experiment.

Before the beginning of the experimental task, the electrocutaneous stimulus used for pain was judged to be painful (M = 8.39, SD = 0.62), unpleasant (M = 8.33, SD = 0.77), and difficult to tolerate (M = 8.20, SD = 0.85). Although there was a significant drop in all ratings when the painful outcome was evaluated after the completion of the first block—all $ps < 0.05$, BFs > 2—all ratings were still high: painfulness (M = 7.43, SD = 1.12), unpleasantness (M = 8, SD = 1.02), and tolerance difficulty (M = 6.89, SD = 1.71).

Regarding the expectancies of the electrocutaneous stimulus, there was a main effect of stimulus, F (2.43, 119.07) = 8.51, $P < 0.001$, $\eta^2_G = 0.108$; $BF_{10} > 1000$, and a main effect of block, F (1.00, 49.00) = 6.55, $P = 0.014$, $\eta^2_G = 0.008$; $BF_{10} = 0.531$. There was no significant stimulus × block interaction, F (2.85, 139.65) = 1.79, $P = 0.155$, $\eta^2_G = 0.004$, $BF_{01} = 19.939$. Pairwise comparisons showed significant differences between the 10% and 50% stimulus; 10% and 70% stimulus; 30% and 50% stimulus; and 30% and 70% stimulus.

A similar pattern of results emerged when participants rated how much they desired to select each square or stimulus. There was a main effect of stimulus, F (2.49, 122.01) = 7.57, $P < 0.001$, $\eta^2_G = 0.096$; $BF_{10} > 1000$, and block, F (1.00, 49.00) = 8.72, $P = 0.005$, $\eta^2_G = 0.008$; $BF_{10} = 0.561$. There was no significant stimulus × block interaction, F (2.79, 136.71) = 1.11, $P = 0.346$, $\eta^2_G = 0.003$, $BF_{01} = 24.178$. The pairwise comparisons showed significant effects between the 10% and the 30% stimulus, the 50% stimulus, and the 70% stimulus. The only other significant difference was between the 30% and the 70% stimulus.

When it came to how much participants were afraid of each square, there was a significant stimulus × block interaction—F (2.85, 139.65) = 2.96, $P = 0.037$, $\eta^2_G = 0.006$—although the Bayesian test showed evidence against this effect, $BF_{01} = 11.995$. The follow-up pairwise comparison revealed significant

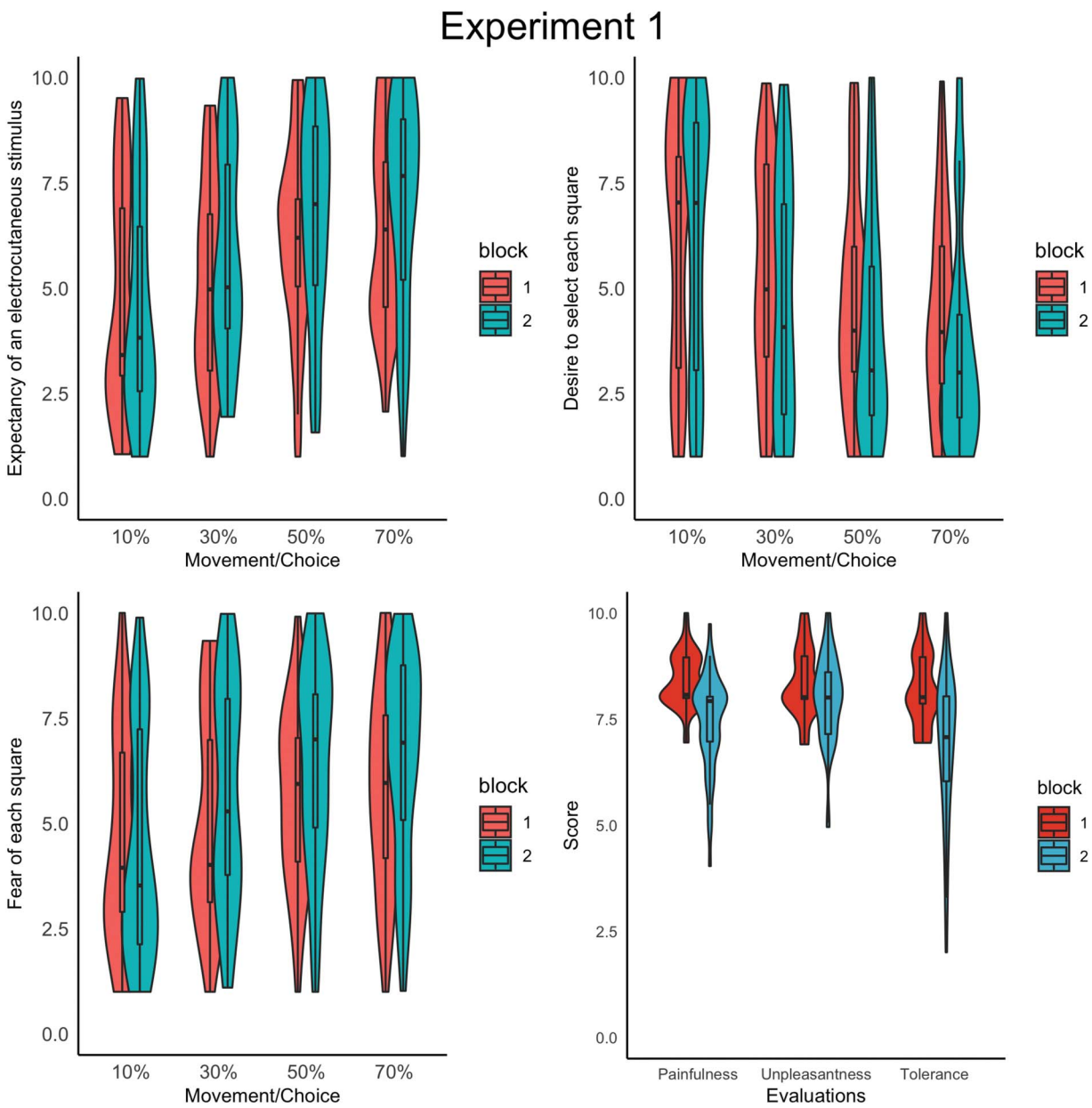

Figure 2. Scores at each rating scale filled at the end of each of the 2 blocks.

differences between the 10% and the 70% stimulus for block 1. For block 2, there were significant differences between the 10% and the 30% stimulus, the 50% stimulus, and the 70% stimulus.

In summary, the results indicate that there were between-stimuli differences between fear of the painful outcome, desire to choose a given square, and expectation of a painful outcome. Please note, however, that there was disagreement between the frequentist and Bayesian analyses, as previously reported in the literature,[53] so these results should be interpreted with caution. Taken together, these findings largely suggest that our manipulation was successful and that participants perceived each stimulus differently according to the probability of the stimulus being followed by a painful outcome.

1.1.8. Performance results

The top panel of **Figure 3** shows the distribution of each choice, while the bottom panel shows the number of trials for each block in which participants switched or did not switch their choices. After computing mean switch vs no-switch responses for each block, we found that participants switched their choices more in the first than in the second block, as revealed by a block × switch (switch vs no-switch) repeated measures ANOVA, F (1.00, 49.00) = 35.34, $P < 0.001$, $\eta^2_G = 0.092$; $BF_{10} > 1000$.

Table 1 reports the LOOIC values for each computational model. The model with the lowest LOOIC values, as well as the one that reached MCMC convergence and sufficient postpredictive checks, is the bandit4arm_lapse_decay model.[2,39]

1.1.9. Correlation with individual differences

Figure 4 displays the plots of the ROPEs of all parameters of the model. In our case, we wanted to compare learning rates and sensitivity values. In both cases, we see that the differences are relevant.

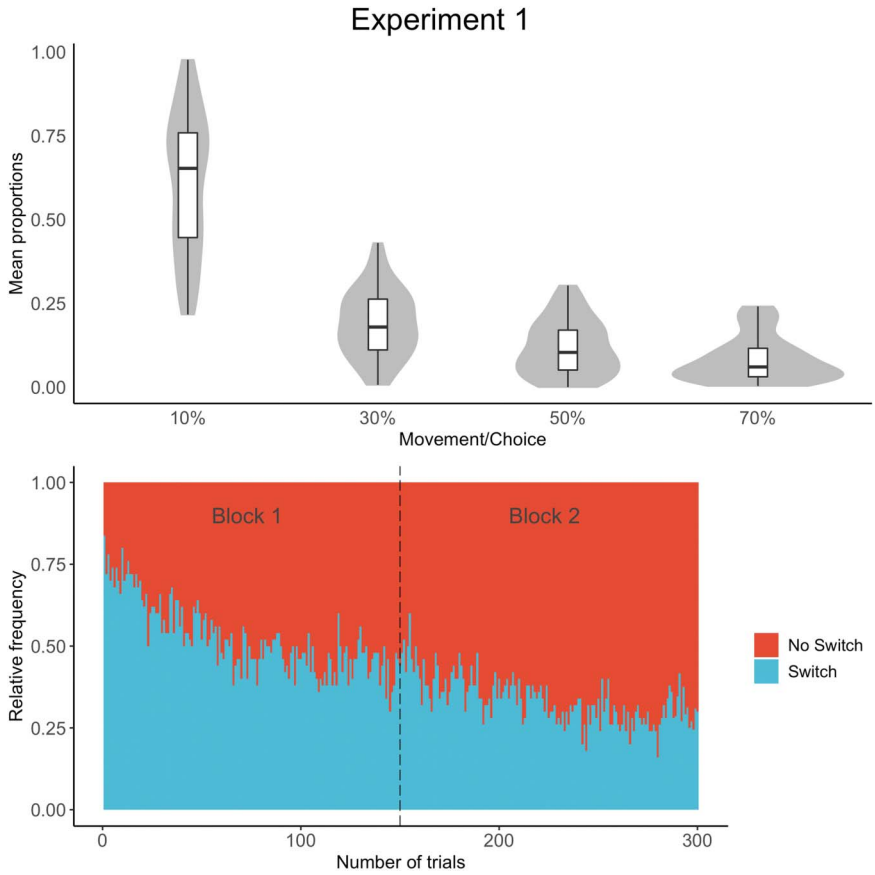

Figure 3. Top panel: Per stimulus means of choosing one of the 4 stimuli. The vertical axis refers to the probabilities of receiving a painful electrocutaneous stimulus after each stimulus for experiment 1. Bottom panel: Proportions of trials, for each block, that participants switched or did not switch their choices for experiment 1.

The supplementary Table 45 (available at http://links.lww.com/PAIN/B394) shows the mean and SDs of all parameter values and all questionnaires for experiment 1. After correlating the parameter values of the winning model and the collected questionnaires, we did not find reliable correlations for any of the variables. Bayes factors yielded results in the same direction, and we found no strong evidence for the data coming from the alternative compared with the null hypothesis. In the supplementary material, we provide the correlation matrix between the collected individual differences and the parameters of the model, as well as the relevant tables with all correlations (refer to Tables 34-44 and Fig. 2 in the supplementary material, available at http://links.lww.com/PAIN/B394).

1.2. Discussion of experiment 1

The first experiment investigated how individuals balanced between exploration and exploitation when moving a cursor towards different on-screen locations, where each location was associated with different probabilities of a painful outcome. Results showed that participants, on average, learned which movement led to the smallest chance of receiving a painful outcome, tending to increasingly repeat these choices in subsequent trials. These results point towards a shift from exploration to exploitation.

The parameters of the model underscore this pattern. First, as revealed by the learning rate parameters, participants tended to be quicker to learn which options were associated with the presentation of the painful outcome: they relied on fewer trials than when learning which options were associated with the absence of the painful outcome. Second, in relation to the sensitivity parameters, the *reward* of avoiding pain was weighted more heavily than the *punishment* of receiving the painful outcome. Third, the low values of both the lapse and decay parameters pointed to low randomness in choices and decay in knowledge, which is also indicative of more exploitation than exploration. Collectively, the results indicate that to avoid a painful outcome, participants exploit more than explore.

Experiment 1 is a first step in investigating the exploration–exploitation dilemma, where rewards were defined as the absence of the painful outcome (negative reinforcement). Experiment 2 extended experiment 1 by delivering actual rewards (ie, positive reinforcement) as possible outcomes of each stimulus or movement.

2. Experiment 2

2.1. Method of experiment 2

2.1.1. Participants

We recruited 50 individuals who had not taken part in experiment 1. The data of one individual were removed because of incomplete responses, reducing the sample to 49 (sex = 8 males and 41 females, age: M = 20.24, SD = 2.37). The same exclusion criteria were applied as in the first experiment. The study was approved by the Ethics Committee of KU Leuven (#G-2018 12 1467). The study's preregistration, the experimental material, and the full data set are available at https://osf.io/5k3yr/.

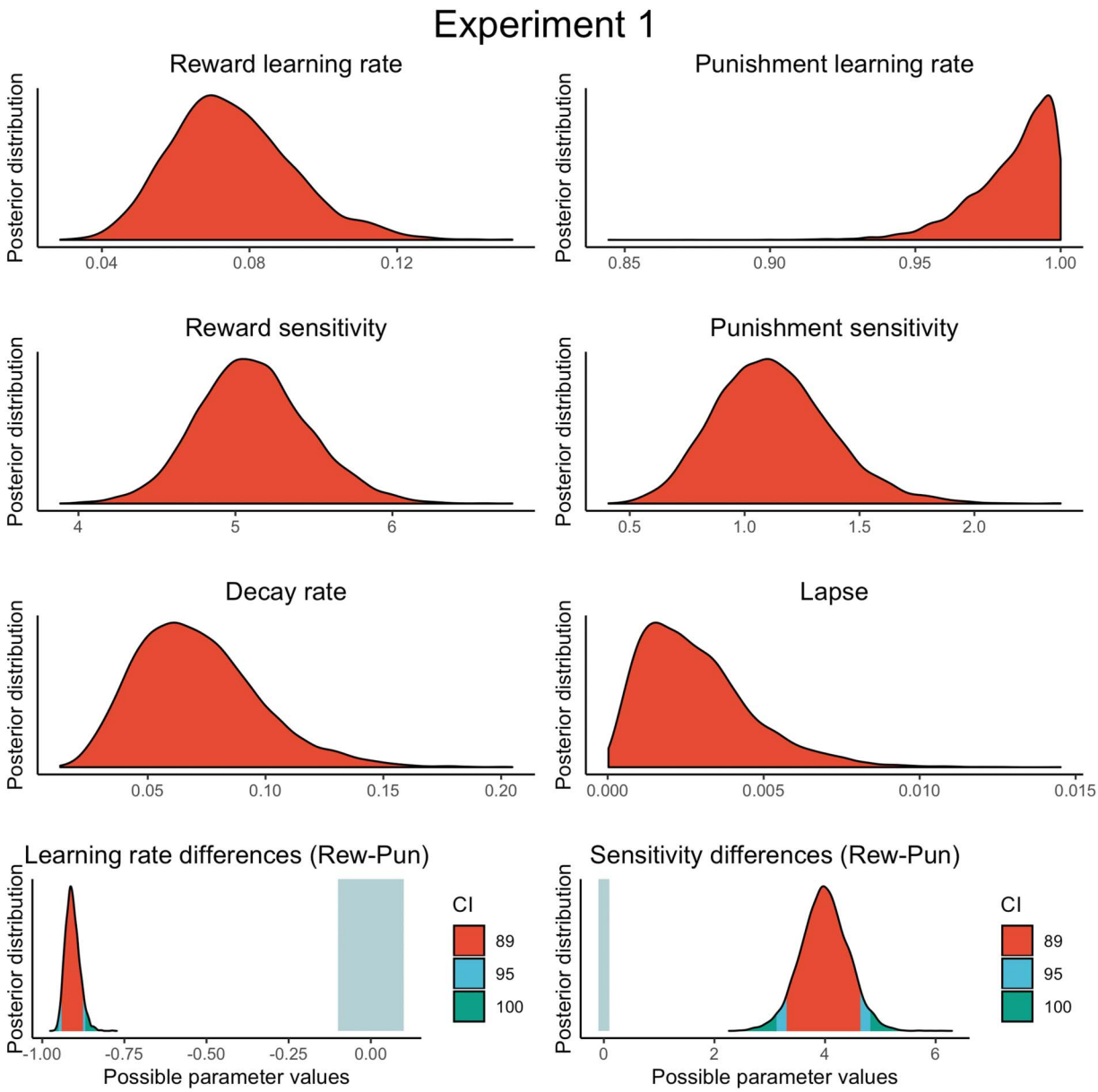

Figure 4. The posterior distributions and ROPEs (bottom 2 panels) for each model parameter for experiment 1. ROPE, region of practical equivalence.

Participants were compensated in a similar manner as in experiment 1.

2.1.2. Material

The same questionnaires, ratings, stimuli, and outcomes were used as in experiment 1, with the following exceptions. First, participants reported on the contingencies between each stimulus and the presentation of both a painful and a rewarding outcome (see below). Second, participants evaluated the rewards,[3,7] here in the form of lottery tickets, in terms of unpleasantness (ie, *How unpleasant did you find the stimulus that was rewarded with the lottery tickets*? 1: not at all unpleasant; 10: extremely unpleasant) (We decided to use the same unpleasantness scale that was used for the electrocutaneous stimulus also for the lottery ticket because this enables a direct comparison of the 2 outcomes. In addition, using one scale rather than using both an unpleasantness and pleasantness scale lowers the burden for the participants), difficulty (ie, *How difficult was it to choose a reward?* 1: not at all difficult at all; 10: extremely difficult), and value (ie, *How valuable is the selected reward to you?* 1: not valuable at all; 10: extremely valuable) and in terms of their interest in winning the reward (ie, *How interested are you in winning the reward?* 1: not interested at all; 10: extremely interested) and their willingness to make an effort (ie, *How much effort are you willing to make to get this reward?* 1: no effort at all; 10: the highest effort).

In addition, at different points in the study, participants were asked to evaluate the square or stimulus that led to the reward in terms of its unpleasantness.

2.1.3. Procedure

The procedure was identical to that of experiment 1, with the following exceptions (refer to the right panel of **Fig. 1**): Before the

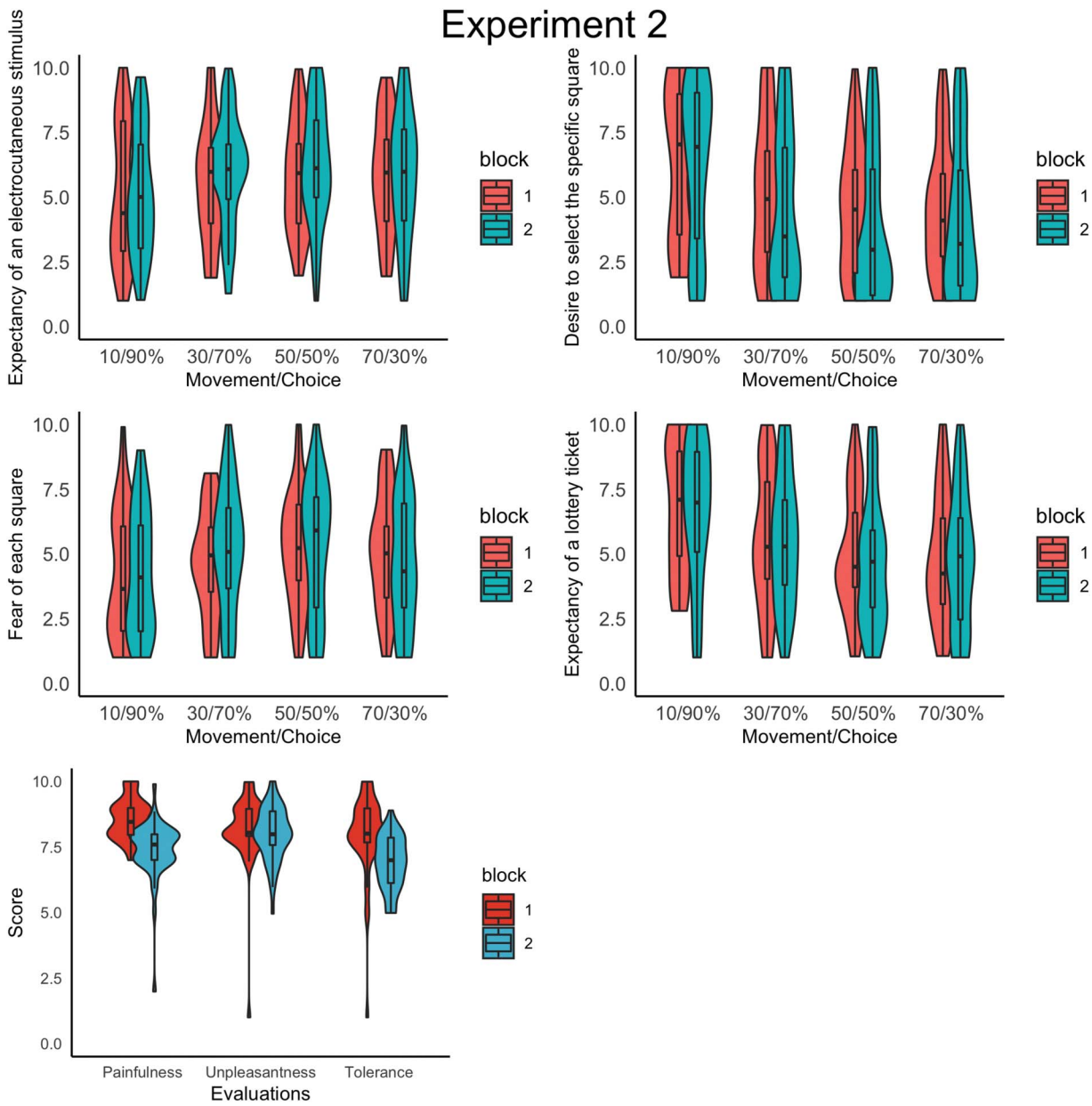

Figure 5. Scores at each questionnaires, filled at the end of each block for experiment 2.

beginning of the experiment, participants were shown a list of rewards and asked to choose which one they found most rewarding. Participants could choose from different gifts—including gift cards, e-readers, or a head massage—with each gift valued up to 100 euros. They were informed that the gift would be given to only one of the participants, and the more frequently they chose a reward, the higher their chances of receiving the gift of their choice. Then, during the main experiment, each square or stimulus was associated with different probabilities of receiving a painful outcome and/or a reward. Specifically, one square or one stimulus had a 10% probability of receiving pain and a 90% probability of receiving a reward. For the other 3 stimuli, the probability of receiving pain increased to 30%, 50%, and 70%, respectively, with the probability of receiving the reward decreasing to 70%, 50%, and 30% respectively. Finally, when participants received the rewarding outcome, the words "loterij ticket" (Dutch for *lottery ticket*) were displayed.

2.1.4. Statistical analyses

We followed the same analytical approach as in experiment 1, with the exception that now the reward was defined as the presence of the lottery ticket (*positive reinforcement*).

2.2. Results of experiment 2

2.2.1. Descriptive statistics

Figure 5 depicts mean scores for each of the visual analogue scales across blocks.

Before the beginning of the experimental task, the electrocutaneous stimulus had been judged as painful (M = 8.56, SD = 0.75), unpleasant (M = 8.18, SD = 1.30), and difficult to tolerate (M = 7.98, SD = 1.53). Although there was a significant drop in painfulness and tolerance difficulty—all *P*s $<$ 0.01, BFs $>$ 1000—all ratings were still high: painfulness (M = 7.39, SD =

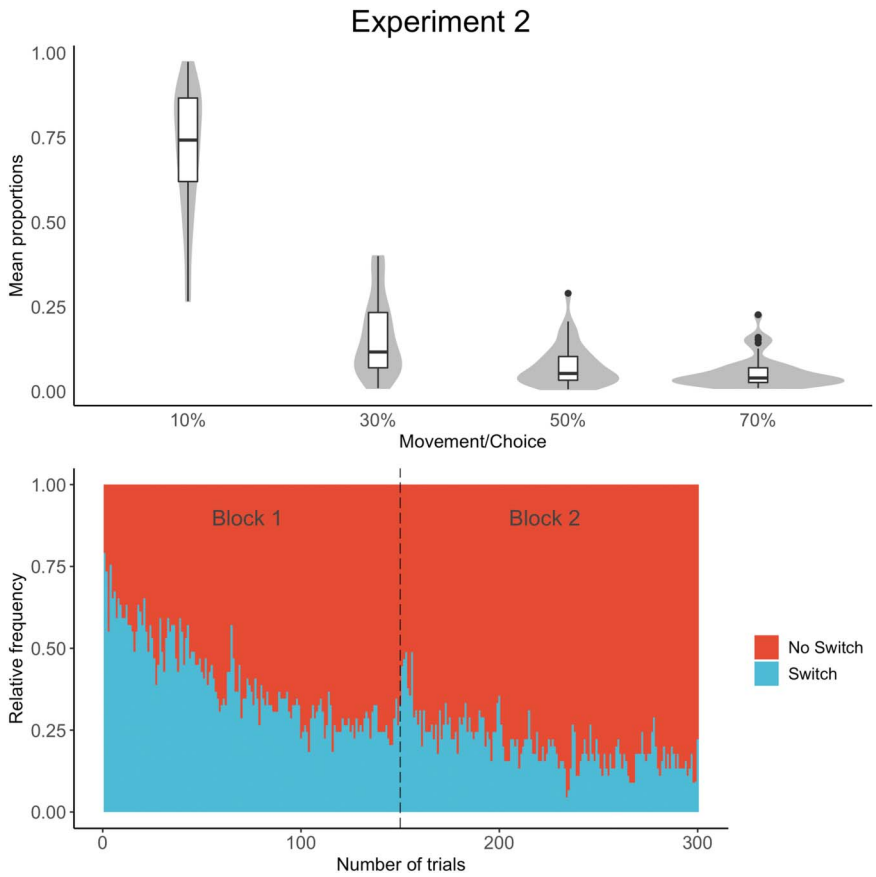

Figure 6. Top panel: Per stimulus means of choosing one of the 4 stimuli. The vertical axis refers to the probabilities of receiving a painful electrocutaneous stimulus after each stimulus for experiment 2. Bottom panel: Proportions of trials, for each block, that participants switched or did not switch their choices for experiment 2.

1.13), unpleasantness (M = 8, SD = 1), and tolerance difficulty (M = 6.89, SD = 0.99). In terms of the rewards, participants evaluated the individually selected rewards as very low in unpleasantness (M = 2.08, SD = 2.29), not difficult to choose (M = 4.41, SD = 2.5), and valuable (M = 7.28, SD = 1.32). Participants were also interested in winning the reward (M = 8.15, SD = 1.53) and willing to make an effort to get the reward (M = 7.83, SD = 1.32). The positive evaluation of the reward also remained during the break between the first and second block (M = 1.74, SD = 2.11).

Regarding the expectancies of an electrocutaneous stimulus, there was no main effect of stimulus, F (2.82, 132.54) = 1.63, P = 0.188, η^2_G = 0.021; BF_{10} = 0.561, block, F (2.82, 132.54) = 1.63, P = 0.188, η^2_G = 0.021; BF_{10} = 0.209, or stimulus × block interaction, F < 1, BF_{01} = 24.832.

For the desire to select a given square or stimulus, there was a main effect of stimulus, F (2.79, 131.13) = 5.70, P = 0.001, η^2_G = 0.087; BF_{10} > 1000, a main effect of block, F (1.00, 47.00) = 10.46, P = 0.002, η^2_G = 0.006; BF_{10} = 0.333, and no stimulus × block interaction, F < 1, BF_{01} = 34.108. After pairwise comparisons, there were significant differences between the 10% and the 30% stimulus, the 50% stimulus, and the 70% stimulus.

Regarding the fear of each square or stimulus, there was only a main effect of stimulus, F (3.00, 141.00) = 3.12, P = 0.028, η^2_G = 0.024; BF_{10} = 4.837, with the other effects failing to reach significance: main effect of block, F < 1; BF_{01} = 8.324, and stimulus × block interaction, F (2.97, 139.59) = 1.41, P = 0.243, η^2_G = 0.004, BF_{01} = 13.973. The only significant pairwise comparison was between the 10% and the 50% stimulus.

Finally, regarding the reward expectancy after selecting a given square or stimulus, there was only a main effect of stimulus, F (2.73, 128.31) = 6.56, P = 0.001, η^2_G = 0.096; BF_{10} > 1000, with the other effects failing to reach significance: main effect of block, F (1.00, 47.00) = 1.40, P = 0.243, η^2_G = 0.001; BF_{01} = 7.193, and stimulus × block interaction, F < 1, BF_{01} = 33.928. The significant pairwise comparisons were between the 10% and the 30% stimulus, the 50% stimulus, and the 70% stimulus.

The above results show that although participants did not learn which square or stimulus was followed by a painful outcome, they nonetheless acquired greater fear regarding movements that had been followed by pain. Moreover, they preferred stimuli with a higher probability of receiving a reward. This result could possibly be explained by the saliency of the reward outcome.

The top panel of **Figure 6** shows the distributions for choosing any one of the 4 squares, while the bottom panel shows the proportion of participants per trial who switched or did not switch. Similar to experiment 1, participants switched choices more often in the first than in the second block: F (1.00, 48.00) = 45.14, P < 0.001, η^2_G = 0.115,BF_{10} > 1000. We include the postpredictive checks in the supplementary material, available at http://links.lww.com/PAIN/B394 (**Fig. 3**).

2.2.2. Performance results

As with the previous study, the winning model is that of *bandit4arm_lapse_decay* (**Table 2**). The posterior distributions

Table 2

Leave-out-one information criterion values per model for experiment 2.

Model	Parameters	LOOIC
bandit4arm_lapse_decay	Reward learning rate, punishment learning rate, reward sensitivity, punishment sensitivity, noise, and decay rate	16,672.28
bandit4arm_4par	Reward learning rate, punishment learning rate, reward sensitivity, and punishment sensitivity	17,112.85
bandit4arm_lapse	Reward learning rate, punishment learning rate, reward sensitivity, punishment sensitivity, and noise	17,134.33
bandit4arm_singleA_lapse	Learning rate, reward sensitivity, punishment sensitivity, and noise	17,434.30
bandit4arm_2par_lapse	Reward sensitivity, punishment sensitivity, choice persistence, and noise	29,080.15

The winning model is the one with the lowest LOOIC value.
LOOIC, leave-out-one information criterion.

of all parameters are shown in **Figure 7**. On the basis of the bottom panel, we see that again the differences between the learning rate and sensitivity parameters were relevant.

2.2.3. Correlation with individual differences

The supplementary Table 57 includes the mean and SDs of all parameter values and all questionnaires for experiment 2 (available at http://links.lww.com/PAIN/B394). Both the NHST and the Bayes tests showed that there was no correlation between the collected individual differences and the parameters of the model (refer to Tables 46-56 and Fig. 4 in the supplementary material for the relevant correlation matrices, available at http://links.lww.com/PAIN/B394).

2.3. Discussion of experiment 2

Experiment 2 replicates and extends the results of experiment 1 by showing that when participants can avoid pain and approach rewards, they tend to exploit more than explore. All parameters are in line with those of experiment 1.

Experiment 3 investigated whether participants are more inclined to stick to exploitation than to exploration, even when the probabilities of receiving the rewarding or painful stimulus change halfway through the experiment. Specifically, we were interested in whether participants continue choosing on the basis of the learned contingencies of receiving a reward or a painful stimulus or whether they adapt their behavior on the basis of the new contingencies. Therefore, we extended the procedure of experiment 2 with 2 new blocks where the contingencies between each stimulus and the probabilities of the outcomes (pain/reward) changed in the middle of the task. If participants do not adapt to the new contingencies, then the choices in the third block would be similar as in the first 2 blocks. In the opposite case, participants' choices should be based on the new contingencies.

3. Experiment 3

3.1. Method of experiment 3

3.1.1. Participants

We recruited 50 individuals who had not participated in any of the previous experiments. One participant was removed because of a mistake during testing (ie, accidental premature termination of the task), reducing the number of participants to 49 (sex = 11 males and 38 females, age: M = 21.16, SD = 2.53). The same exclusion criteria were applied as in the previous experiments. This study was approved by the Ethics Committee of KU Leuven (#G-2018 12 1467). The study's preregistration, all experimental material, and the full data set are available at https://osf.io/pv69j/. Because each session of experiment 3 lasted 1.5 hours, participants were compensated with 12 euros or 1.5 research credits.

3.1.2. Procedure

The procedure was identical to that of experiment 2, with the following differences: Participants completed 4 blocks and the contingencies changed between the first 2 and the last 2 blocks so that different stimuli corresponded to the different contingencies. For example, if in trial blocks 1 and 2 the top right square was associated with a 10% probability of a painful outcome and 90% probability of receiving a reward, in trial blocks 3 and 4 these contingencies applied to another of the 3 squares.

3.1.3. Statistical analyses

We followed the same analytical approach as in experiments 1 and 2, but this time the block factor for all repeated measures ANOVAs was set to 4 (block 1 vs block 2 vs block 3 vs block 4). Regarding the analyses in the model, we ran the same models as before separately for blocks 1 and 2 (the blocks with the original contingencies) and for blocks 3 and 4 (the blocks with the new contingencies). We also computed the differences in parameters between the different blocks by computing the differences between the estimated parameters of blocks 1 and 2 vs blocks 3 and 4.

3.2. Results of experiment 3

3.2.1. Descriptive statistics

Before the beginning of the experimental task, the electrocutaneous stimulus had been judged as painful (M = 7.72, SD = 1.19), unpleasant (M = 7.57, SD = 1.39), and difficult to tolerate (M = 7.19, SD = 1.42). Although there was a significant drop in painfulness and tolerance difficulty—all *P*s < 0.01, BFs > 1000—all ratings remained high: US painfulness (M = 6.95, SD = 1.19), US unpleasantness (M = 7.29, SD = 1.68), and tolerance difficulty (M = 6.38, SD = 1.74). In terms of the rewards,

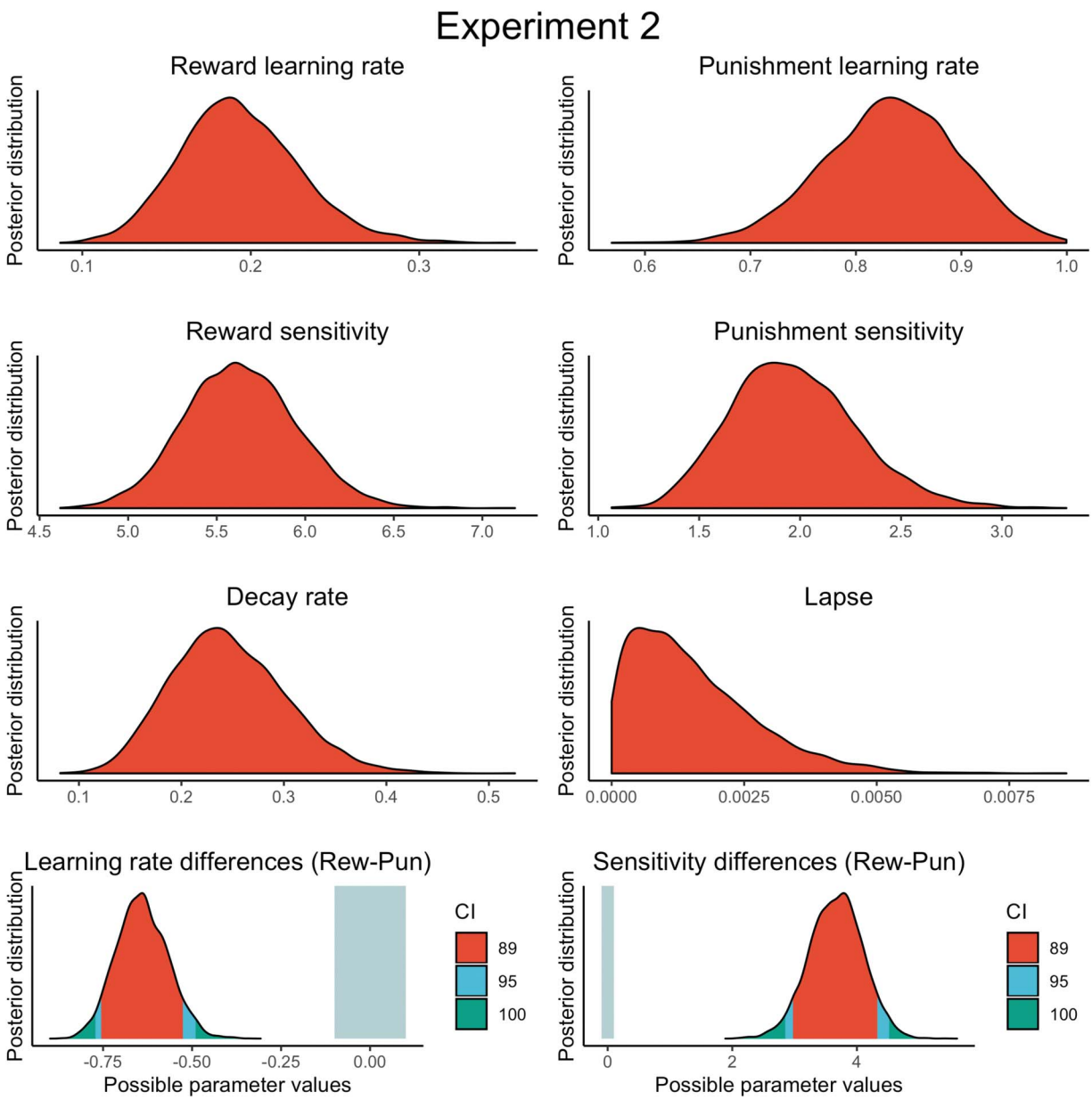

Figure 7. The posterior distributions and ROPEs (bottom 2 panels) for each model parameter for experiment 2. ROPE, region of practical equivalence.

participants evaluated the individually selected rewards as low in unpleasantness (M = 1.43, SD = 0.98), not difficult to choose (M = 4.58, SD = 2.45), and valuable (M = 6.12, SD = 2.77). Participants were also interested in winning the reward (M = 6.99, SD = 2.7) and willing to make an effort to get the reward (M = 6.87, SD = 1.96). The positive evaluation of the reward remained during the reward evaluation in the second block (M = 1.40, SD = 0.86), third block (M = 1.50, SD = 1.10), and fourth block (M = 1.55, SD = 1.74).

Figure 8 depicts the scores of all self-reports across blocks. Regarding the expectancies of a painful outcome, there was a main effect of stimulus, F (2.49, 119.52) = 6.83, $P = 0.001$, $\eta^2_G = 0.043$; $BF_{10} > 1000$, and a main effect of block, F (2.49, 119.52) = 6.83, $P = 0.001$, $\eta^2_G = 0.043$; $BF_{10} = 0.081$, whereas there was no stimulus × block interaction, $F < 1$, $BF_{01} = 616.484$. Pairwise comparisons showed that there were significant differences between the 10% and the 30% stimulus, the 50% stimulus, and the 70% stimulus.

Similar results emerged when participants rated how much they wanted to select a given square or stimulus. There was a main effect of stimulus, F (2.49, 119.52) = 14.60, $P < 0.001$, $\eta^2_G = 0.099$, $BF_{10} > 1000$, and block, F (2.10, 100.80) = 10.73, $P < 0.001$, $\eta^2_G = 0.01$, $BF_{10} = 0.165$, but no stimulus × block interaction, F (4.14, 198.72) = 1.90, $P = 0.109$, $\eta^2_G = 0.02$, $BF_{01} = 7.25$. Pairwise comparisons showed significant differences between the 10% and the 30% stimulus, the 10% and the 50% stimulus, the 10% and the 70% stimulus, the 30% and the 50% stimulus, and the 30% and the 70% stimulus.

Regarding fear of moving towards a given square or stimulus, there was a main effect of stimulus, F (2.43, 116.64) = 5.24, $P = 0.004$, $\eta^2_G = 0.022$; $BF_{10} = 129.113$, and a main effect of block, F (2.43, 116.64) = 5.24, $P = 0.004$, $\eta^2_G = 0.022$; $BF_{10} = 0.099$, whereas there was no stimulus × block interaction, F (4.77, 228.96) = 1.46, $P = 0.208$, $\eta^2_G = 0.013$, $BF_{01} = 28.126$. There was no significant pairwise comparison for the block factor. For the stimulus factor, there were again significant differences

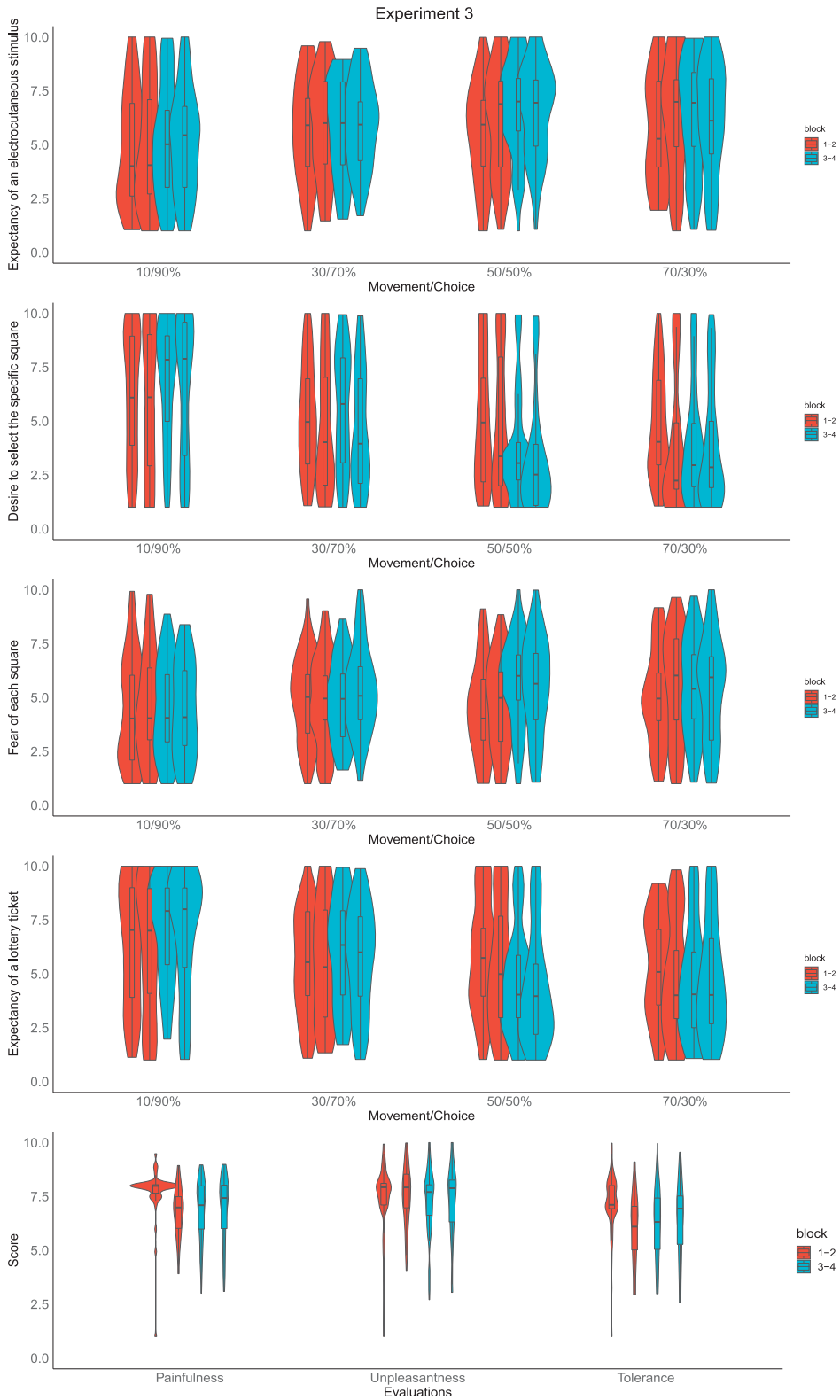

Figure 8. Scores at each questionnaires, filled at the end of each block for experiment 3.

between the 10% and the 30% stimulus, the 10% and the 50% stimulus, and the 10% and the 70% stimulus.

Regarding the reward expectancies for each square or stimulus, there was only a main effect of stimulus, F (2.52, 120.96) = 11.06, $P < 0.001$, $\eta^2_G = 0.078$; $BF_{10} > 1000$, and block, F (2.55, 122.40) = 3.09, $P = 0.037$, $\eta^2_G = 0.003$; $BF_{01} = 70.19$. There was no stimulus × block interaction, F (4.59, 220.32) = 1.87, $P = 0.107$, $\eta^2_G = 0.019$, $BF_{01} = 10.189$. Pairwise comparisons showed that the only significant differences were for the third block between the 10% and the 50% stimulus, the 10%

and the 70% stimulus, the 30% and the 50% stimulus, and the 30% and the 70% stimulus. For the fourth block, there were significant interactions between the 10% and the 50% stimulus and the 10% and the 70% stimulus.

Finally, participants switched more in their choices across blocks: F (1.62, 77.76) = 29.99, $P < 0.001$, $\eta^2_G = 0.159$,$BF_{10} > 1000$. Pairwise comparisons showed significant differences in blocks 2, 3, and 4, with participants more likely than not to repeat the same options. Importantly, as can be seen in **Figure 9**, after the contingencies changed in block 3, participants again started switching their options, signaling more exploratory behavior, before reverting to not switching, a sign of exploitation behavior.

Collectively, these results show that between the different stimuli there were differences between the fear of selecting a given square, the desire to select a given square, and expectancies of receiving a painful or rewarding outcome. These findings suggest that our manipulation was successful and that participants perceived each stimulus differently according to the chances of each stimulus being followed by a painful outcome and/or a reward.

3.2.2. Performance results

On the basis of the results of the model (refer to **Table 3** for blocks 1 and 2 and **Table 4** for blocks 3 and 4), the winning model for both blocks is bandit4arm_lapse_decay. Refer to Figures 3 to 4 in the supplementary material for the postpredictive checks (available at http://links.lww.com/PAIN/B394).

The results replicate the results of experiment 2 for blocks 1 to 2 (**Fig. 10**) and blocks 3 to 4 (**Fig. 11**). Specifically, all results suggest that the differences in learning rate and sensitivity parameters are relevant, with the punishment learning rate being higher than the reward learning rate and reward sensitivity being higher than punishment sensitivity.

We then compared the different parameters between blocks (**Fig. 12**). Although there was no conclusive evidence for differences between the parameters, there were differences in the reward learning rate where there was a difference between blocks. The results reveal a difference between the blocks, with the reward learning being higher in the last 2 blocks (M = 0.35) than in the first 2 learning blocks (M = 0.15). There were no reliable differences for the lapse or the decay parameters, with the values suggesting more exploitation (ie, low randomness in decisions and low forgetfulness after an option has not been chosen, respectively) than exploration.

3.2.3. Correlation with individual differences

The supplementary Table 69 includes the mean and SDs of all parameter values and all questionnaires for experiment 3 (available at http://links.lww.com/PAIN/B394). Similar to the previous experiments, we did not find any evidence for a correlation between the parameters of the model and any of the collected individual differences (refer to Tables 58-68 and Fig. 7 in the supplementary material for the relevant correlation matrices, available at http://links.lww.com/PAIN/B394).

3.3. Discussion of experiment 3

The results of experiment 3 are in line with the results of experiment 2. In addition, the switch in contingencies resulted in individuals updating their behaviors to the new contingencies, with no significant differences between the different blocks in the exploration–exploitation distribution. In addition, there did not seem to be any differences in the parameter values after the change in contingencies, apart from the reward learning rate being higher in the latter portion of the experiment. Collectively, the results show that individuals are quick to learn the new contingencies and stick to exploiting the options that lead to the highest chance of receiving a reward and the lowest chance of receiving a painful outcome.

4. Discussion

We investigated how individuals balance their decisions between exploration and exploitation in an instrumental learning paradigm in which participants were repeatedly required to choose between performing different movements. Each movement was associated with different probabilities of a painful outcome (experiment 1) and/or a reward (experiment 2), with the contingencies between each movement and the presentation of the painful or rewarding outcome changing during the task (experiment 3). Across all experiments, participants tended to exploit more than explore. Positive reinforcement (experiment 2) seemed to lead to faster learning than negative reinforcement (experiment 1). In addition, participants were quick to stop exploiting and start exploring once the contingencies between each movement and the painful or rewarding outcome changed (experiment 3). No individual differences were found between performance and the individual characteristics tested for. Collectively, the results indicate that in the presence of pain, individuals exploit more than explore.

Computational modeling allowed us to break down participants' performance into parameters that reflect decision-making parameters.[24] The learning parameter showed that participants relied on fewer trials when learning which movement was followed by a painful outcome, as compared to those followed by a reward, even in the absence of either the painful outcome or a positive reward. Across all experiments, the lapse and decay parameters were low, suggesting low randomness in participants' choices and good retention of the movement–outcome contingencies.

Our results are in line with motivational accounts of pain[34,49] and corroborate findings from a considerable number of studies in this domain.[5,6,34,36,37,43] According to such accounts, pain motivates actions that stop or cancel the impact of harmful events. Extending previous studies, here we have modeled the dynamic nature of searching for the appropriate action to avoid impending pain. In the laboratory, avoidance is typically studied through instrumental learning procedures, where individuals learn to avoid aversive outcomes by emitting an experimenter-defined response. Although such procedures are valid for studying avoidance,[26] they still fall short in addressing *how* such avoidance arises. In everyday life, individuals enter a cascade of decisions as to which action is appropriate for pain avoidance, often in an unpredictable environment. Exploration–exploitation paradigms seem to be more fitting with regard to the learning processes of avoidance in the context of pain. The exploration–exploitation dilemma typically represents a choice between 2 behavioral options in a dynamic environment. Exploitation refers to the maintenance of the current status, with the risk of missing out on opportunities for improvement. Exploration involves gathering novel information, with the opportunity of improving the current situation but with the risk of worsening it. The current experiments studied avoidance with an experimental task that for the first time combined instrumental conditioning procedures with the *n*-bandit task that is often used to study the exploration–exploitation dilemma. This allowed us to break

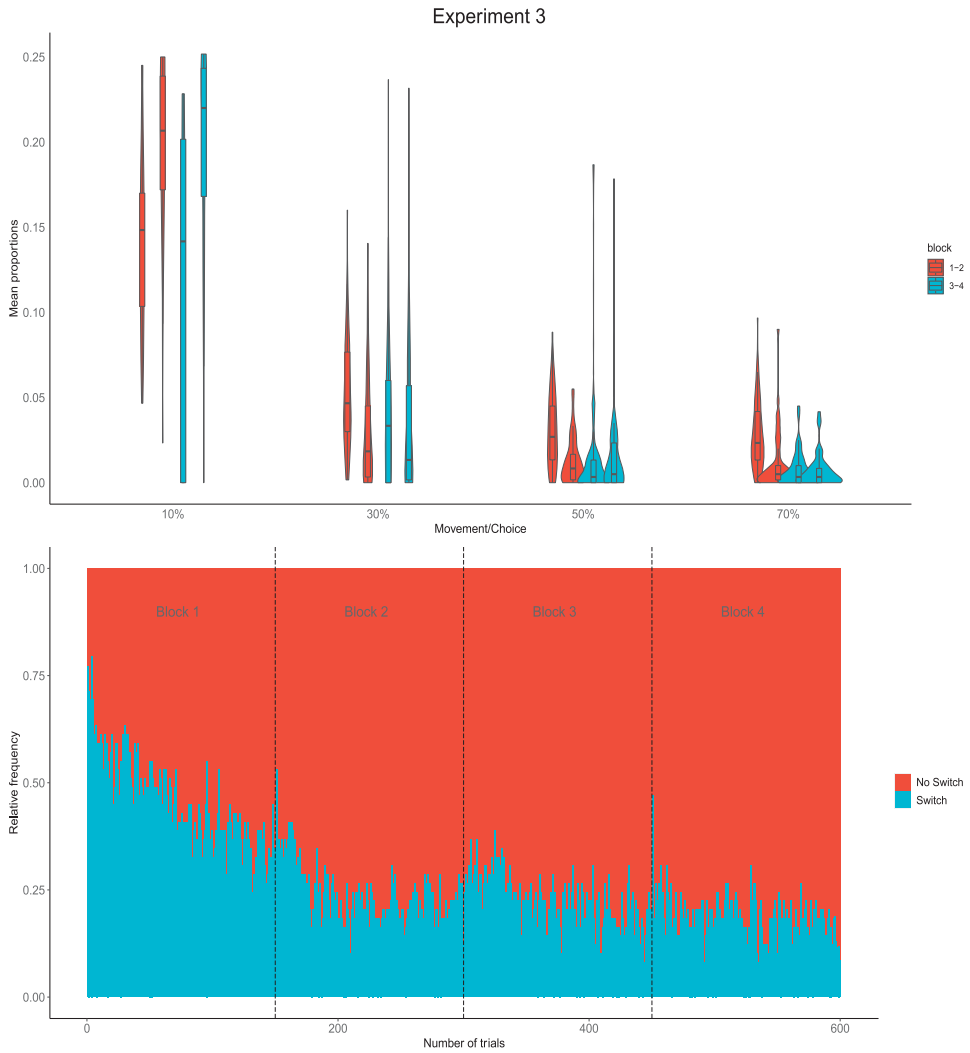

Figure 9. Top panel: Per stimulus means of choosing one of the 4 stimuli. The vertical axis refers to the probabilities of receiving a painful stimulus after each stimulus for experiment 3. Bottom panel: Proportions of trials, for each block, that participants switched or did not switch their choices for experiment 3.

down avoidance performance into different parameters,[24] providing a deeper insight into avoidance decisions than more traditional avoidance paradigms that monitor binominal performance using button pressing (avoid vs not avoid)[28] or that assess gradients of avoidance behavior with a continuous measure.[35]

Our work also extends studies investigating exploration–exploitation decisions in individuals diagnosed with a mental

Table 3

Leave-out-one information criterion values per model for blocks 1 and 2 of experiment 3.

Model	Parameters	LOOIC
bandit4arm_lapse_decay	Reward learning rate, punishment learning rate, reward sensitivity, punishment sensitivity, noise, decay rate	19,953.82
bandit4arm_4par	Reward learning rate punishment learning rate, reward sensitivity, punishment sensitivity	20,098.24
bandit4arm_lapse	Reward learning rate, punishment learning rate, reward sensitivity, punishment sensitivity, noise	20,107.48
bandit4arm_singleA_lapse	Learning rate, reward sensitivity, punishment sensitivity, noise	20,336.10
bandit4arm_2par_lapse	Reward sensitivity, punishment sensitivity, choice persistence, noise	31,246.24

The winning model is the one with the lowest LOOIC value.
LOOIC, leave-out-one information criterion.

Table 4

Leave-out-one information criterion values per model for blocks 3 and 4 of experiment 3.

Model	Parameters	LOOIC
bandit4arm_lapse_decay	Reward learning rate, punishment learning rate, reward sensitivity, punishment sensitivity, noise, decay rate	12,981.98
bandit4arm_lapse	Reward learning rate, punishment learning rate, reward sensitivity, punishment sensitivity, noise	13,255.30
bandit4arm_4par	Reward learning rate punishment learning rate, reward sensitivity, punishment sensitivity	13,266.47
bandit4arm_singleA_lapse	Learning rate, reward sensitivity, punishment sensitivity, noise	13,403.50
bandit4arm_2par_lapse	Reward sensitivity, punishment sensitivity, choice persistence, noise	28,532.37

The winning model is the one with the lowest LOOIC value.
LOOIC, leave-out-one information criterion.

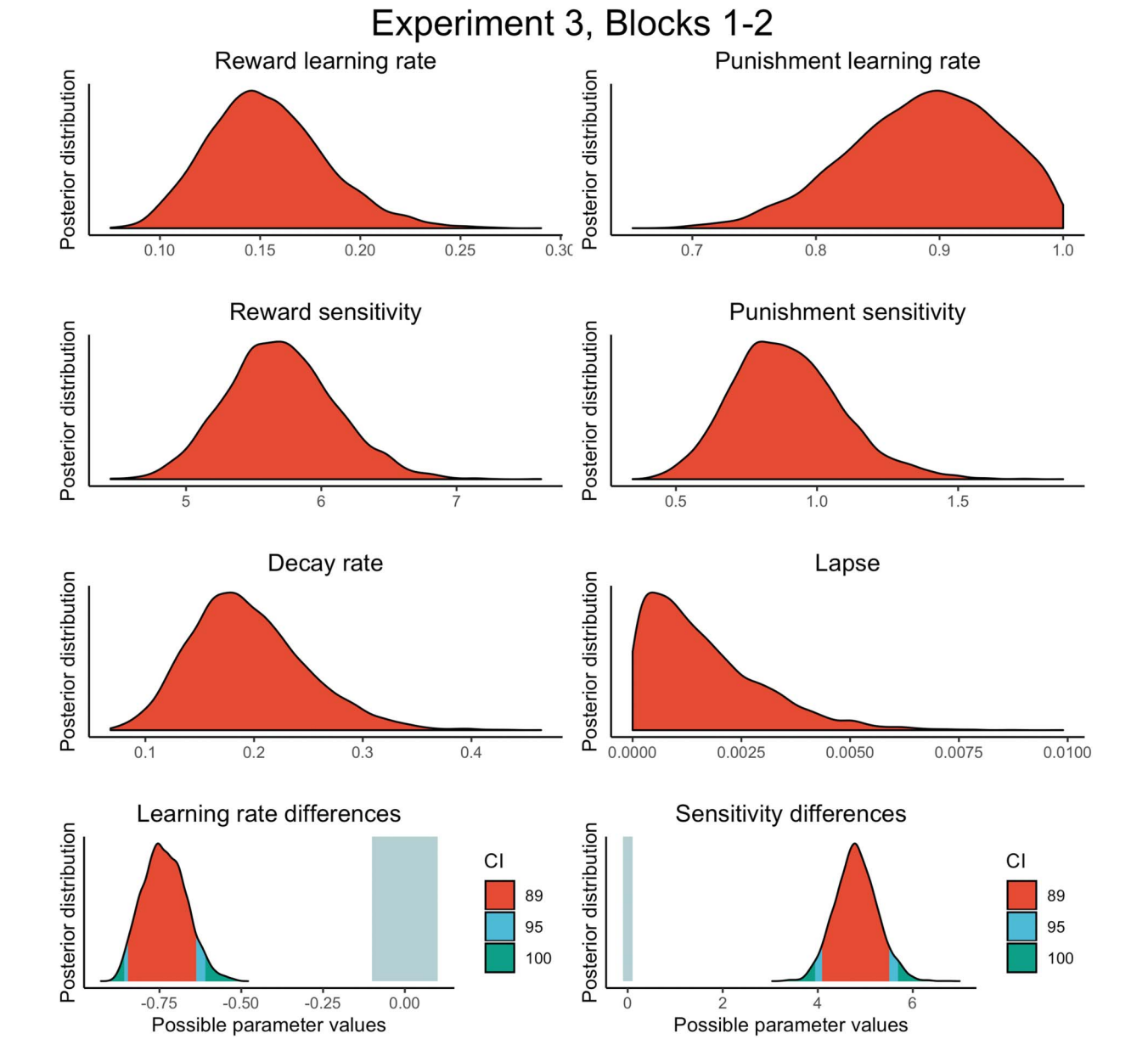

Figure 10. The posterior distributions and ROPEs (bottom 2 panels) for each model parameter for experiment 3 and blocks 1 and 2. ROPE, region of practical equivalence.

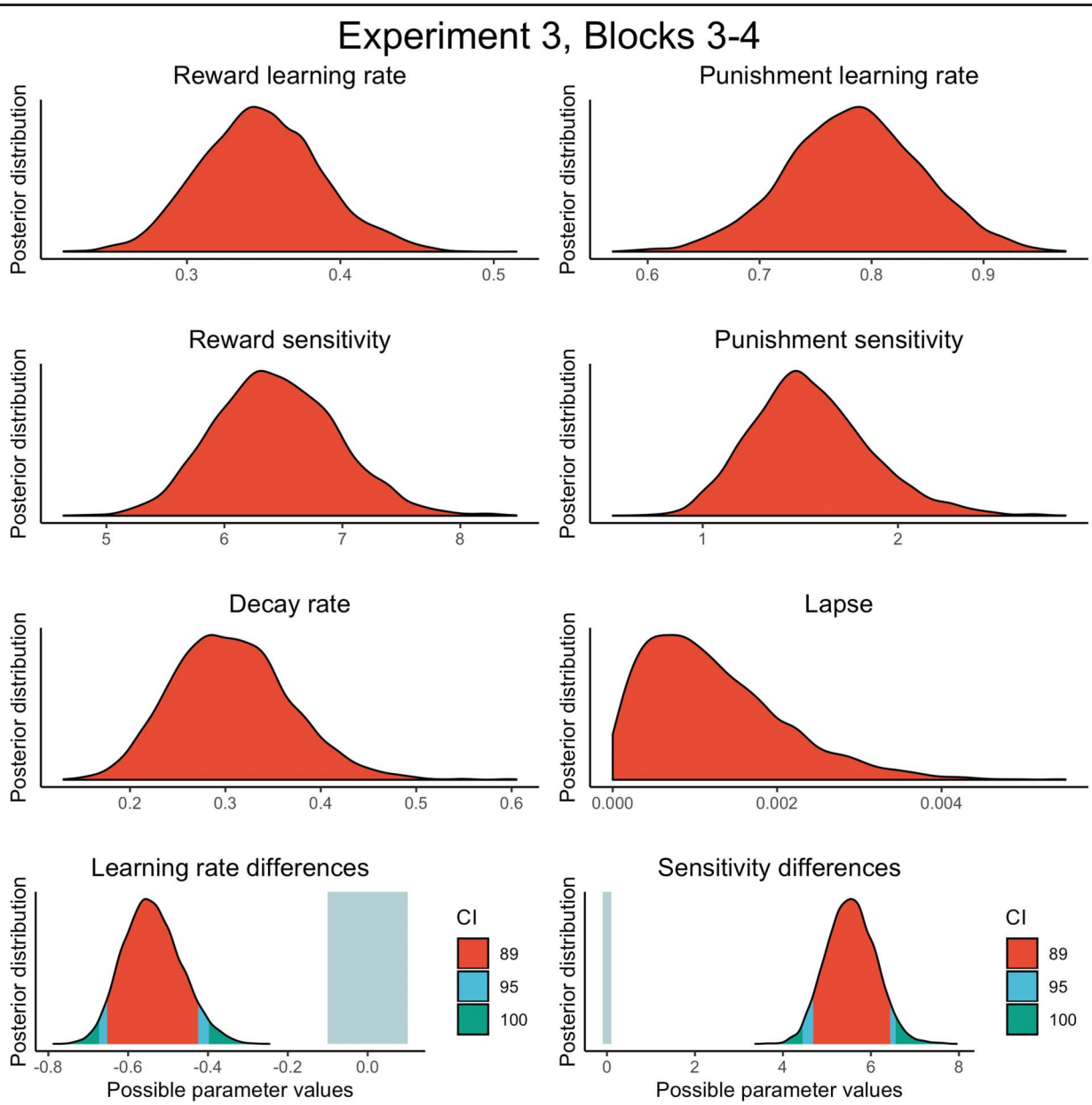

Figure 11. The posterior distributions and ROPEs (bottom 2 panels) for each model parameter for experiment 3 and blocks 3 and 4. ROPE, region of practical equivalence.

disorder. Aylward et al.[2] investigated exploration–exploitation behaviors in individuals with an anxiety-related disorder and in healthy controls. By applying a similar modeling approach, they found that participants with anxiety symptomatology had higher punishment learning rates than controls. The authors suggested that their results implied that one way to reduce anxiety symptomatology might be to assist individuals to integrate information over a longer period. Similarly, Morris et al.[38] tested the exploration–exploitation dilemma in individuals with obesity or drug dependency problems. The results of their study revealed that drug dependency reduced exploration, which often led to choices leading to unfavorable outcomes. The performance of individuals with obesity did not differ from that of healthy controls. Taken together, these results encourage the idea that a better understanding of how individuals solve the exploration–exploitation dilemma may be an important step towards understanding dysfunctional actions.

In this study, none of the individual differences collected predicted any of the performance parameters. There are several possible explanations for this. First, it may be that although the individual differences tested are relevant for avoidance learning, they are not relevant for exploration–exploitation. Second, it is possible that the task used in these experiments is not tailored to capture variations in exploration–exploitation between individuals, at least the variations measured through self-reports.

The results of our study corroborate those of Roy et al.[44] Using a pain-related instrumental task, these authors showed that after receiving a painful outcome, participants switched their choices, which indicates exploratory behavior. Extending these findings, here we used a task with different types of rewards (ie, both positive and negative reinforcement), task switching (ie, a design that allowed us to test the relearning of contingencies between action and reward/punishment), a comparative approach within models (ie, testing different models with different parameters/

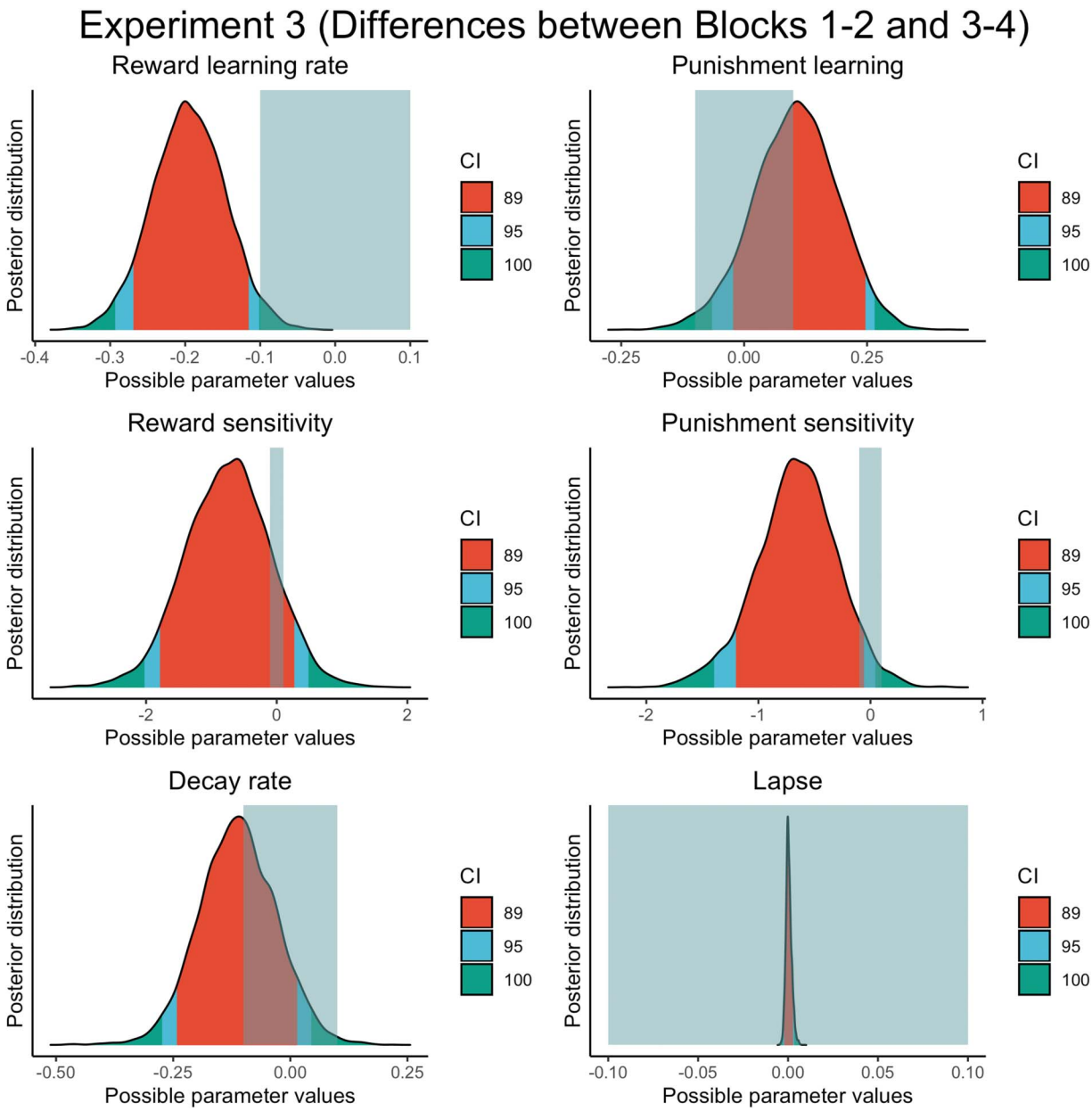

Figure 12. ROPEs for parameter differences between the different blocks. ROPE, region of practical equivalence.

latent constructs), and a more complex task with multiple choice options and pain movements as responses (ie, making our task relevant for acute/chronic pain).

Although here we focused on acute experimental pain, the experimental task used could also form the basis for future studies in individuals with chronic pain. To illustrate, according to the fear-avoidance model of pain,[8,50] individuals suffering from chronic pain may hold catastrophic misinterpretations of pain that can lead to pain-related fear and excessive avoidance. In such cases, even if rewards are available, they may not be sufficient for individuals to overcome their fears (and inhibit their urge to avoid pain) and to focus on performing actions that may result in pain. For example, if someone believes that bending will "break their back," this would seem a greater cost than merely losing an individually tailored reward offered in an experimental set-up. It would be interesting to test how solving the exploration–exploitation dilemma may predict long-term outcomes in these individuals. Even then, the challenge will be to find ways to encourage these individuals to tilt their decisions towards exploration or even towards exploitation of rewards. The prospect of again attaining valuable life goals might be a good candidate in this respect. Research of this kind might reveal to what extent exploration could enhance exposure-based treatments that have generally proved effective in individuals with chronic pain.[10,16,46]

Our study also has limitations. First, in experiments 2 and 3 the probabilities of receiving pain and/or a positive reward were concomitant, meaning that a high probability of receiving pain was associated with a low probability of receiving a positive reward and vice versa. Although this choice was intentional to test positive and negative reinforcement, this is not in line with situations where the chances of pain and reward may be opposite. A second limitation is that in experiment 2, the movement expectancy data did not show that participants

learned which movement was associated with the lowest probability of a painful outcome. Although this finding can be probably explained by random noise (see also the results of experiment 3, where the same manipulation was successful), the rest of the self-report data support the idea that participants learned which movement was associated with the different outcomes. However, the rest of the measures (eg, desire to select each square/stimulus) pointed to successful acquisition of the contingencies. The difference in expectancy ratings between experiments 2 and 3 is an interesting finding in its own right, given that there is wide debate in the literature about whether learning can be achieved with or without contingency awareness.[3,33] Third, we cannot exclude the possibility that not observing individual differences across our sample relates to the homogeneous nature of our sample; repeating the task in a more heterogeneous group (eg, including individuals with chronic pain complaints) is warranted. Finally, we note that our experiment used mostly a relatively healthy student sample, which may potentially limit the extension of our results to other age and education groups.

There are several directions for future studies. One of the main assumptions of the exploration–exploitation dilemma is that behavior is goal-directed and that individuals achieve a balance between these 2 types of behaviors to attain their desired goal. From this perspective, one question is whether and how the balance between exploration and exploitation flexibly changes when a movement that was initially judged to be rewarding stops being rewarding. This could be modeled by devaluing the reward value of the initially rewarding outcome. Specifically, and in line with the literature on habitual and goal-directed behavior,[27] a movement that has been paired reliably with a reward could subsequently be followed by a painful outcome, and vice versa. Persisting with the same learned movement, despite the now aversive outcome, might entail a risk for future behavioral problems because that behavior might no longer serve the goal being pursued. Another question is how the balance between exploitation and exploration is influenced by the presence of learned cues predicting aversive or rewarding outcomes.[52]

Conflict of interest statement

The authors have no conflicts of interest to declare.

Acknowledgements

The authors thank Sandra Seelen, Sophie De Windt, and Alexia Daem for data collection. 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). A.-M. Krypotos is supported by a senior postdoctoral grant from FWO (Reg. #12X5320N) and a replication grant from NWO (Reg. #401.18.056). The authors report no competing interests. Author contributions: A.-M. Krypotos, G. Crombez, N. Claes, and J.W.S. Vlaeyen designed all studies. A.-M. Krypotos supervised the data collection, performed all analyses, and drafted the paper. All authors provided critical comments to the manuscript.

Appendix A. Supplemental digital content

Supplemental digital content associated with this article can be found online at http://links.lww.com/PAIN/B394.

Article history:
Received 31 August 2020
Received in revised form 19 May 2021
Accepted 21 May 2021
Available online 8 June 2021

References

[1] Ahn W-Y, Haines N, Zhang L. Revealing neurocomputational mechanisms of reinforcement learning and decision-making with the hBayesDM package. Comp Psychiatry 2017;1:24–57.
[2] Aylward J, Valton V, Ahn WY, Bond RL, Dayan P, Roiser JP, Robinson OJ. Altered learning under uncertainty in unmedicated mood and anxiety disorders. Nat Hum Behav 2019;3:1116–23.
[3] Baeuchl C, Hoppstädter M, Meyer P, Flor H. Contingency awareness as a prerequisite for differential contextual fear conditioning. Cog Aff Beh Neurosc 2019;19:811–28.
[4] Buhr K, Dugas MJ. The intolerance of uncertainty scale: psychometric properties of the English version. Behav Res Ther 2002;40:931–45.
[5] Claes N, Crombez G, Vlaeyen JW. Pain-avoidance versus reward-seeking: an experimental investigation. PAIN 2015;156:1449–57.
[6] Claes N, Karos K, Meulders A, Crombez G, Vlaeyen JW. Competing goals attenuate avoidance behavior in the context of pain. J Pain 2014;15:1120–9.
[7] Claes N, Vlaeyen JWS, Crombez G. Pain in context: cues predicting a reward decrease fear of movement related pain and avoidance behavior. Behav Res Ther 2016;84:35–44.
[8] Crombez G, Eccleston C, Van Damme S, Vlaeyen JW, Karoly P. Fear-avoidance model of chronic pain: the next generation. Clin J Pain 2002;28:475–83.
[9] Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ. Cortical substrates for exploratory decisions in humans. Nature 2006;441:876–9.
[10] den Hollander M, Goossens M, Jong Jde, Ruijgrok J, Oosterhof J, Onghena P, Vlaeyen JW. Expose or protect? A randomized controlled trial of exposure in vivo vs pain-contingent treatment as usual in patients with complex regional pain syndrome type 1. PAIN 2016;157:2318–29.
[11] Dienes Z. Bayesian versus orthodox statistics: which side are you on? Persp Psycho Sci 2011;6:274–90.
[12] Eysenck HJ, Eysenck SBG. Manual of the Eysenck personality questionnaire. San Diego, CA: Educational; Industrial Testing Service, 1975.
[13] Franken IH, Muris P, Rassin E. Psychometric properties of the Dutch BIS/BAS scales. J Psychopathology Behav Assess 2005;27:25–30.
[14] Garrett N, Daw ND. Biased belief updating and suboptimal choice in foraging decisions. Nat Commun 2020;11:3417–12.
[15] Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. Stat Sci 2002;7:457–72.
[16] Glombiewski JA, Holzapfel S, Riecke J, Vlaeyen JWS, de Jong J, Lemmer G, Rief W. Exposure and CBT for chronic back pain: an RCT on differential efficacy and optimal length of treatment. J Consult Clin Psychol 2018;86:533–45.
[17] Helsen K, Van den Bussche E, Vlaeyen JW, Goubert L. Confirmatory factor analysis of the Dutch intolerance of uncertainty scale: comparison of the full and short version. J Behav Ther Exp Psychiatry 2013;44:21–9.
[18] van den Hout JH, Vlaeyen JW, Houben RM, Soeters AP, Peters ML. The effects of failure feedback and pain-related fear on pain report, pain tolerance, and pain avoidance in chronic low back pain patients. PAIN 2001;92:247–57.
[19] Huys QJ, Moutoussis M, Williams J. Are computational models of any use to psychiatry? Neur Net 2011;24:544–51.
[20] Katehakis MN, Veinott AF Jr. The multi-armed bandit problem: decomposition and computation. Math Oper Res 1987;12:262–8.
[21] Kim SH, Yoon H, Kim H, Hamann S. Individual differences in sensitivity to reward and punishment and neural activity during reward and avoidance learning. Soc Cogn Affect Neurosci 2015;10:1219–27.
[22] Kruschke JK. Bayesian assessment of null values via parameter estimation and model comparison. Persp Psych Sci 2011;6:299–312.
[23] Krypotos AM, Blanken TF, Arnaudova I, Matzke D, Beckers T. A primer on Bayesian analysis for experimental psychopathologists. J Exp Psychopathol 2017;8:140–57.
[24] Krypotos A-M, Crombez G, Meulders A, Claes N, Vlaeyen JWS. Decomposing conditioned avoidance performance with computational models. Behav Res Ther In press.
[25] Krypotos AM, Engelhard IM. Targeting avoidance via compound extinction. Cogn Emot 2019;33:1523–30.

[26] Krypotos AM, Klugkist I, Engelhard IM. Bayesian hypothesis testing for human threat conditioning research: an introduction and the condir R package. Eur J Psychotraumatol 2017;8:1314782.
[27] Krypotos AM, Klugkist I, Mertens G, Engelhard IM. A step-by-step guide on preregistration and effective data sharing for psychopathology research. J Abn Psych 2018;128:517–27.
[28] Krypotos AM, Vervliet B, Engelhard IM. The validity of human avoidance paradigms. Behav Res Ther 2018;111:99–105.
[29] Lee MD, Wagenmakers E-J. Bayesian cognitive modeling: a practical course. Cambridge, United Kingdom: Cambridge University Press, 2014.
[30] LimeSurvey Project Team/Carsten Schmitz. LimeSurvey: an open source survey tool. Hamburg, Germany: LimeSurvey Project, 2012. Available at: http://www.limesurvey.org. Accessed January 2021.
[31] Lommen MJ, Engelhard IM, van den Hout MA. Neuroticism and avoidance of ambiguous stimuli: better safe than sorry? Person Individ Differ 2010;49:1001–6.
[32] Mehlhorn K, Newell BR, Todd PM, Lee MD, Morgan K, Braithwaite VA, Gonzalez C. Unpacking the exploration–exploitation tradeoff: a synthesis of human and animal literatures. Decision 2015;2:191–215.
[33] Mertens G, Engelhard IM. A systematic review and meta-analysis of the evidence for unaware fear conditioning. Neurosci Biobehav Rev 2020; 108:254–68.
[34] Meulders A. From fear of movement-related pain and avoidance to chronic pain disability: a state-of-the-art review. Curr Opin Behav Sci 2019;26:130–6.
[35] Meulders A, Franssen M, Fonteyne R, Vlaeyen JW. Acquisition and extinction of operant pain-related avoidance behavior using a 3 degrees-of-freedom robotic arm. PAIN 2016;157:1094–104.
[36] Meulders A, Vansteenwegen D, Vlaeyen JW. The acquisition of fear of movement-related pain and associative learning: a novel pain-relevant human fear conditioning paradigm. PAIN 2011;152:2460–9.
[37] Meulders A, Vlaeyen JW. The acquisition and generalization of cued and contextual pain-related fear: an experimental study using a voluntary movement paradigm. PAIN 2013;154:272–82.
[38] Morris LS, Baek K, Kundu P, Harrison NA, Frank MJ, Voon V. Biases in the explore–exploit tradeoff in addictions: the role of avoidance of uncertainty. Neuropsychopharmac 2016;41:940–8.
[39] Niv Y, Daniel R, Geana A, Gershman SJ, Leong YC, Radulescu A, Wilson RC. Reinforcement learning in multidimensional environments relies on attention mechanisms. J Neurosci 2015;35:8145–57.
[40] Peirce JW. PsychoPy—psychophysics software in Python. J Neurosc Methods 2007;162:8–13.
[41] R Core Team. R. A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing, 2018.
[42] Richard DM, Jeffrey NR. BayesFactor: Computation of bayes factors for common designs, 2018. Available at: https://CRAN.R-project.org/package=BayesFactor. Accessed January 2021.
[43] Roy M. Weighting pain avoidance and reward seeking: a neuroeconomical approach to pain. J Neurosci 2010;30:4185–6.
[44] Roy M, Shohamy D, Daw N, Jepma M, Wimmer GE, Wager TD. Representation of aversive prediction errors in the human periaqueductal gray. Nat Neurosci 2014;17:1607–12.
[45] Simons LE, Sieberg CB, Carpino E, Logan D, Berde C. The fear of pain questionnaire (FOPQ): assessment of pain-related fear among children and adolescents with chronic pain. J Pain 2011;12: 677–86.
[46] Simons LE, Vlaeyen JW, Declercq L, Smith AM, Beebe J, Hogan M, Kronman C, Mahmud F, Corey J, Sieberg C, Ploski C. Avoid or engage? Outcomes of graded exposure in youth with chronic pain using a sequential replicated single-case randomized design. PAIN 2020;161: 520–31.
[47] Sutton RS, Barto AG. Reinforcement learning: An introduction. Cambridge, MA: MIT press, 2018.
[48] Vehtari A, Gelman A, Gabry J. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. Stat Comp 2017;27:1413–32.
[49] Vlaeyen JW, Crombez G. Behavioral conceptualization and treatment of chronic pain. Ann Rev Clin Psych 2020;16:187–212.
[50] Vlaeyen JW, Crombez G, Linton SJ. The fear-avoidance model of pain. PAIN 2016;157:1588–9.
[51] Wang XJ, Krystal JH. Computational psychiatry. Neuron 2014;84: 638–54.
[52] Watson P, Wiers RW, Hommel B, de Wit S. Motivational sensitivity of outcome-response priming: experimental research and theoretical models. Psychon Bull Rev 2018;25:2069–82.
[53] Wetzels R, Matzke D, Lee MD, Rouder JN, Iverson GJ, Wagenmakers EJ. Statistical evidence in experimental psychology: an empirical comparison using 855 t tests. Perspect Psychol Sci 2011;6:291–8.
[54] van Wijk AJ, Hoogstraten J. Dutch translation of the fear of pain questionnaire: factor structure, reliability and validity. Eur J Pain 2006;10: 479–86.