



# Reward and avoidance learning in the context of aversive environments and possible implications for depressive symptoms

Miriam Sebold<sup>1,2</sup> · M. Garbusow<sup>1</sup> · P. Jetzschmann<sup>1</sup> · D. J. Schad<sup>3</sup> · S. Nebe<sup>4</sup> · F. Schlagenhauf<sup>1,5</sup> · A. Heinz<sup>1</sup> · M. Rapp<sup>2</sup> · N. Romanczuk-Seiferth<sup>1</sup>

Received: 11 January 2019 / Accepted: 5 June 2019  
© The Author(s) 2019

## Abstract

**Background** Aversive stimuli in the environment influence human actions. This includes valence-dependent influences on action selection, e.g., increased avoidance but decreased approach behavior. However, it is yet unclear how aversive stimuli interact with complex learning and decision-making in the reward and avoidance domain. Moreover, the underlying computational mechanisms of these decision-making biases are unknown.

**Methods** To elucidate these mechanisms, 54 healthy young male subjects performed a two-step sequential decision-making task, which allows to computationally model different aspects of learning, e.g., model-free, habitual, and model-based, goal-directed learning. We used a within-subject design, crossing task valence (reward vs. punishment learning) with emotional context (aversive vs. neutral background stimuli). We analyzed choice data, applied a computational model, and performed simulations.

**Results** Whereas model-based learning was not affected, aversive stimuli interacted with model-free learning in a way that depended on task valence. Thus, aversive stimuli increased model-free avoidance learning but decreased model-free reward learning. The computational model confirmed this effect: the parameter lambda that indicates the influence of reward prediction errors on decision values was increased in the punishment condition but decreased in the reward condition when aversive stimuli were present. Further, by using the inferred computational parameters to simulate choice data, our effects were captured. Exploratory analyses revealed that the observed biases were associated with subclinical depressive symptoms.

**Conclusion** Our data show that aversive environmental stimuli affect complex learning and decision-making, which depends on task valence. Further, we provide a model of the underlying computations of this affective modulation. Finally, our finding of increased decision-making biases in subjects reporting subclinical depressive symptoms matches recent reports of amplified Pavlovian influences on action selection in depression and suggests a potential vulnerability factor for mood disorders. We discuss our findings in the light of the involvement of the neuromodulators serotonin and dopamine.

**Keywords** Reward learning · Avoidance learning · Reinforcement learning · Computational psychiatry · Decision-making · Affective modulation · Depression symptoms

---

This article belongs to a Special Issue on Translational Computational Psychopharmacology.

---

**Electronic supplementary material** The online version of this article (<https://doi.org/10.1007/s00213-019-05299-9>) contains supplementary material, which is available to authorized users.

---

✉ Miriam Sebold  
miriam.sebold@charite.de

<sup>1</sup> Department of Psychiatry and Psychotherapy, Charité – Universitätsmedizin Berlin, Charitéplatz 1, 10117 Berlin, Germany

<sup>2</sup> Department for Social and Preventive Medicine, University of Potsdam, Potsdam, Germany

<sup>3</sup> Cognitive Science, University of Potsdam, Potsdam, Germany

<sup>4</sup> Zurich Center for Neuroeconomics, Department of Economics, University of Zurich, Zurich, Switzerland

<sup>5</sup> Max Planck Institute for Human Cognitive and Brain Sciences, 04303 Leipzig, Germany

## Introduction

In animals and humans, the appearance of a predator or other dangers typically promote withdrawal. These stimulus-induced avoidance responses are hard-wired and can be very beneficial in the presence of a mortal threat. Conceptually, holding back from threat reflects a Pavlovian response because a stimulus automatically elicits an action that is independent from the organism's behavior. Pavlovian responses are disentangled from operant responses, where actions are elicited because they have repeatedly been paired with a specific consequence (reward or punishment).

Interestingly, interactions between Pavlovian avoidance responses and operant actions have been reported. Examples for these interactions are evident in real-life situations, as negative facial expressions reduce consumption behavior (Winkielman et al. 2005) and bad weather on election days reduces voter turnouts (Bassi 2013).

Experimentally, two types of such interactions have been reported: first, a stimulus-induced avoidance response can exaggerate another avoidance action. Consequently, negative stimuli increase an acquired withdrawal behavior, a phenomenon entitled as aversive Pavlovian-to-instrumental transfer (PIT) effect (Geurts et al. 2013b; Lewis et al. 2013; Nord et al. 2018). Second, stimulus-induced avoidance responses can interfere with approach behavior. Accordingly, negative stimuli decrease appetitive responses (Lee et al. 2005; Nelson and Sanjuan 2006), a phenomenon termed as conditioned suppression. Both of these phenomena provide evidence that negative valence can interact with actions. Further substantiation for these interactions come from studies showing impaired go but facilitated no-go responses when subjects anticipate punishment (Crockett et al. 2012a; Crockett et al. 2009; Guitart-Masip et al. 2012; Guitart-Masip et al. 2011).

Much attention has been drawn to phenomena where aversive stimuli interact with operant actions, mostly because such interactions also play a role in psychopathologies (Dayan et al. 2006). In mood disorders, for instance, patients are biased towards negative environmental stimuli and these biases amplify avoidance behavior (Eshel and Roiser 2010). Indeed, in depressive patients, aversive cues exert increased control over avoidance behavior (Nord et al. 2018).

So far, previous studies investigating valence-action interactions have realized actions along an axis of activation, thus as approach/avoidance or go/no-go responses. No study has so far investigated how aversive stimuli interacted with complex learning and decision-making, where subjects have to choose between several options in order to optimize their decisions. The analogy between complex decision-making and activation is at hand, as the former can also be mapped on a valence axis, ranging from avoidance- (minimize punishment) to reward- (maximize reward) based decision-making.

Importantly, recent computational accounts of complex learning and decision-making have emphasized the existence of two distinct control mechanisms, where actions can be guided by either model-free or model-based computations (Daw et al. 2011). Model-free controllers are reflexive and habitual and involve the retrospective updating of action values. Model-based controllers on the other side are adaptive and goal-directed and choose actions based on an internal model.

With regard to valence-action interactions, the concept of model-free control is specifically interesting because it shares many features of Pavlovian control including automaticity and rigidity (Friedel et al. 2014; Gillan et al. 2015). Moreover, model-free and Pavlovian control processes are assumed to rely on very similar neurobiological and computational mechanisms (Dayan and Berridge 2014).

Interestingly, certain internalizing pathologies such as depression have been associated with the dominance of the habitual system at the expense of the goal-directed system (Huys et al. 2015a; Huys et al. 2015b). In line with this, negative affective states such as acute stress shift decision-making away from model-based control, especially in subjects with depressive symptoms (Heller et al. 2018) or previous adverse life experiences (Radenbach et al. 2015).

The overarching aim of the present study was to elucidate the computational mechanisms of how aversive stimuli interacted with learning and decision-making. Given previous valence-action interactions (Guitart-Masip et al. 2014), we expected opposing roles on reward and avoidance learning. More precisely, we hypothesized impairments in reward learning, but enhancements in avoidance learning in the presence of negative stimuli. With regard to the computational mechanism of modulation, we asked whether model-free or model-based controllers would be affected by aversive stimuli.

As previous research has shown increased modulation of avoidance responses in the presence of aversive stimuli (Nord et al. 2018) as well as stress-induced disruptions in model-based control in depression (Heller et al. 2018), we further asked whether interactions between aversive influences on reward and avoidance learning were moderated by subclinical depression scores.

## Methods

### Subjects

Subjects were  $n = 55$  21-year-old males, who had previously taken part in a larger study investigating learning in alcohol dependence (Lead study, [ClinicalTrials.gov](https://clinicaltrials.gov/ct2/show/study/NCT01744834) identifier: NCT01744834). All subjects were free from psychotropic medication and had no axis-1 psychiatric disorder as indicated by the Composite International Diagnostic Interview (CIDI,

(Jacobi et al. 2013; Wittchen and Pfister 1997), see S1 for full inclusion criteria). In the Lead study, the 2-step task in its original form (no background stimuli, reward condition only) had previously been administered twice (Nebe et al. 2018). Therefore, subjects were familiar with the general task structure. Due to technical problems, the experiment was aborted in one subject, leaving a final sample of  $n = 54$ . Subjects were compensated for their participation with a fixed amount of 20 € plus an additional sum contingent on task performance (max. 15 €). Participants' demographic and clinical characteristics are outlined in Table S1.

## Task

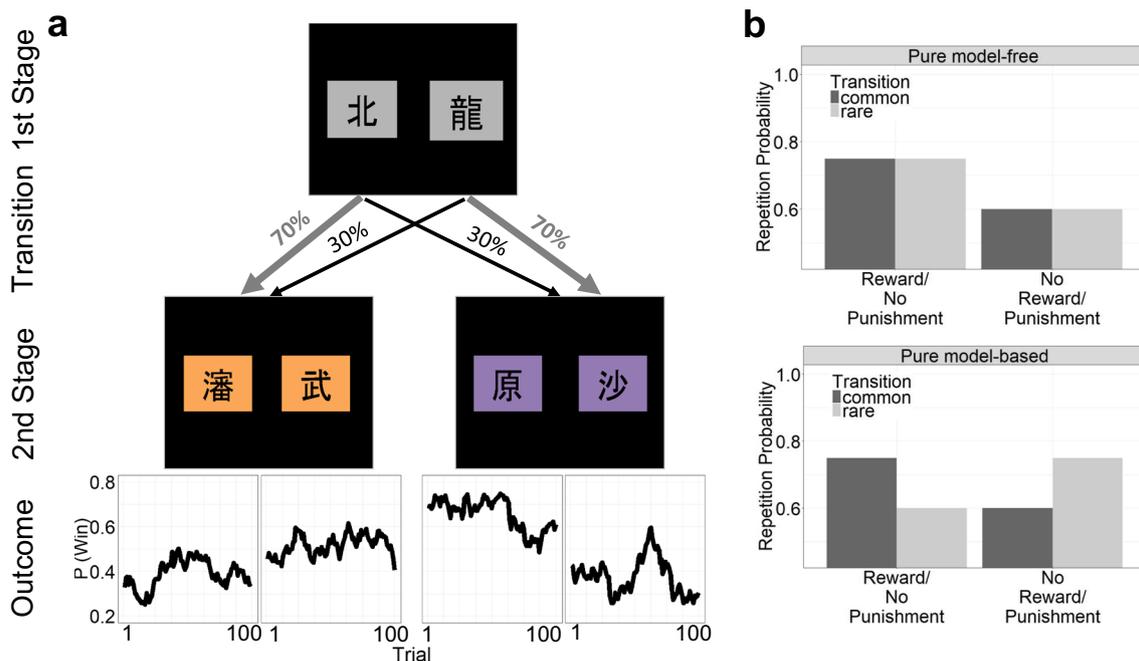
We used a two-stage Markov decision task with separate reward and punishment conditions as previously described (Worbe et al. 2016). Briefly, on each trial in the first stage, subjects made a choice between two stimuli, which led to one of two pairs of second-stage stimuli with fixed probabilities (70 and 30% of choices). Each of the four second-stage stimuli was associated with probabilistic 20 Euro Cent monetary reward in the reward condition and loss in the punishment condition or no monetary outcome (Fig. 1a).

Model-free and model-based learning make distinct predictions about first-stage choice repetition probabilities (Fig. 1b). A model-free strategy predicts a higher repetition probability of first-stage choices that led to reward in the previous trial,

therefore predicting a main effect of outcome on first-stage repetitions. However, a model-based strategy predicts the incorporation of the transition frequency of the previous trial and therefore the interaction between outcome and transition on first-stage repetitions. For instance, it predicts lower repetition probabilities when the transition frequency of the previous trial was rare but rewarded, because the unchosen first-stage stimulus has a higher probability in leading to the rewarding second-stage stimulus pair. The same is true for the other task valence when choices lead to no punishment.

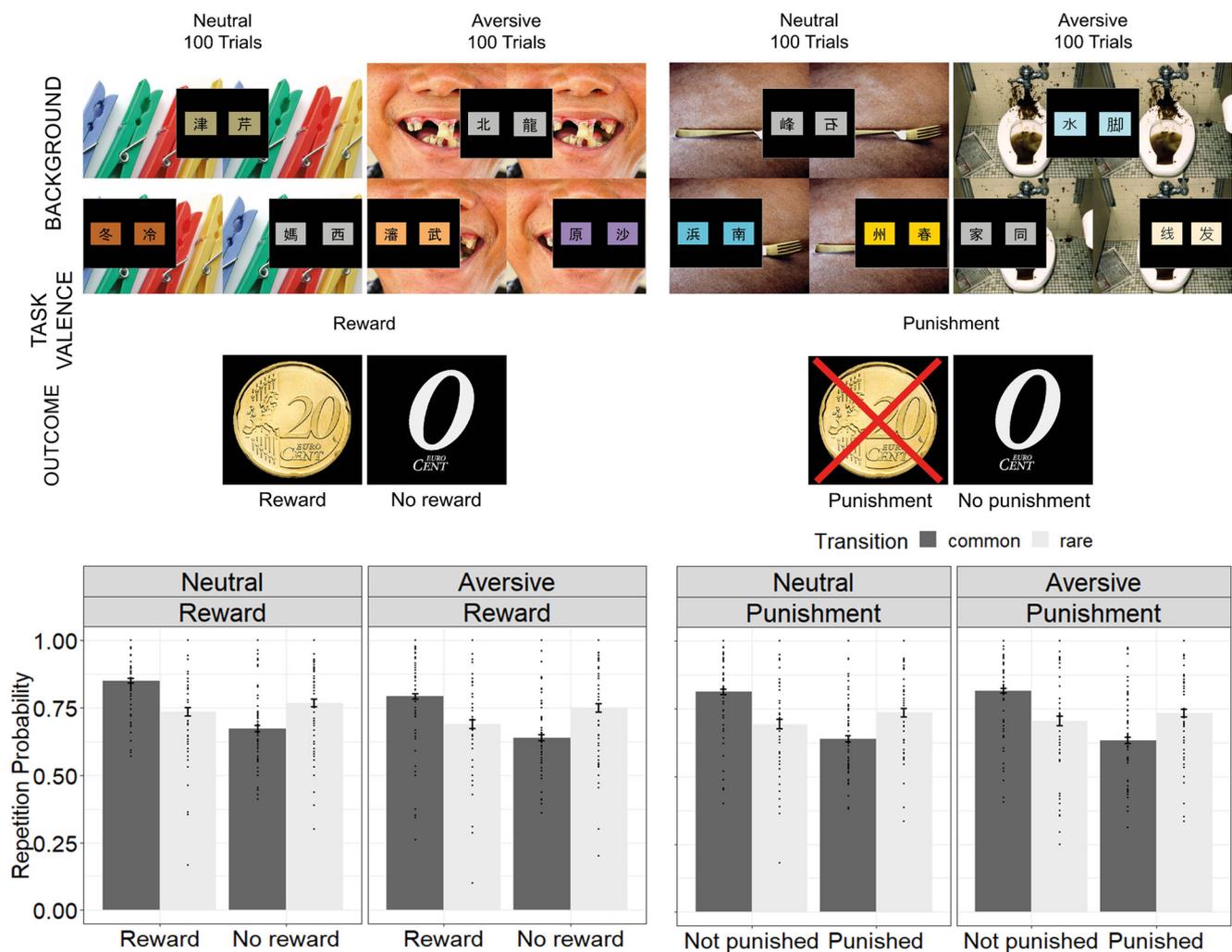
Reward and punishment conditions were either deposited with neutral or aversive background stimuli (Fig. 2), resulting in a 2 (background)  $\times$  2 (task valence) within-subject design (neutral reward, aversive reward, neutral punishment, aversive punishment). Each background stimulus was shown four times to fill the entire screen, as previously done in human PIT paradigms (Garbusow et al. 2016; Garbusow et al. 2014). The background stimulus appeared at the beginning of each trial and stayed there during the complete two-step trial (first, second stage, and outcome).

All four conditions of the task had the same transition probabilities and dynamic range of the reward or punishment probability. The courses of these second-stage outcome probabilities (random walks) were slightly different between all four conditions but were randomly assigned to each condition. All conditions had different color code and stimuli set on the first and second task stages. The used sets were identical to Liu et al.



**Fig. 1** **a** Trial configuration: in each trial, subjects had to make two consecutive choices. At a first stage, subjects chose one stimulus over the other and then proceeded to a second stage where they chose between two stimuli. Second-stage choices were probabilistically rewarded or punished according to Gaussian random walks. Transition probabilities

between first and second stages varied for first-stage choices: whereas one stimulus choice led commonly (70% of all trials) to one second-stage stimulus pair and rarely (30% of all trials) to the second-stage stimulus pair, the opposite was true for the other first-stage choice. **b** Expected stay probabilities of pure model-free and pure model-based learning



**Fig. 2** **a** Schematic presentation of the task. Subjects performed 100 trials within each condition (neutral reward, aversive reward, neutral punishment, aversive punishment). Stimulus sets and courses of second-stage reward probabilities (random walks) varied between all conditions. First and second stages were presented consecutively and are

presented simultaneously here for illustrative purposes of the different stimulus sets. **b** Overall performance: across all conditions, subjects showed a mixture of model-free and model-based behavior (main effect outcome and transition  $\times$  outcome interaction)

(2016). Participants completed 100 trials for each condition divided into two sessions. The order of the reward and punishment conditions as well as the order of the aversive or neutral background conditions was counterbalanced across subjects.

Neutral and aversive background stimuli were drawn from the international affective picture system (IAPS, Lang et al. 2005). Picture selection was based on the ratings of the male college cohort as reported in Lang et al. (2008). To exclude pictures in the aversive condition with potential disruptive or harmful content, we limited our selection to pictures with ratings of moderate arousal. Our final aversive image selection included pictures with lowest ratings in valence and lower arousal ratings than the mean + 0.2 standard deviations. Our final neutral image selection included pictures with lowest arousal ratings and valence ratings between mean valence of the cohort and the mean + 0.2 standard deviations.

Based on visual inspection, we excluded pictures with black borders, sexual content, and duplicates. Each condition consisted of a final set of 50 pictures (see Table S2). Each picture was shown twice, once in the reward and once in the punishment condition. Examples of stimuli in the neutral set are household articles such as a tissue, a basket, or a lamp whereas the aversive set includes images of contaminated toilets, sick or dead animals, and violence. Post hoc comparisons of the ratings by Lang et al. (2008) of our final stimulus set (using Wilcoxon rank sum test) indicated that pictures of the aversive condition were rated as more negative ( $p < 0.0001$ ) and more arousing ( $p < 0.0001$ ) than neutral pictures.

Before the experiment, all subjects underwent the self-paced computer-based instructions explaining the structure

of the task and providing examples. Overall, the subjects were instructed to win as much money as they could in the reward condition and to avoid monetary loss in the punishment condition of the task. Participants were told that they would be paid for the experiment depending on their cumulative performance in the reward and the punishment condition. After an extensive training session, subjects were confronted with 3 multiple choice questions on the structure and the reward versus punishment condition of the task (adapted from Gillan et al. (2016)). Subjects only advanced to the experiment if they answered all questions correctly.

## Questionnaire

In a separate session which was conducted prior to the experimental session, subjects answered (among other questionnaires) the German version of the Hospital anxiety and depression scale (HADS, Herrmann-Lingen 2005; Zigmond and Snaith 1983), which includes 14 items, of which 7 items are related to anxiety symptoms and 7 items are related to depressive symptoms. Subjects can respond to each item on a scale between 0 (do not agree at all) and 3 (fully agree). Scores are summed up and can vary between 0 and 21 for both subscales. Previous studies have found that depressive symptoms increase the impact of negative states on decision-making (Heller et al. 2018). Moreover, negative Pavlovian stimuli decreased approach but increased avoidance responses in depressive patients, but not in matched healthy controls (Nord et al. 2018). Based on these findings, we explored the association between subclinical depressive symptoms and task behavior on an explorative basis.

## Analyses

We performed two sets of analyses. The first was a mixed effects logistic regression where first-stage choices (stay/switch) were regressed on the previous trial outcome (reward/no reward or no punishment/punishment), transition frequency (common/rare), the background (aversive/neutral), and task valence (reward/punishment). Within-subject factors (intercept, main effect of outcome, main effect of transition, main effect of background, task valence and their interaction) were taken as random effects across participants. For visual purposes, we extracted model-free and model-based scores as previously described (Sebold et al. 2014). Briefly, model-free scores reflected the individual main effect of outcome on first-stage repetition probabilities (% reward/not punished common + % rewarded/not punished rare – % unrewarded/punished common – % unrewarded/punished rare), whereas model-based scores reflected the interaction between transition frequency and outcome on first-stage repetition probabilities (% reward/not punished common + % unrewarded/punished rare – % rewarded rare/not punished – % unrewarded/punished

common). As we were specifically interested in how aversive stimuli alter these scores depending on task valence, we calculated these scores between aversive and neutral conditions for the reward and the punishment condition respectively (separate for model-free and model-based control).

The second analysis was the fit of the original Daw et al. (2011) reinforcement learning model (7-parameter hybrid model, see S4) to the data. We also fitted two reduced models for model comparison and selection (see S5). We used an expectation maximization algorithm to find maximum a posteriori estimates. We fitted the model simultaneously for all four conditions and compared parameters of interest. Practically, the model consists of two different sets of parameters: the reinforcement learning parameters that capture the internal learning and evaluation processes and the response (softmax) parameters that associate the result of the internal valuations to choices. We hypothesized that background and task valence would specifically influence learning parameters but had no hypothesis on how it would affect softmax parameters. The learning parameters in the hybrid model included two learning rate parameters ( $\alpha_1$  and  $\alpha_2$  for first and second stages, respectively); the weighting parameter  $\omega$ , which indicates the balance between model-free and model-based control; and the eligibility trace parameter  $\lambda$  from the model-free algorithm, which indicates how much second-stage outcomes update first-stage action values.

These four reinforcement parameters ( $\alpha_1$ ,  $\alpha_2$ ,  $\omega$ ,  $\lambda$ ; for their distribution, see Table S3) were subjected to a multivariate  $2 \times 2$  ANOVA analysis with background and task valence as within-subject effects and tested for interactions. Post hoc tests included univariate  $2 \times 2$  ANOVAs with background and task valence as repeated measure factors for all four parameters.

A key scientific question of reinforcement learning concerns how accurate RL fits actually are. Indeed, using rather complex algorithms can lead to ambiguous or false parameter identifications that mischaracterize participants actual learning strategies. One way to evaluate this is to generate simulation data from the computational model with the parameters that have been estimated from the real data. In an ideal world, the simulation data should faithfully match the real data (Gureckis and Love 2009). Thus, we ran additional simulation analyses, where individual first- and second-stage actions were generated from the computational model with the previously estimated parameters. For each subject, we computed 100 simulation data sets for each condition (neutral reward, aversive reward, neutral punishment, aversive punishment). Consequently, we compared model-free (main effect outcome) and model-based (interaction between outcome and transition) scores between the real data and the simulated data. Moreover, we tested for interaction effects between task valence, outcome, background, and transition in the simulated data.

Regression analyses were conducted using generalized linear mixed-effects models implemented with the lme4 package (Bates et al. 2015) in the R programming language, version 3.1.2 (cran.us.r-project.org). Computational modeling was performed in Matlab 2014 (8.3., 2014a). MANOVA for modeling parameters was computed using the RM function from the MANOVA.RM package (Friedrich et al. 2017) in R.

## Results

### Behavioral data

Subjects missed comparably few trials (mean = 6.5, sd = 7.4 trials). Twenty-six subjects began with the punishment condition, and 28 subjects began with the reward condition. Moreover, 20 subjects began with aversive and 34 began with the neutral background stimuli, suggesting that the randomization was effective.

The logistic regression indicated a main effect of outcome ( $p < 0.0001$ ), a main effect of transition frequency ( $p < 0.05$ ), and an interaction between outcome and transition frequency ( $p < 0.0001$ ). Thus, subjects showed a mixture between model-free and model-based learning (Fig. 2b), as previously shown in this cohort (Nebe et al. 2018).

Neither task valence ( $p = 0.1$ ) nor background stimuli per se ( $p = 0.2$ ) influenced choice behavior. However, we found an interaction between task valence and background ( $p = 0.011$ ), suggesting that aversive stimuli influenced first-stage repetition behavior depending on task valence (Fig. 3a, right panel). Post hoc analyses indicated that aversive stimuli decreased repetition in the reward condition ( $p = 0.002$ ), but not in the punishment condition ( $p = 0.41$ ).

Importantly, the interaction between background and task valence was further modulated by outcome, as indicated by the three-way interaction of outcome  $\times$  task valence  $\times$  background ( $p = 0.041$ ). This result indicates that aversive stimuli affect model-free learning in a way that depends on task valence, justifying further post hoc analyses. These post hoc analyses indicated an interaction between outcome  $\times$  task valence in the presence of aversive background stimuli ( $p = 0.04$ ), but not neutral background stimuli ( $p = 0.44$ ), indicating that aversive stimuli promoted model-free learning in the punishment condition, but decreased model-free learning in the reward condition (Fig. 3a, left panel). The interaction between transition frequency and outcome was not modulated by task valence ( $p = 0.6$ ), background ( $p = 0.9$ ), or the interaction between task valence and background ( $p = 0.9$ ) indicating that model-based performance was not modulated by either of the two manipulations (Fig. 3a, middle panel). Complete results of the regression analysis can be found in Table 1 (real data).

### Computational modeling

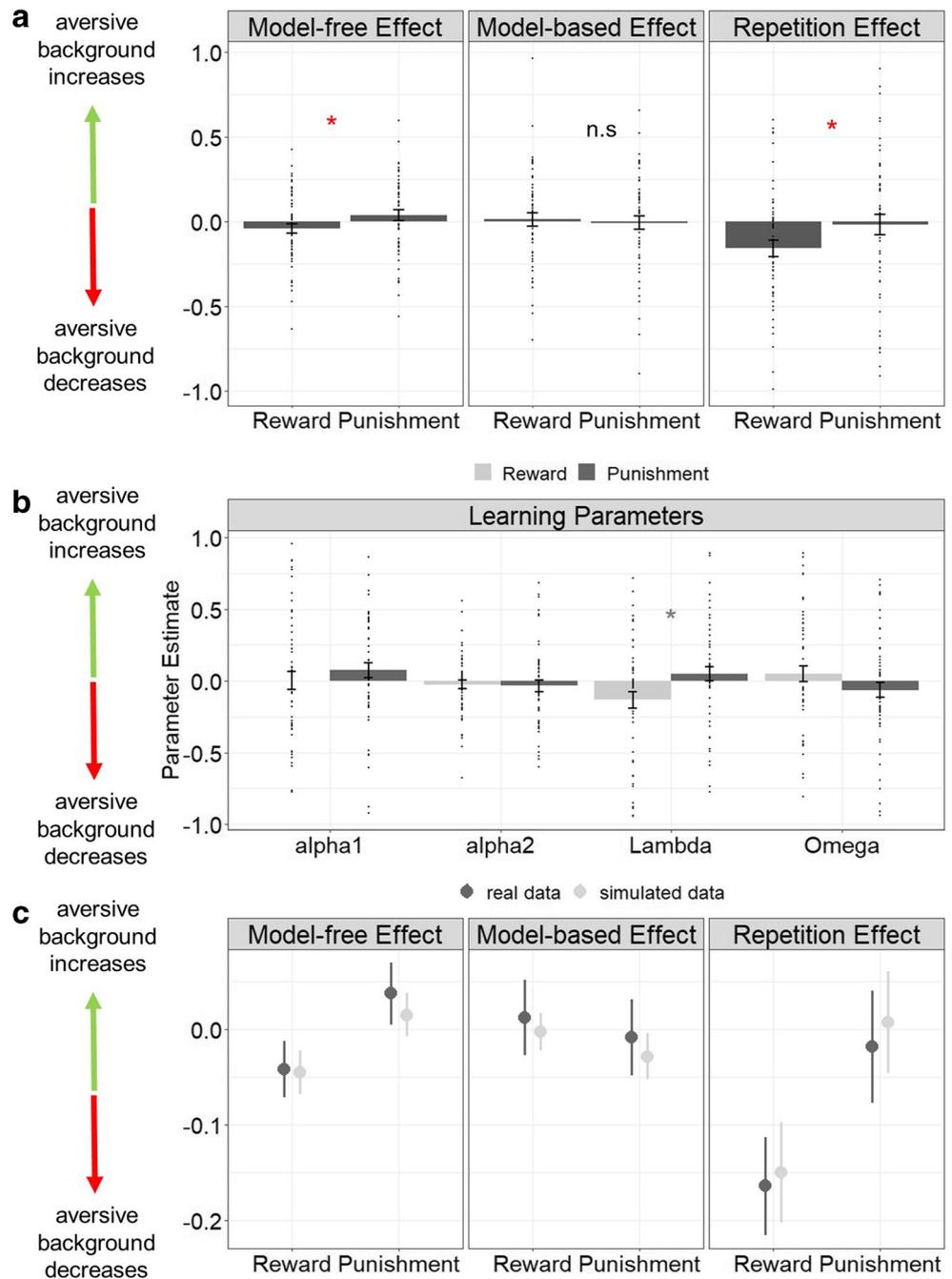
The MANOVA indicated a significant interaction between background and task valence (ANOVA-type statistic (ATS) = 2.78,  $p = 0.042$ ). Post hoc univariate tests showed no difference for the balance of model-free and model-based control  $\omega$  ( $F_{(1,159)} = 2.27$ ,  $p = 0.14$ ), for first-stage learning rates  $\alpha_1$  ( $F_{(1,159)} = 0.84$ ,  $p = 0.36$ ) nor for second-stage learning rates  $\alpha_2$  ( $F_{(1,159)} = 0.03$ ,  $p = 0.86$ ), but significant differences in stage-skipping update  $\lambda$  ( $F_{(1,159)} = 4.73$ ,  $p = 0.03$ , Fig. 3b). The parameter  $\lambda$  signifies a stronger influence of reward prediction errors at the second stage on first-stage decision values and accounts for the main effect of outcome observed in first-stage stay behavior. In line with raw data analysis, this speaks for a subtle, albeit significant, elevation of model-free learning in the presence of aversive stimuli in the punishment condition, but impairments in the reward condition. Explorative comparison of parameters of the softmax decision model ( $\beta_1$ ,  $\beta_2$ ,  $\rho$ ) showed no significant modulation by task valence or background (ANOVA-type statistic (ATS) = 1.39,  $p = 0.249$ , see Fig. S1).

Simulation analyses confirmed the observed effect, as, overall, subjects showed decreased model-free learning in the aversive environment in the reward condition, but increased model-free learning in the aversive environment in the punishment condition (Fig. 3C, left panel). Moreover, in line with the real data, model-based learning was not influenced by task valence and aversive environments (Fig. 3c middle panel). Simulation analyses further demonstrated that the aversive environment decreased the repetition effect on first stages in the reward condition (Fig. 3c, right panel). Moreover, we found high correlations between model-free scores of the real data and model-free scores of the simulated data ( $\rho = 0.71$ ,  $p = < 0.0001$ ) and similarly high correlations between model-based scores of the measured data and model-based scores of the simulated data ( $\rho = 0.7$ ,  $p = < 0.0001$ ). When performing the linear mixed effects logistic regression with the simulated data, we again found a 2-way interaction between task valence and background ( $p = 0.034$ ) and a 3-way interaction between outcome, task valence, and background ( $p = 0.018$ ), suggesting that our model parameters indeed captured the observed effects (Table 1, simulation data).

### Association with subclinical depression scores

We further wanted to test the association between the observed influence of aversive stimuli on model-free learning in the punishment and reward condition and interindividual variation in subclinical depression scores. In accordance with the fact that none of the subjects met the criteria for major depression, subjects scored very low on the HADS subscale. Most subjects stated that they had no depressive symptoms at

**Fig. 3** Results of how aversive stimuli interacted with model-free and model-based reward and punishment learning and first-stage repetition, which indicates stochasticity or exploration behavior, independent of learning. Bar plots indicate means and standard errors of means for the difference scores between aversive versus neutral background condition **a Left:** aversive background enhanced model-free learning in the punishment condition, but reduced model-free learning in the reward condition. **Middle:** aversive background did not influence model-based learning in the reward or punishment condition. **Right:** aversive background selectively decreased first-stage repetition behavior in the reward condition. **b** The stage-skipping update parameter  $\lambda$  was differently influenced by aversive background in the reward and punishment condition. Other learning parameters were not affected by aversive background in the reward or punishment condition. **c** Behavioral data simulations from the extracted parameters captured the observed reinforcement-dependent effect of aversive background on model-free learning (**left**) and the repetition effect (**right**) and replicated the null finding regarding model-based learning (**middle**)



all ( $n = 19$ ), whereas 16 subjects reported one depressive symptom and 17 subjects reported two or more depressive symptoms. For exploratory analyses, we tested whether the extreme ends of this distribution (none vs. some depressive symptoms) would differ in the observed effect. Therefore, we performed the abovementioned regression analysis with an additional between-subject factor, indicating the occurrence of depressive symptoms. Indeed, this regression analysis revealed a 4-way interaction between the observed effect and the presence of depressive symptoms (task valence  $\times$  outcome  $\times$  background  $\times$  group,  $p =$

0.032, Fig. 4a). Post hoc analyses where interactions were separately tested for both groups indicated that only subjects with mild depressive scores showed valence-dependent influences of the background on model-free learning (task valence  $\times$  outcome  $\times$  background,  $p = 0.002$ ), whereas this effect was absent in subjects with no depressive symptoms (task valence  $\times$  outcome  $\times$  background,  $p = 0.772$ ). There was no interaction between group, task valence, and background ( $p = 0.938$ ) indicating that modulation of first-stage repetition by task valence and background was not affected by depressive symptoms (Fig. 4b).

**Table 1** Fixed effects results from the mixed effects logistic regression, where stay behavior on first stages was regressed on previous trial's outcome  $\times$  transition  $\times$  task valence  $\times$  background. Significant main effects and interactions are displayed in italics. Effects are displayed for the measured data (left column) and the simulated data from the computational modeling parameters (right column). Effects of the simulated data match observed effects from the measured data

	Real data		Simulation data	
	Estimate	<i>p</i> value	Estimate	<i>p</i> value
<i>Intercept</i>	1.191	< 0.0001	1.259	< 0.0001
<i>Outcome</i>	0.432	< 0.0001	0.547	< 0.0001
<i>Transition</i>	0.121	0.019	0.160	< 0.0001
Task valence	0.072	0.092	0.170	0.062
Background	0.087	0.152	0.131	0.061
<i>Outcome <math>\times</math> transition</i>	1.403	< 0.0001	1.279	< 0.0001
Task valence $\times$ transition	-0.510	0.546	0.041	0.178
Outcome $\times$ task valence	-0.059	0.466	0.021	0.803
Transition $\times$ background	-0.002	0.983	-0.008	0.689
Outcome $\times$ background	0.005	0.947	0.124	0.105
<i>Task valence <math>\times</math> background</i>	0.329	0.011	0.379	0.034
Transition $\times$ outcome $\times$ task valence	0.014	0.595	0.205	0.170
Transition $\times$ outcome $\times$ background	0.013	0.947	0.086	0.459
Transition $\times$ task valence $\times$ background	0.011	0.942	-0.036	0.437
<i>Outcome <math>\times</math> task valence <math>\times</math> background</i>	0.314	0.041	0.320	0.018
Transition $\times$ outcome $\times$ task valence $\times$ background	-0.047	0.898	0.020	0.928

## Discussion

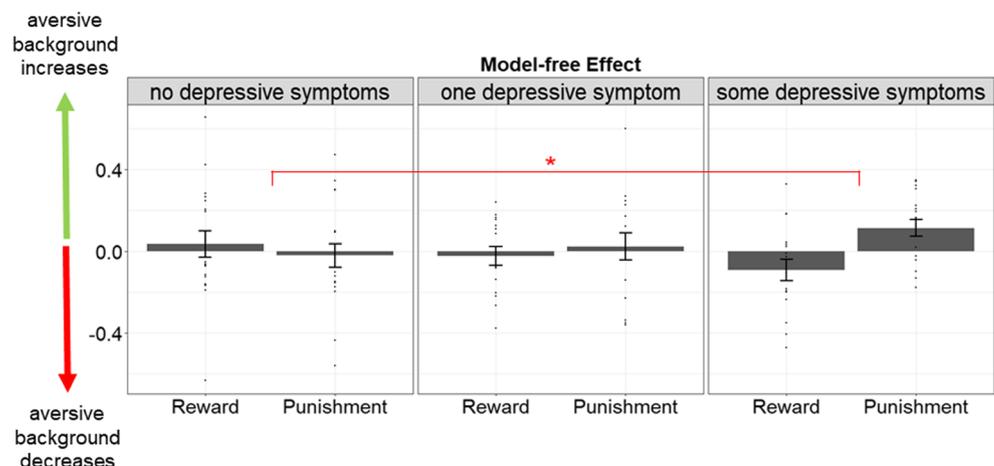
Here, we show that aversive context stimuli influence complex learning and decision-making in a way that depends on task valence. We found that aversive stimuli increased model-free learning when subjects had to avoid punishments but decreased model-free learning when subjects had to approach rewards.

This result is in accordance with a number of studies demonstrating valence-action interactions in the negative domain, as facilitation of avoidance responses have been reported in the presence of aversive stimuli (Campese et al. 2017; Dickinson and Pearce 1977). Moreover, our finding matches reports about aversive Pavlovian-to-instrumental transfer

(PIT) effects in humans (Geurts et al. 2013a; Lally et al. 2017; Nord et al. 2018; Rigoli et al. 2012), where conditioned aversive stimuli decrease operant responses. The finding of our study extends these previous reports and shows that aversive stimuli do not only interact with rather simple approach/avoidance responses but also with more complex learning and decision-making processes. This is particularly important given the fact that humans are often required to make challenging decisions in various emotional states or environments.

According to a recent computational account, complex learning and decision-making can be dichotomized along two control systems, a habitual model-free system and a goal-directed model-based system (Daw et al. 2011). The here applied two-step task enables to disentangle the relative

**Fig. 4** Exploratory analysis comparing subjects at the two ends of a depression score showed that mainly subjects with some depressive symptoms showed aversive background influences of aversive background stimuli on model-free learning depending on the task valence



contribution of these control systems. Interestingly, in our study, only model-free, but not model-based learning, was modulated in a valence-dependent way by aversive stimuli.

This finding is particularly interesting given the large debate on whether valence-action interactions as seen in PIT relate to model-free or model-based control mechanisms (Dayan and Berridge 2014). One recent study showed that aversive PIT in humans reflected habitual rather than goal-directed behavior (Garofalo and Robbins 2017), which speaks to an association between model-free mechanisms and PIT. Although our procedure was fundamentally different from the PIT procedure, it shares one feature, that is, Pavlovian control over operant responding. Pavlovian and model-free control mechanisms are assumed to be innately specified. Interestingly, bodily freezing—a major innately specified defensive responses to threat (Blanchard et al. 2001)—is associated with Pavlovian control over operant responding (Ly et al. 2014), which speaks to a subtle (although no direct) evidence that Pavlovian control over operant behavior relates to the model-free system. Although, in our study, aversive stimuli did not impact model-based decision-making, some other studies have shown that negative states, such as acute stress (Heller et al. 2018; Otto et al. 2013), decrease model-based control (but see Gillan et al. 2019). One possible explanation for the discrepant results is the qualitative different manipulation types (acute stress vs. aversive stimuli) which might tax different control mechanisms.

A computational model substantiated our findings of a valence-dependent effect of aversive stimuli on model-free control. The computational parameter  $\lambda$ , which codes influences of second-stage reward prediction errors on first-stage action values, was increased under aversive influences in the punishment condition but decreased in the reward condition. No other learning parameters were affected by task valence and background.

Beyond learning, we found an interaction between task valence and background on choice behavior which was independent from learning. More precisely, in our study, aversive environments increased switch behavior in the reward condition. Increased first-stage switch behavior could either reflect exploration behavior, stochasticity, or noise, and unfortunately, the two-step task does not enable to disentangle these components. However, two parameters from the softmax could potentially have mirrored the effect:  $\beta_1$ , indicating how first-stage choices are guided by expected values and the repetition parameter  $\rho$ , indicating general stickiness. None of these parameters was modulated by task valence and environment. We believe that the failure of the computational model to capture the effect from the regression analysis relates to a power problem as the large variance within the softmax parameters makes it harder to detect between condition variations. However, despite the failure of the computational model to replicate the repetition effect, our simulation analyses demonstrated

the ability of the model to predict all raw data results, which shows that the model could recover all behavioral effects.

Our computational model further provides a mechanistic framework on how aversive stimuli affect reward and avoidance learning. A famous biologically based computational model proposes distinct dopaminergic cortico-striatal pathways for both types of learning (Frank et al. 2004). According to this model, dopamine bursts increase synaptic plasticity in the direct pathway and thereby support reward learning whereas dopaminergic dips reduce synaptic plasticity in the indirect pathway through which avoidance learning is strengthened. Crucially, reductions in tonic dopamine levels, as seen in Parkinson's disease, lead to increased avoidance learning but disrupted reward learning, a functional status that is reversed through dopaminergic medication (Frank 2005; Frank et al. 2004). Interestingly, this finding is in line with our data of increased model-free punishment learning but decreased model-free reward learning in the presence of aversive stimuli. As it seems unlikely that aversive pictures change individual's tonic dopamine levels, one might speculate that aversive stimuli alter phasic dopaminergic activity, e.g., increase dopaminergic dips induced by punishment (model-free negative prediction errors) and decreased reward-related dopaminergic bursts (model-free positive prediction errors). Future studies should investigate this hypothesis in animals using microdialysis or single cell recordings.

Interestingly, exploration behavior (which might be reflected by the here observed effect of increased switching during reward learning with aversive background) has also been associated with tonic dopamine. L-DOPA selectively increased first- (Wunderlich et al. 2012) and second- (Kroemer et al. 2019) stage switch behavior in the reward version of the two-step task and remediated task switching problems in Parkinson patients (Cools et al. 2001). Studies investigating the catechol-O-methyltransferase (COMT) gene further support a model in which exploration has a dopaminergic basis (Frank et al. 2009; Kayser et al. 2015). Whereas the mentioned studies suggest a positive association between dopamine and exploration, one recent study suggested the opposite direction of this association as dopaminergic blockade in rats led to increases of random exploration behavior (Cinotti et al. 2019). Thus, future studies should further investigate the association between dopamine and the influence of negative states on exploration behavior.

With regard to interindividual differences, our exploratory analyses showed that valence-dependent effects of how aversive stimuli affect model-free learning were only evident in subjects reporting mild, subclinical symptoms of depression, but not in subjects with no symptoms in depression. This effect is in line with studies demonstrating increased aversive PIT effects in depression (Nord et al. 2018) and matches a recent study that reports increased influences of negative states on complex decision-making in subjects with depressive symptoms (Heller et al. 2018).

The dysfunctional serotonergic system in depression (Belmaker and Agam 2008; Graeff et al. 1996; Mann 1999) might relate to these altered influences of aversive stimuli on learning and decision-making. Evidence comes from studies using tryptophan depletion to lower central 5-HT levels (Crockett et al. 2012b). Some studies have shown that tryptophan depletion lowers the impact of aversive stimuli (Crockett et al. 2012a) or expected punishments (Crockett et al. 2012a; Crockett et al. 2009; den Ouden et al. 2015) on inhibition and decreases aversive PIT effects (Geurts et al. 2013b), suggesting that low serotonin reduces the influence of aversive stimuli on actions. However, increases in the influence of aversive stimuli on response inhibition have also been reported under tryptophan depletion (Cools et al. 2008; Hebart and Glascher 2015). Moreover, two studies have shown that tryptophan depletion impairs model-based reward learning (Worbe et al. 2015) while it enhances model-based avoidance learning (Worbe et al. 2016), suggesting that low serotonin may have valence-dependent effects. Moreover, in line with the assumption that low serotonin induces negative biases, genetic studies show that subjects carrying the less functional *s* allele in the promoter region of the serotonin transporter gene (5-HTTLPR) show increased avoidance learning (Finger et al. 2007) and amplified neural responses to aversive stimuli (Canli et al. 2005; Hariri et al. 2005; Heinz et al. 2005). In the current study, we did not assess central serotonin levels or genetic variations of the serotonergic transporter gene and future studies should further investigate the interactions between serotonin deficiency, depressive symptoms, and emotional biases of decision-making to elucidate potential vulnerability markers for depression.

Our study has several limitations: first, the statistical evidence for our observed effects are tentative. The targeted interaction effects in this study are of higher order (4- and 5-fold interactions), and large sample sizes are usually needed to detect these kinds of effects. Additional analyses where we tested the hypothesis that our observed effects were stronger for model-free than for model-based control (see S3) failed to reach statistical significance. Therefore, although we have direct evidence that model-free control was affected by task valence and background, we cannot conclude that our effects were statistically stronger for model-free compared with model-based control. As this might again be a power problem, our findings ultimately need replication in bigger sample sizes.

Moreover, we used negative IAPS pictures that were rated as more negative in valence, but also as more arousing compared with neutral pictures. Therefore, we cannot rule out that the observed effects were due to arousal instead of valence. The arousal system has been closely linked to fight or flight reactions (Roelofs 2017), and a recent study has shown that arousal but not valence of stimuli affects approach behavior (Bouman et al. 2015). Moreover, arousal as indexed by pupil diameters is positively related to exploration (Jepma and Nieuwenhuis 2011), which is in line with our observation on increased switching

behavior (potentially indexing exploration) in aversive compared with neutral environments.

However, two findings speak against the arousal hypothesis. First, a u-shaped relationship between arousal and reaction times has been reported, where moderate and high arousal decreases and increases RTs, respectively (Nishisato 1966; Welford 1971). Second, an inverted u-shape association between arousal and task performance has been reported (Broadhurst 1957; Neiss 1988), where moderate and high arousal increases and decreases performance respectively. Model-based control is the more sophisticated control system in the two-step task and might be related to task performance as it correlates with processing speed and working memory capacity (Schad et al. 2014). As we used moderately arousing pictures, we could have expected decreases in first-stage RTs but increased model-based control. However, aversive environments neither affected reaction times (see S2) nor model-based control, which speaks against the arousal hypothesis. The additional use of positive, highly arousing pictures would allow ruling out the alternative arousal instead of valence explanation. Due to time restriction, we could not implement this additional manipulation in the current experiment.

One second limitation of the study is the extremely homogeneous sample (male 21-year-old subjects), which limits the generalizability of its findings. Also, because we only included subjects with no current axis-1 psychiatric disorder, the distribution of the depression score was heavily skewed with a strong zero inflation. Beyond this, one further limitation was that depressive symptoms were assessed in a separate session prior to the assessment of the task. A recent study on almost 200 young adults (Jinnin et al. 2017) showed that mild depressive symptoms remained stable over a 1-year follow-up period, which makes us believe that this could also hold for this sample. However, further studies that aim at investigating depression vulnerability should sample subjects from a high-risk population, e.g., first-degree relatives of depressive patients.

Besides these limitations, our study is the first to provide a computational model of how negative environmental stimuli influence complex decision-making. Thus, these results are of high interest for a better understanding of context-specific effects on learning mechanisms in everyday life. Further, our study adds a mechanistic model of how alterations of these effects can be understood as a potential vulnerability marker for depression.

**Funding** This work was supported by the following institutions: German Research Foundation (Deutsche Forschungsgemeinschaft, DFG, HE2597/14-1, HE2597/14-2, RA1047/2-1, RA1047/2-2, RO 5046/2-2, SCHA1971/1-2, SCHL 1969/2-2, SCHL 1969/4-1); Federal Ministry of Health (Bundesministerium für Gesundheit, BMG, ZMV11-2516DSM223).

## Compliance with ethical standards

**Conflict of interests** The authors declare that they have no conflict of interest.

**Open Access** This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

## References

- Bassi A (2013) Weather, mood, and voting: an experimental analysis of the effect of weather beyond turnout
- Bates D, Mächler M, Bolker B, Walker S (2015) Fitting linear mixed-effects models using lme4 2015 67:48 doi:<https://doi.org/10.18637/jss.v067.i01>
- Belmaker RH, Agam G (2008) Major depressive disorder. *N Engl J Med* 358:55–68. <https://doi.org/10.1056/NEJMra073096>
- Blanchard DC, Griebel G, Blanchard RJ (2001) Mouse defensive behaviors: pharmacological and behavioral assays for anxiety and panic. *Neurosci Biobehav Rev* 25:205–218
- Bouman D, Stins JF, Beek PJ (2015) Arousal and exposure duration affect forward step initiation. *Front Psychol* 6:1667. <https://doi.org/10.3389/fpsyg.2015.01667>
- Broadhurst PL (1957) Emotionality and the Yerkes-Dodson law. *J Exp Psychol* 54:345–352
- Campese VD, Kim IT, Rojas G, LeDoux JE (2017) Pavlovian extinction and recovery effects in aversive Pavlovian to instrumental transfer. *Front Behav Neurosci* 11:179. <https://doi.org/10.3389/fnbeh.2017.00179>
- Canli T, Omura K, Haas BW, Fallgatter A, Constable RT, Lesch KP (2005) Beyond affect: a role for genetic variation of the serotonin transporter in neural activation during a cognitive attention task. *Proc Natl Acad Sci U S A* 102:12224–12229. <https://doi.org/10.1073/pnas.0503880102>
- Cinotti F, Fresno V, Aklil N, Coutureau E, Girard B, Marchand AR, Khamassi M (2019) Dopamine blockade impairs the exploration-exploitation trade-off in rats. *Sci Rep* 9:6770. <https://doi.org/10.1038/s41598-019-43245-z>
- Cools R, Barker RA, Sahakian BJ, Robbins TW (2001) Enhanced or impaired cognitive function in Parkinson's disease as a function of dopaminergic medication and task demands. *Cereb Cortex* 11:1136–1143
- Cools R, Roberts AC, Robbins TW (2008) Serotonergic regulation of emotional and behavioural control processes. *Trends Cogn Sci* 12: 31–40. <https://doi.org/10.1016/j.tics.2007.10.011>
- Crockett MJ, Clark L, Apergis-Schoute AM, Morein-Zamir S, Robbins TW (2012a) Serotonin modulates the effects of Pavlovian aversive predictions on response vigor. *Neuropsychopharmacology* 37: 2244–2252. <https://doi.org/10.1038/npp.2012.75>
- Crockett MJ, Clark L, Robbins TW (2009) Reconciling the role of serotonin in behavioral inhibition and aversion: acute tryptophan depletion abolishes punishment-induced inhibition in humans. *J Neurosci* 29:11993–11999. <https://doi.org/10.1523/JNEUROSCI.2513-09.2009>
- Crockett MJ, Clark L, Roiser JP, Robinson OJ, Cools R, Chase HW, den Ouden H, Apergis-Schoute A, Campbell-Meikelljohn D, Seymour B, Sahakian BJ, Rogers RD, Robbins TW (2012b) Converging evidence for central 5-HT effects in acute tryptophan depletion. *Mol Psychiatry* 17:121–123. <https://doi.org/10.1038/mp.2011.106>
- Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ (2011) Model-based influences on humans' choices and striatal prediction errors. *Neuron* 69:1204–1215. <https://doi.org/10.1016/j.neuron.2011.02.027>
- Dayan P, Berridge KC (2014) Model-based and model-free Pavlovian reward learning: reevaluation, revision, and revelation. *Cogn Affect Behav Neurosci* 14:473–492. <https://doi.org/10.3758/s13415-014-0277-8>
- Dayan P, Niv Y, Seymour B, Daw ND (2006) The misbehavior of value and the discipline of the will. *Neural Netw* 19:1153–1160. <https://doi.org/10.1016/j.neunet.2006.03.002>
- den Ouden HE, Swart JC, Schmidt K, Fekkes D, Geurts DE, Cools R (2015) Acute serotonin depletion releases motivated inhibition of response vigour. *Psychopharmacology* 232:1303–1312. <https://doi.org/10.1007/s00213-014-3762-4>
- Dickinson A, Pearce JM (1977) Inhibitory interactions between appetitive and aversive stimuli. *Psychol Bull* 84:690–711. <https://doi.org/10.1037/0033-2909.84.4.690>
- Eshel N, Roiser JP (2010) Reward and punishment processing in depression. *Biol Psychiatry* 68:118–124. <https://doi.org/10.1016/j.biopsych.2010.01.027>
- Finger EC, Marsh AA, Buzas B, Kamel N, Rhodes R, Vythilingham M, Pine DS, Goldman D, Blair JR (2007) The impact of tryptophan depletion and 5-HTTLPR genotype on passive avoidance and response reversal instrumental learning tasks. *Neuropsychopharmacology* 32:206–215. <https://doi.org/10.1038/sj.npp.1301182>
- Frank MJ (2005) Dynamic dopamine modulation in the basal ganglia: a neurocomputational account of cognitive deficits in medicated and nonmedicated parkinsonism. *J Cogn Neurosci* 17:51–72. <https://doi.org/10.1162/0898929052880093>
- Frank MJ, Doll BB, Oas-Terpstra J, Moreno F (2009) Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nat Neurosci* 12:1062–1068. <https://doi.org/10.1038/nn.2342>
- Frank MJ, Seeberger LC, O'Reilly RC (2004) By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* 306: 1940–1943. <https://doi.org/10.1126/science.1102941>
- Friedel E, Koch SP, Wendt J, Heinz A, Deserno L, Schlagenhaut F (2014) Devaluation and sequential decisions: linking goal-directed and model-based behavior. *Front Hum Neurosci* 8:587. <https://doi.org/10.3389/fnhum.2014.00587>
- Friedrich S, Brunner E, Pauly M (2017) Permuting longitudinal data in spite of the dependencies. *J Multivar Anal* 153:255–265
- Garbusow M, Schad DJ, Sebold M, Friedel E, Bernhardt N, Koch SP, Steinacher B, Kathmann N, Geurts DEM, Sommer C, Müller DK, Nebe S, Paul S, Wittchen HU, Zimmermann US, Walter H, Smolka MN, Sterzer P, Rapp MA, Huys QJM, Schlagenhaut F, Heinz A (2016) Pavlovian-to-instrumental transfer effects in the nucleus accumbens relate to relapse in alcohol dependence. *Addict Biol* 21: 719–731. <https://doi.org/10.1111/adb.12243>
- Garbusow M, Schad DJ, Sommer C, Jünger E, Sebold M, Friedel E, Wendt J, Kathmann N, Schlagenhaut F, Zimmermann US, Heinz A, Huys QJM, Rapp MA (2014) Pavlovian-to-instrumental transfer in alcohol dependence: a pilot study. *Neuropsychobiology* 70:111–121. <https://doi.org/10.1159/000363507>
- Garofalo S, Robbins TW (2017) Triggering avoidance: dissociable influences of aversive Pavlovian conditioned stimuli on human instrumental behavior. *Front Behav Neurosci* 11:63. <https://doi.org/10.3389/fnbeh.2017.00063>
- Geurts DE, Huys QJ, den Ouden HE, Cools R (2013a) Aversive Pavlovian control of instrumental behavior in humans. *J Cogn Neurosci* 25:1428–1441. [https://doi.org/10.1162/jocn\\_a\\_00425](https://doi.org/10.1162/jocn_a_00425)
- Geurts DE, Huys QJ, den Ouden HE, Cools R (2013b) Serotonin and aversive Pavlovian control of instrumental behavior in humans. *J Neurosci* 33:18932–18939. <https://doi.org/10.1523/JNEUROSCI.2749-13.2013>
- Gillan C et al. (2019) Experimentally-induced and real-world acute anxiety have no effect on goal-directed behaviour bioRxiv:606145 <https://doi.org/10.1101/606145>

- Gillan CM, Otto AR, Phelps EA, Daw ND (2015) Model-based learning protects against forming habits. *Cogn Affect Behav Neurosci* 15: 523–536. <https://doi.org/10.3758/s13415-015-0347-6>
- Gillan CM, Robbins TW, Sahakian BJ, van den Heuvel OA, van Wingen G (2016) The role of habit in compulsivity. *Eur Neuropsychopharmacol* 26:828–840. <https://doi.org/10.1016/j.euroneuro.2015.12.033>
- Graeff FG, Guimaraes FS, De Andrade TG, Deakin JF (1996) Role of 5-HT in stress, anxiety, and depression. *Pharmacol Biochem Behav* 54:129–141
- Guitart-Masip M, Chowdhury R, Sharot T, Dayan P, Duzel E, Dolan RJ (2012) Action controls dopaminergic enhancement of reward representations. *Proc Natl Acad Sci U S A* 109:7511–7516. <https://doi.org/10.1073/pnas.1202229109>
- Guitart-Masip M, Duzel E, Dolan R, Dayan P (2014) Action versus valence in decision making. *Trends Cogn Sci* 18:194–202. <https://doi.org/10.1016/j.tics.2014.01.003>
- Guitart-Masip M, Fuentemilla L, Bach DR, Huys QJ, Dayan P, Dolan RJ, Duzel E (2011) Action dominates valence in anticipatory representations in the human striatum and dopaminergic midbrain. *J Neurosci* 31:7867–7875. <https://doi.org/10.1523/JNEUROSCI.6376-10.2011>
- Gureckis TM, Love BC (2009) Short-term gains, long-term pains: how cues about state aid learning in dynamic environments. *Cognition* 113:293–313. <https://doi.org/10.1016/j.cognition.2009.03.013>
- Hariri AR, Drabant EM, Munoz KE, Kolachana BS, Mattay VS, Egan MF, Weinberger DR (2005) A susceptibility gene for affective disorders and the response of the human amygdala. *Arch Gen Psychiatry* 62:146–152. <https://doi.org/10.1001/archpsyc.62.2.146>
- Hebart MN, Glascher J (2015) Serotonin and dopamine differentially affect appetitive and aversive general Pavlovian-to-instrumental transfer. *Psychopharmacology* 232:437–451. <https://doi.org/10.1007/s00213-014-3682-3>
- Heinz A, Braus DF, Smolka MN, Wrase J, Puls I, Hermann D, Klein S, Grüsser SM, Flor H, Schumann G, Mann K, Büchel C (2005) Amygdala-prefrontal coupling depends on a genetic variation of the serotonin transporter. *Nat Neurosci* 8:20–21. <https://doi.org/10.1038/nn1366>
- Heller AS, Ezie CEC, Otto AR, Timpano KR (2018) Model-based learning and individual differences in depression: the moderating role of stress. *Behav Res Ther* 111:19–26. <https://doi.org/10.1016/j.brat.2018.09.007>
- Herrmann-Lingen C (2005) Hospital anxiety and depression scale : HADS-D ; deutsche Version ; ein Fragebogen zur Erfassung von Angst und Depressivität in der somatischen Medizin. Huber
- Huys QJ, Daw ND, Dayan P (2015a) Depression: a decision-theoretic analysis. *Annu Rev Neurosci* 38:1–23. <https://doi.org/10.1146/annurev-neuro-071714-033928>
- Huys QJM, Guitart-Masip M, Dolan RJ, Dayan P (2015b) Decision-theoretic psychiatry. *Clin Psychol Sci* 3:400–421. <https://doi.org/10.1177/2167702614562040>
- Jacobi F, Mack S, Gerschler A, Scholl L, Höfler M, Siegert J, Bürkner A, Preiss S, Spitzer K, Busch M, Hapke U, Gaebel W, Maier W, Wagner M, Zielasek J, Wittchen HU (2013) The design and methods of the mental health module in the German Health Interview and Examination Survey for Adults (DEGS1-MH). *Int J Methods Psychiatr Res* 22:83–99. <https://doi.org/10.1002/mpr.1387>
- Jepma M, Nieuwenhuis S (2011) Pupil diameter predicts changes in the exploration-exploitation trade-off: evidence for the adaptive gain theory. *J Cogn Neurosci* 23:1587–1596. <https://doi.org/10.1162/jocn.2010.21548>
- Jinnin R, Okamoto Y, Takagaki K, Nishiyama Y, Yamamura T, Okamoto Y, Miyake Y, Takebayashi Y, Tanaka K, Sugiura Y, Shimoda H, Kawakami N, A Furukawa T, Yamawaki S (2017) Detailed course of depressive symptoms and risk for developing depression in late adolescents with subthreshold depression: a cohort study. *Neuropsychiatr Dis Treat* 13:25–33. <https://doi.org/10.2147/NDT.S117846>
- Kayser AS, Mitchell JM, Weinstein D, Frank MJ (2015) Dopamine, locus of control, and the exploration-exploitation tradeoff. *Neuropsychopharmacology* 40:454–462. <https://doi.org/10.1038/npp.2014.193>
- Kroemer NB, Lee Y, Poese S, Eppinger B, Goschke T, Smolka MN (2019) L-DOPA reduces model-free control of behavior by attenuating the transfer of value to action. *Neuroimage* 186:113–125. <https://doi.org/10.1016/j.neuroimage.2018.10.075>
- Lally N, Huys QJM, Eshel N, Faulkner P, Dayan P, Roiser JP (2017) The neural basis of aversive Pavlovian guidance during planning. *J Neurosci* 37:10215–10229. <https://doi.org/10.1523/JNEUROSCI.0085-17.2017>
- Lang P, Bradley M, Cuthbert BN (2008) International affective picture system (IAPS): affective ratings of pictures and instruction manual.
- Lang PJ, Bradley MM, Cuthbert BN (2005) International affective picture system (IAPS) : affective ratings of pictures and instruction manual. NIMH. In: Center for the Study of Emotion & Attention. Gainesville, Fla
- Lee JL, Dickinson A, Everitt BJ (2005) Conditioned suppression and freezing as measures of aversive Pavlovian conditioning: effects of discrete amygdala lesions and overtraining. *Behav Brain Res* 159: 221–233. <https://doi.org/10.1016/j.bbr.2004.11.003>
- Lewis AH, Niznikiewicz MA, Delamater AR, Delgado MR (2013) Avoidance-based human Pavlovian-to-instrumental transfer. *Eur J Neurosci* 38:3740–3748. <https://doi.org/10.1111/ejn.12377>
- Liu S, Schad DJ, Kuschpel MS, Rapp MA, Heinz A (2016) Music and video gaming during breaks: influence on habitual versus goal-directed decision making. *PLoS One* 11:e0150165. <https://doi.org/10.1371/journal.pone.0150165>
- Ly V, Huys QJ, Stins JF, Roelofs K, Cools R (2014) Individual differences in bodily freezing predict emotional biases in decision making. *Front Behav Neurosci* 8:237. <https://doi.org/10.3389/fnbeh.2014.00237>
- Mann JJ (1999) Role of the serotonergic system in the pathogenesis of major depression and suicidal behavior. *Neuropsychopharmacology* 21:99S–105S. [https://doi.org/10.1016/S0893-133X\(99\)00040-8](https://doi.org/10.1016/S0893-133X(99)00040-8)
- MATLAB (2014) Matlab version 8.3.0 (2014a). the MathWorks Inc., Natick, Massachusetts
- Nebe S, Kroemer NB, Schad DJ, Bernhardt N, Sebold M, Müller DK, Scholl L, Kuitunen-Paul S, Heinz A, Rapp MA, Huys QJM, Smolka MN (2018) No association of goal-directed and habitual control with alcohol consumption in young adults. *Addict Biol* 23:379–393. <https://doi.org/10.1111/adb.12490>
- Neiss R (1988) Reconceptualizing arousal: psychobiological states in motor performance. *Psychol Bull* 103:345–366
- Nelson BJ, Sanjuan MC (2006) A context-specific latent inhibition effect in a human conditioned suppression task. *Q J Exp Psychol (Hove)* 59:1003–1020. <https://doi.org/10.1080/17470210500417738>
- Nishisato S (1966) Reaction time as a function of arousal and anxiety. *Psychon Sci* 6:157–158. <https://doi.org/10.3758/bf03328005>
- Nord CL, Lawson RP, Huys QJM, Pilling S, Roiser JP (2018) Depression is associated with enhanced aversive Pavlovian control over instrumental behaviour. *Sci Rep* 8:12582. <https://doi.org/10.1038/s41598-018-30828-5>
- Otto AR, Raio CM, Chiang A, Phelps EA, Daw ND (2013) Working-memory capacity protects model-based learning from stress. *Proc Natl Acad Sci U S A* 110:20941–20946. <https://doi.org/10.1073/pnas.1312011110>
- Radenbach C, Reiter AMF, Engert V, Sjoerds Z, Villringer A, Heinze HJ, Deserno L, Schlagenhaut F (2015) The interaction of acute and chronic stress impairs model-based behavioral control. *Psychoneuroendocrinology* 53:268–280. <https://doi.org/10.1016/j.psyneuen.2014.12.017>

- Rigoli F, Pavone EF, Pezzulo G (2012) Aversive pavlovian responses affect human instrumental motor performance. *Front Neurosci* 6: 134. <https://doi.org/10.3389/fnins.2012.00134>
- Roelofs K (2017) Freeze for action: neurobiological mechanisms in animal and human freezing. *Philos Trans R Soc Lond Ser B Biol Sci* 372:20160206. <https://doi.org/10.1098/rstb.2016.0206>
- Schad DJ, Janger E, Sebold M, Garbusow M, Bernhardt N, Javadi AH, Zimmermann US, Smolka MN, Heinz A, Rapp MA, Huys QJM (2014) Processing speed enhances model-based over model-free reinforcement learning in the presence of high working memory functioning. *Front Psychol* 5:1450. <https://doi.org/10.3389/fpsyg.2014.01450>
- Sebold M, Deserno L, Nebe S, Schad DJ, Garbusow M, Hagele C, Keller J, Junger E, Kathmann N, Smolka M, Rapp MA, Schlagenhauf F, Heinz A, Huys QJM (2014) Model-based and model-free decisions in alcohol dependence. *Neuropsychobiology* 70:122–131. <https://doi.org/10.1159/000362840>
- Welford AT (1971) What is the basis of choice reaction-time? *Ergonomics* 14:679–693. <https://doi.org/10.1080/00140137108931291>
- Winkielman P, Berridge KC, Wilbarger JL (2005) Unconscious affective reactions to masked happy versus angry faces influence consumption behavior and judgments of value. *Personal Soc Psychol Bull* 31: 121–135. <https://doi.org/10.1177/0146167204271309>
- Wittchen HU, Pfister H (1997) DIA-X-Interviews: Manual fur Screening-Verfahren und Interview; Interviewheft. In: Swets & Zeitlinger, Frankfurt
- Worbe Y, Palminteri S, Savulich G, Daw ND, Fernandez-Egea E, Robbins TW, Voon V (2016) Valence-dependent influence of serotonin depletion on model-based choice strategy. *Mol Psychiatry* 21: 624–629. <https://doi.org/10.1038/mp.2015.46>
- Worbe Y, Savulich G, de Wit S, Fernandez-Egea E, Robbins TW (2015) Tryptophan depletion promotes habitual over goal-directed control of appetitive responding in humans. *Int J Neuropsychopharmacol* 18:pyv013. <https://doi.org/10.1093/ijnp/pyv013>
- Wunderlich K, Smittenaar P, Dolan RJ (2012) Dopamine enhances model-based over model-free choice behavior. *Neuron* 75:418–424. <https://doi.org/10.1016/j.neuron.2012.03.042>
- Zigmond AS, Snaith RP (1983) The hospital anxiety and depression scale. *Acta Psychiatr Scand* 67:361–370

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.