Supplementary Text and Figures

Contents

Supplementary Methods **2**
Participants . . . 2

Sample size **3**

Demographic and psychopathology measures **4**
Demographic information . . . 4
Psychometric questionnaires . . . 4

Task design **6**
Additional Task Information . . . 6
Stimuli . . . 6

Computational Models **7**
RLDDM . . . 7
Estimation . . . 8
Priors . . . 8

Data analysis **9**
Quality control . . . 9
Models comparison . . . 9
Collinearity check . . . 11
Posterior predictive checks . . . 12
Additional figures . . . 15

Comorbidity **20**

Medication **22**

Biased choices **23**

Outcome-irrelevant learning: spatial-motor associations **25**

References **27**

Supplementary Methods

Participants

To ensure the relevance of our findings to the broader psychiatric population, individuals with R-AN who had comorbid psychiatric conditions were included in the study. The presence of comorbidities was determined by specialized psychiatrists and psychologists using a semi-structured interview based on the Mini International Neuropsychiatric Interview (MINI). Among the 40 individuals with R-AN, comorbidities included anxiety disorder (n=16), obsessive-compulsive disorder (OCD) (n=8), social phobia (n=1), and depressive and anxiety disorders (n=1). Some R-AN patients were also taking medication, including anxiolytic antidepressants (n=10), Selective Serotonin Reuptake Inhibitors (SSRIs) (n=6), benzodiazepines (n=1), and mood stabilizers (lithium) (n=1).

To be eligible for participation, individuals needed to be proficient in spoken and written Italian. Exclusion criteria for all participants included a history of alcohol or drug abuse or dependence, neurological disorders, past or present psychiatric diagnosis, and intellectual or developmental disability. Cognitive function within the normal range was assessed using the Raven's Standard Progressive Matrices test (Raven, 2000). The eligibility criteria for all participants were evaluated through psychologist interviews. Body mass index (BMI) values were determined in the laboratory.

The study included a predominantly Caucasian sample, with 97.7% of the participants identifying as Caucasian. A smaller proportion of participants identified as Asian-Italian (1.7%) and African-Italian (0.6%). Additionally, all selected participants were right-handed and were unaware of the specific objectives of the study, ensuring a blind study design.

Sample size

Before conducting the present study, we conducted a parallel but separate study with two distinct groups. The first group included 29 anorexic patients, and the second group consisted of 124 healthy controls (these participants were different from those in the current study). In this prior study, each participant completed 160 trials per condition in a PRL task, where the content of the pair of images presented in each trial was manipulated. In both groups, the difference in the learning rate (which is the main focus of the current study) was measured to be 0.54 (on a logit scale).

To determine the sample size needed to detect a similar effect, we carried out a parameter recovery study following the method outlined by Pedersen & Frank (2020). We simulated the data of two groups of 30 participants with different values of α (lower and higher) with a difference of 0.54. The other parameters of the RLDDM model (*i.e.*, a, t, v) were set to the values estimated from the empirical data of the 29 anorexic patients and 124 healthy controls. For the simulation, we used the `hddm.generate.gen_rand_rlddm_data` function of the `hddm` module with following parameters:

```
subjects = 30
trials = 160

data = hddm.generate.gen_rand_rlddm_data(
    a=1.5,
    alpha=0.79, # or 0.25
    scaler=2.25,
    t=0.25,
    size=trials,
    subjs=subjects,
    p_upper=0.7,
    p_lower=0.3,
)
```

We used the `HDDMrl` function of the `hddm` module to estimate the RLDDM parameters based on the simulated data. We repeated this procedure 100 times, and in each iteration, the parameters for the lower and higher values of α were completely separated. This simulation suggest that our study had enough participants and trials to detect the effect size on α that had been observed in the previous study.

A parameter recovery study and a frequentist power analysis are two distinct approaches. However, since Bayesian methods prioritize estimation over hypothesis testing, it is comforting to see that with the current number of participants and trials, the RLDDM model can detect an effect size similar to the one found in a separate study with a different group of participants but with a similar experimental manipulation.

Demographic and psychopathology measures

Demographic information

Mean age and Body Mass Index (BMI) for each group of participant were as follows: patients with R-AN, mean age = 21.11 (SD = 4.33), average Body Mass Index (BMI) = 17.79 (SD = 2.85); HCs, mean age = 19.49 (SD = 2.32), average BMI = 21.78 (SD = 3.53); healthy individuals at risk of developing eating disorders, mean age = 21.31 (SD = 4.82), average BMI = 21.64 (SD = 4.11).

Bayesian statistical analysis revealed no credible age differences among the three groups (R-AN, HC, and RI). R-AN participants displayed a lower mean BMI than HC participants. No noteworthy difference in BMI was observed between HC and RI participants. Furthermore, there is credible evidence that the Rosenberg Self-Esteem Scale scores of the two groups (R-AN and RI) are smaller than those of the HC group. We also found credible evidence that individuals with R-AN and RI exhibited higher levels of dissatisfaction with their body shape, as measured by the BSQ-14 questionnaire, when compared to the HCs.

Individuals with R-AN displayed higher stress, anxiety, and depression levels (as measured by the DASS-21) than HCs. Additionally, individuals with AN showed credibly higher levels of social interaction anxiety (as measured by the SIAS) than HCs. All R-AN and RI groups exhibited higher levels of Concerns over mistakes and doubts scores of the MPS scale compared to HCs. Individuals with R-AN also showed higher levels of Personal standard scores of the MPS scale compared to HCs. Moreover, individuals with R-AN displayed higher values on all three subscales of the EAT-26 questionnaire relative to HCs.

Sixteen individuals with R-AN were diagnosed with a comorbid anxiety disorder, 8 with obsessive-compulsive disorder, 1 with social phobia, and 1 with panic disorder.

Psychometric questionnaires

The *Eating Attitude Test-26* (EAT-26, Garner et al., 1982) consists of 26 items assessing levels and types of eating disturbances in the past three mouths. The EAT-26 is characterized by three subscales: the Dieting Scale, the Bulimia and Food Preoccupation Scale and the Oral Control Scale. Scores ≥ 20 point out the presence of an eating disorder. Respondents are required to rate intensity associated with the items on a 6-point Likert scale (0 = never, rarely, sometimes; 3 = always). The Italian version of the EAT-26 demonstrated good psychometric properties (Dotti & Lazzari, 1998). In fact, Cronbach's alpha was high in an undergraduate sample for the Dieting scale (.87), for Bulimia and Food Preoccupation scale (.70), for Oral Control Scale (.62). Cronbach's alpha for the total scores was 0.86.

The *Body Shape Questionnaire-14* (BSQ-14, Dowson & Henderson, 2001) is a 14-item self-report scale assessing the global body satisfaction in the past two weeks. Respondents are required to rate intensity of concerns about own appearance associated with the items on a

6-point Likert scale (1 = never, 6 = always). The Italian version of the BSQ-14 demonstrated good psychometric properties (Matera et al., 2013). In the present sample, $\omega = 0.978$. For the 40-item BSQ, a score below 80 is considered "no concern", a score of 80 to 110 is considered "slight concern", a score of 111 to 140 is considered "moderate concern", and a score above 140 is considered "marked concern".

The *Social Interaction Anxiety Scale* (SIAS, Mattick & Clarke, 1998) is a 20-item self-report questionnaire assessing social interaction anxiety. Respondents are required to rate intensity associated with the items on a 4-point Likert scale from 0 (not at all true) to 4 (extremely true). Higher scores denote greater social interaction anxiety levels. Both original version and the Italian version (Sica et al., 2007) show acceptable psychometric properties (in the present sample $\omega = 0.938$). Heimberg et al. (1992) have suggested a cut-off of 34 on the 20-item SIAS to denote a clinical level of social anxiety (32.3 for the Italian 19 item version).

The *Depression Anxiety Stress Scale-21* (DASS-21, Lovibond & Lovibond, 1995) is a 21-item self-report measure assessing depression, anxiety, and stress over the previous week. Items are rated on a 4-point scale ranging from 0 (did not apply to me at all) to 3 (applied to me very much). Both the original and the Italian version (Bottesi et al., 2015) demonstrate adequate reliability. In the present sample $\omega_{\text{anxiety}} = 0.875$, $\omega_{\text{depression}} = 0.914$, $\omega_{\text{stress}} = 0.899$; for the total scale, $\omega = 0.945$.

The *Rosenberg Self-Esteem Scale* (RSES, Rosenberg, 1965) is a 10-item scale designed to assess person's overall self-esteem. It comprises five straightforwardly worded and five reverse-worded items each rated on a 4-point Likert scale ranging from 4 (strongly agree) to 1 (strongly disagree). Increased values indicate increased self-esteem. In the present sample, $\omega = 0.949$.

The *Multidimensional Perfectionism Scale* (MPS, Frost et al., 1990) is a 35-item assessing perfectionism tendencies. According to Stöber (1998), MPS-F is composed of four underlying factors: Concerns over Mistakes and Doubts (CMD), Parental Expectations and Criticism (PEC), Personal Standards (PS), and Organization (O). Both the original MPS and the Italian version (Lombardo, 2008) demonstrate adequate reliability. In the present sample, $\omega_{\text{CMD}} = 0.919$, $\omega_{\text{PS}} = 0.851$, $\omega_{\text{PEPC}} = 0.946$, $\omega_{\text{OR}} = 0.931$; for the total scale, $\omega = 0.932$.

Task design

Additional Task Information

Participants were introduced to the study as a way to evaluate their cognitive functions using a computer-based "game" and additional questionnaires. Their goal in the PRL task was to maximize their earnings, which were shown at the end of each trial block. When participants felt uncertain, they were instructed to rely on their instincts. During the PRL task, feedback was provided in a probabilistic manner. The correct image was rewarded in 70% of the trials, while negative feedback was given in the remaining 30% of the trials. Each block of the task consisted of four epochs, with 40 trials in each epoch where the same image was considered correct. Within each block, there were three rule changes known as reversal phases. Participants were aware that the stimulus-reward contingencies would change, but they were not provided with specific details about when or how these changes would occur. Prior to the actual experiment, participants underwent a training block consisting of 20 trials. The PRL tasks were programmed using the Psychtoolbox extensions in MATLAB (MathWorks) (Brainard & Vision, 1997).

Stimuli

The food-related category consisted of images of french fries, cake, pancake, cheeseburger, and cupcake (IAPS #7461, 7260, 7470, 7451, 7405), while the food-unrelated category included images of a lamp, book, umbrella, basket, and clothespin (IAPS #7175, 7090, 7150, 7041, 7052). For the control task, five images were used for each of the two food-unrelated categories, *i.e.*, five images of flowers (IAPS #5000, 5001, 5020, 5030, 5202) and five images of objects (IAPS #7010, 7020, 7034, 7056, 7170).

Computational Models

The main goal of this study was to use computational models of reinforcement learning to compare the learning outcomes of two different decision-making contexts: those involving disorder-relevant information and those that did not. We used a reinforcement learning drift diffusion model (RLDDM, Pedersen & Frank, 2020) to investigate the impact of disorder-related information (which is irrelevant to the outcome) on decision-making in individuals with restrictive eating disorders (R-AN).

RLDDM

The RLDDM consists of two key components: one describes how reward feedback is employed to update value expectations and the other describes how an agent uses these expectations to arrive at a decision.

The model assumes that subjective option values (Q values) are learned through reward prediction errors (PEs), which measure the disparity between expected and obtained outcomes (Sutton & Barto, 2018). The update of subjective option values follows a delta learning rule (Rescorla & Wagner, 1972):

$$Q_{a,i} = Q_{a,i-1} + \alpha(I_{a,i-1} - Q_{a,i-1}),$$

where Q refers to the expected values for option a on trial i, I represents the reward (with values 1 or 0), and α is the leaning rate, which scales the difference between the expected and actual rewards. A higher learning rate results in rapid adaptation to reward expectations, while a lower learning rate results in slow adaptation. We included in the model different learning rates for positive and negative prediction errors: The parameter α^+ is computed from reinforcements, whereas α^+ is computed from punishments.

The second component describes the selection rule for reinforced options. Typically, a softmax function is used, where the probability of selecting option a depends on its expected value relative to other options n, scaled by the inverse temperature parameter β:

$$p_{a,i} = \frac{e^{\beta Q_{a,i}}}{\sum_{j=1}^{n} e^{\beta Q_{j,i}}}.$$

In the RLDDM, instead, this second component of decision-making is replaced by a Drift-Diffusion Model (DDM, Ratcliff & McKoon, 2008) which assumes a stochastic accumulation of evidence on each trial. The DDM includes four parameters: A drift rate parameter (v), which describes the rate of (noisy) evidence accumulation; a decision threshold parameter (a), which represents the amount of evidence needed to make a decision; a non-decision time parameter (t), which accounts for the time devoted to sensory processing, motor preparation,

and motor output, and a starting point parameter (z), which accounts for any predispositions in the initial decision variable towards either boundary.

Estimation

The RLDDM is the most advanced method for analyzing performance in the probabilistic reversal learning (PRL) task. The RLDDM was estimated in a hierarchical Bayesian framework using the `HDDMrl` module of the `HDDM` (version 0.9.7) Python package (Fengler et al., 2021; Wiecki et al., 2013). Hierarchical modeling of reinforcement learning tasks has been demonstrated to yield superior predictive accuracy compared to alternative methods (Geen & Gerraty, 2021; Gershman, 2016).

The posterior distribution of group and individual parameters of the RLHDDM model were estimated in a hierarchical Bayesian framework using the "HDDMrl" module of the "HDDM" (version 0.9.8) Python package (for a detailed description of the model, see Pedersen & Frank, 2020; Wiecki et al., 2013).

Priors

The Bayesian posterior estimations of the RLDDM rely on informative priors obtained from a prior meta-analysis (Wiecki et al., 2013) for the DDM aspect of the model. On the other hand, non-informative broad normal distributions, centered at 0.5 after transformation, are used for the learning rate parameters (positive and negative).

Data analysis

Quality control

To maintain the integrity of our dataset, we applied rigorous exclusion criteria prior to analysis. Participants whose accuracy on the PRL task did not exceed the chance performance of 50% were excluded. This decision aligns with benchmarks established in the literature, such as that reported by Geisler et al. (2017). In addition to performance-based exclusions, we also excluded individuals whose data contributed to convergence issues within our Bayesian hierarchical models. These issues were indicated by high $\hat{R}$ values, which suggest a lack of model convergence.

After applying these criteria, our analytical sample was composed of 117 participants. This included 40 individuals with restrictive R-AN. From this R-AN subgroup, 4 participants were excluded: 2 due to their performance not surpassing the chance level and 2 because of convergence issues in the model. The remaining sample consisted of 45 HC and 36 individuals identified as at-risk (RI) for developing eating disorders. Among the at-risk group, 2 participants were excluded based on their chance-level performance on the PRL task.

Models comparison

The study used a minimal model as a starting point for constructing new models. In these new models, all parameters were allowed to vary by condition (*i.e.*, disorder-related vs. disorder-unrelated information) and group. This decision was made because there was no prior knowledge to suggest that the effects of outcome-irrelevant information were limited to a particular parameter of the RLDDM or that they varied across groups. Markov chain Monte Carlo (MCMC) sampling was used to estimate these models. 2000 traces were sampled after a burn-in period of 500. The Deviance Information Criterion (DIC) was calculated for each model. The model with the lowest DIC was selected. The winning model was then re-estimated using 15000 traces and a 5000 burn-in (Kruschke & Liddell, 2018):

```
m8 = hddm.HDDMrl(
    data,
    # bias=True,
    depends_on={
        "a": ["group", "stim"],
        "v": ["group", "stim"],
        "t": ["group", "stim"],
        "alpha": ["group", "stim"],
        "pos_alpha": ["group", "stim"],
    },
    dual=True,
```

```
    p_outlier=0.05,
    informative=True,
    include=["v", "a", "t"],
)

m8.find_starting_values()
m8.sample(15000, burn=5000, dbname="models/ddm8_final.db", db="pickle")
```

Collinearity check

As shown in the following figure, for all three groups the correlation between the parameters is generally low.

Figure 1: Joint posterior distribution of RLDMM parameters.

Posterior predictive checks

To assess model validity, we performed posterior predictive checks. This involved simulating data using estimated parameters and comparing observed and simulated results. We generated the simulated dataset by repeating the simulation process 500 times for each subject in a sample dataset.

PPC for learning rate

To evaluate the choice proportion for the best option across learning in both observed and simulated data, we binned the trials and plotted the 90% highest density intervals of the mean responses. The following figures illustrate the rate of selecting the best option during learning. The 90% highest density interval of the means across simulated datasets captures the uncertainty present in the generated data.

Figure 2: Observed and predicted learning rates across conditions: R-AN group.

Figure 3: Observed and predicted learning rates across conditions: HC group.

Figure 4: Observed and predicted learning rates across conditions: RI group.

PPC for reaction times

The density plots of observed and predicted reaction time across conditions are presented in the following figures. To distinguish upper and lower bound responses, reaction times for lower boundary choices (*i.e.*, worst option choices) were set to be negative (0-RT). There is a good agreement between the observed and predicted values.

Figure 5: Observed and predicted reaction times across conditions: R-AN group.

Figure 6: Observed and predicted reaction times across conditions: HC group.

Figure 7: Observed and predicted reaction times across conditions: RI group.

Additional figures

Within-group comparisons

Figure 8: Comparative plots of the posterior distribution for the Within-group effects of parameter ν ($\nu_{food} - \nu_{neutral}$), across three groups of participants. The upper plot represents the within-group effect for the Restricting-Anorexia Nervosa group, the central plot for the Healthy Controls, and the lower plot for the At-Risk Participants. Each plot illustrates the distribution of effects and highlights the variance in parameter ν within each respective group.

Figure 9: Comparative plots of the posterior distribution for the Within-group effects of parameter t ($t_{food} - t_{neutral}$), across three groups of participants. The upper plot represents the within-group effect for the Restricting-Anorexia Nervosa group, the central plot for the Healthy Controls, and the lower plot for the At-Risk Participants. Each plot illustrates the distribution of effects and highlights the variance in parameter t within each respective group.

Between-group comparisons

Figure 10: Comparative plots of the posterior distribution for the Between-group effects of parameter ν ($\nu_{R-AN} - \nu_{HC}$ and $\nu_{R-AN} - \nu_{RI}$), in two different contexts. The left plot represents the food-related context, while the right plot corresponds to the neutral context, each illustrating the distribution of effects and highlighting the variance in parameter ν between groups.

Figure 11: Comparative plots of the posterior distribution for the Between-group effects of parameter t ($t_{R-AN} - t_{HC}$ and $t_{R-AN} - t_{RI}$), in two different contexts. The left plot represents the food-related context, while the right plot corresponds to the neutral context, each illustrating the distribution of effects and highlighting the variance in parameter t between groups.

Comorbidity

Individuals with eating disorders often have comorbid psychiatric conditions, including depression (up to 75%), bipolar disorder (10%), anxiety disorders, obsessive-compulsive disorder (40%), panic disorder (11%), social anxiety disorder/social phobia, post-traumatic stress disorder (prevalence varies with eating disorder), and substance abuse (15-40%) – see Woodside & Staab (2006).

In this study, we included patients with comorbidities in our sample in order to increase the generalizability of our findings to the broader psychiatric population: 16 individuals with R-AN were diagnosed with comorbid anxiety disorder, 8 with OCD, 1 with social phobia, and 1 with DAP.

Here, we present the results of applying a modified version of model M7 to the R-AN data, where patients were categorized into two groups based on the presence or absence of comorbidities. The potential effects of comorbidities were evaluated using the following model.

```
m = hddm.HDDMrl(
    data,
    depends_on={
        "a": ["comorbidity", "stim"],
        "v": ["comorbidity", "stim"],
        "t": ["comorbidity", "stim"],
        "alpha": ["comorbidity", "stim"],
        "pos_alpha": ["comorbidity", "stim"],
    },
    dual=True,
    p_outlier=0.05,
    informative=True
)
```

The results show no credible differences in the posterior estimates of the model parameters between the group with comorbidities and those without comorbidities. Specifically, for disorder-related choices, the difference in the a parameter was -0.034 (95% CI [-0.188, 0.124]), for v it was 0.230 (95% CI [-1.203, 1.586]), for t it was 0.002 (95% CI [-0.050, 0.055]), for α^- it was 2.614 (95% CI [-3.173, 8.364]), and for α^+ it was -0.635 (95% CI [-4.301, 2.449]). For disorder-unrelated choices, the differences in the respective parameters were -0.126 (95% CI [-0.281, 0.025]) for a, 0.744 (95% CI [-0.452, 1.885]) for v, -0.003 (95% CI [-0.057, 0.052]) for t, -0.768 (95% CI [-6.570, 4.401]) for α^-, and -1.739 (95% CI [-6.184, 1.654]) for α^+.

The 95% credibility intervals of these parameter differences encompass zero, suggesting a lack of credible evidence for any substantial difference in the RLDDM parameters between patients with and without comorbid diagnoses. These findings indicate that the conservative

RL patterns observed for disorder-related choices are not attributable to the presence of comorbidities in individuals with R-AN.

Medication

A similar statistical analysis to the previous section was conducted, categorizing individuals with R-AN into two groups based on the presence or absence of medication.

The results show no credible differences in the posterior estimates of the model parameters between the group with medication and those without medication. Specifically, for disorder-related choices, the difference in the a parameter was -0.074 (95% CI [-0.233, 0.078]), for v it was 0.370 (95% CI [-0.952, 1.701]), for t it was 0.005 (95% CI [-0.048, 0.059]), for α^- it was 1.010 (95% CI [-4.506, 6.468]), and for α^+ it was -0.440 (95% CI [-3.705, 2.547]). For disorder-unrelated choices, the differences in the respective parameters were -0.056 (95% CI [-0.208, 0.100]) for a, 0.837 (95% CI [-0.360, 2.031]) for v, -0.007 (95% CI [-0.066, 0.047]) for t, -1.843 (95% CI [-7.674, 3.351]) for α^-, and -3.346 (95% CI [-7.475, 0.241]) for α^+.

The 95% credibility intervals of these parameter differences encompass zero, suggesting a lack of credible evidence for any substantial difference in the RLDDM parameters between patients with and without medication. These findings indicate that the conservative RL patterns observed for disorder-related choices are not attributable to the presence of medication in individuals with R-AN.

Biased choices

To investigate whether the conservative learning behavior seen in patients with R-AN during the PRL task, when making choices relevant to their disorder, could be due to a general preference for non-food-related images regardless of their past action-outcome history, we conducted a specific analysis. Our focus was on the frequency of food-related choices in the PRL blocks. More precisely, we examined the PRL blocks where participants were presented with a choice between a food image and a neutral image. We asked whether individuals with R-AN tended to choose the food image less frequently compared to the other control groups.

We focused on specific PRL blocks where participants had to select between a food image and a non-food image. We calculated the percentage of times each participant chose the food image. This percentage was then used as the dependent variable in a robust Bayesian regression model.

```
priors <- c(
  set_prior("student_t(4, 0, 1.0)", class = "b")
)

bmod <- brm(
  bf(
    freq ~ diag_cat,
    sigma ~ diag_cat
  ),
  data = bysubj_freq,
  family = student(),
  control = list(adapt_delta = 0.99),
  prior = priors,
  warmup = 1000,
  iter = 5000,
  cores = parallel::detectCores(),
  seed = "12345",
  chains = 4
)
```

In the R-AN group, the proportion of food choices was 0.48, with a 95% credible interval ranging from 0.46 to 0.51. Since the 95% CI includes the value of 0.5, the data do not provide credible evidence of a bias in the R-AN group away from food choices.

The comparisons between the R-AN group and the two control groups are given by the following contrasts.

contrast	estimate	lower.HPD	upper.HPD
R-AN - HC	-0.010	-0.041	0.021
R-AN - RI	0.012	-0.020	0.045

In both cases, the 95% credibility interval is inclusive of 0, indicating no credible difference between the R-AN group and the other two groups in the likelihood of selecting the food image compared to the neutral image.

Outcome-irrelevant learning: spatial-motor associations

Shahar et al. (2019) investigated the impact of spatial-motor associations on participants' RL. Optimal decision-making prioritizes rewards regardless of spatial-motor associations, such as the choice of response key in the previous trial. Shahar et al. (2019) found that rewards had a greater impact on the probability of choosing between two images presented in each trial when the chosen image was linked to the same response key in both the n-1 and n trials.

In order to investigate whether the likelihood of selecting 'stay' was greater for 'same' versus 'flipped' response/key mapping in our data, when contrasting rewarded and unrewarded responses, we reproduced the statistical analyses conducted by Shahar et al. (2019) and Ben-Artzi et al. (2022).

We performed two Bayesian multilevel regression models using the `brm` function from the brms package in R, which utilizes Stan on the back-end. One model included a group effect, while the other did not.

```
priors_0 <- c(
  set_prior("student_t(3, 0, 0.2)", class = "b", coef = "Intercept"),
  set_prior("student_t(3, 0, 0.2)", class = "b"),
  set_prior("student_t(3, 0, 0.2)", class = "sd"),
  set_prior("lkj(1)", class = "cor"),
  set_prior("gamma(0.01, 0.01)", class = "phi"),
  set_prior("beta(2, 2)", class = "coi"),
  set_prior("beta(2, 2)", class = "zoi")
)

mod_0 <- brm(
  stay ~ 0 + Intercept + mapping * feedback * group +
    (1 + mapping * feedback | subj_idx),
  family = zero_one_inflated_beta(),
  backend = "cmdstanr",
  data = bysubj_ed,
  prior = priors_0,
  iter = 2000
)
```

We observed no credible effect of group (HC, RI, R-AN) or a credible interaction between group, previous outcome, and mapping (ELPD difference = -4.5, SE = 2.8). The simpler model used `stay` as the dependent variable (1 if the same response key was chosen in trial i as in trial $i-1$, 0 otherwise), with `mapping` (repeated or flipped image-to-key mapping), `feedback` (reward, punishment), and their interaction as fixed-effects, and `subj_idx` as a random variable. The marginal effects of the `mapping` $\times$ `feedback` interaction ($\beta = 0.93$,

95% CI [0.81, 1.06]) are shown in the following figure.

The figure shows the probability of choosing the same image at trial $n + 1$, depending on the outcome (rewarded vs. unrewarded) and mapping (flipped vs. same). Results indicate a tendency to repeat image selection after a rewarded trial. Notably, a stronger effect of reward was observed when the image was mapped to the same response key, compared to the alternative key. Error bars represent 95% credibility intervals.

In conclusion, our study successfully replicated the findings reported by Shahar et al. (2019) and Ben-Artzi et al. (2022), providing evidence that the effects of reward extend not only to the relevant image but also to the outcome-irrelevant response key. This highlights the impact of outcome-irrelevant factors, such as contextual effects in our study and spatial-motor mapping in the study of Shahar et al. (2019), on reinforcement learning in a PRL task.

References

Ben-Artzi, I., Luria, R., & Shahar, N. (2022). Working memory capacity estimates moderate value learning for outcome-irrelevant features. *Scientific Reports*, *12*(1), 1–10.

Bottesi, G., Ghisi, M., Altoè, G., Conforti, E., Melli, G., & Sica, C. (2015). The Italian version of the Depression Anxiety Stress Scales-21: Factor structure and psychometric properties on community and clinical samples. *Comprehensive Psychiatry*, *60*, 170–181.

Brainard, D. H., & Vision, S. (1997). The psychophysics toolbox. *Spatial Vision*, *10*(4), 433–436.

Dotti, A., & Lazzari, R. (1998). Validation and reliability of the italian EAT-26. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*, *3*(4), 188–194.

Dowson, J., & Henderson, L. (2001). The validity of a short version of the Body Shape Questionnaire. *Psychiatry Research*, *102*(3), 263–271.

Fengler, A., Govindarajan, L. N., Chen, T., & Frank, M. J. (2021). Likelihood approximation networks (LANs) for fast inference of simulation models in cognitive neuroscience. *Elife*, *10*, e65074.

Frost, R. O., Marten, P., Lahart, C., & Rosenblate, R. (1990). The dimensions of perfectionism. *Cognitive Therapy and Research*, *14*(5), 449–468.

Garner, D. M., Olmsted, M. P., Bohr, Y., & Garfinkel, P. E. (1982). The Eating Attitudes Test: Psychometric features and clinical correlates. *Psychological Medicine*, *12*(4), 871–878.

Geen, C. van, & Gerraty, R. T. (2021). Hierarchical bayesian models of reinforcement learning: Introduction and comparison to alternative methods. *Journal of Mathematical Psychology*, *105*, 102602.

Geisler, D., Ritschel, F., King, J. A., Bernardoni, F., Seidel, M., Boehm, I., Runge, F., Goschke, T., Roessner, V., Smolka, M. N., & Ehrlich, S. (2017). Increased anterior cingulate cortex response precedes behavioural adaptation in anorexia nervosa. *Scientific Reports*, *7*(1), 1–10.

Gershman, S. J. (2016). Empirical priors for reinforcement learning models. *Journal of Mathematical Psychology*, *71*, 1–6.

Heimberg, R. G., Mueller, G. P., Holt, C. S., Hope, D. A., & Liebowitz, M. R. (1992). Assessment of anxiety in social interaction and being observed by others: The social interaction anxiety scale and the social phobia scale. *Behavior Therapy*, *23*(1), 53–73.

Kruschke, J. K., & Liddell, T. M. (2018). Bayesian data analysis for newcomers. *Psychonomic Bulletin & Review*, *25*(1), 155–177.

Lombardo, C. (2008). Adattamento italiano della multidimensional perfectionism scale (MPS). *Psicoterapia Cognitiva e Comportamentale*, *14*(3), 31–46.

Lovibond, P. F., & Lovibond, S. H. (1995). The structure of negative emotional states: Comparison of the Depression Anxiety Stress Scales (DASS) with the Beck Depression and Anxiety Inventories. *Behaviour Research and Therapy*, *33*(3), 335–343.

Matera, C., Nerini, A., & Stefanile, C. (2013). The role of peer influence on girls' body dissatisfaction and dieting. *Revue Européenne De Psychologie Appliquée/European Review of Applied Psychology*, *63*(2), 67–74.

Mattick, R. P., & Clarke, J. C. (1998). Development and validation of measures of social phobia scrutiny fear and social interaction anxiety. *Behaviour Research and Therapy*, *36*(4), 455–470.

Pedersen, M. L., & Frank, M. J. (2020). Simultaneous hierarchical bayesian parameter estimation for reinforcement learning and drift diffusion models: A tutorial and links to neural data. *Computational Brain & Behavior*, *3*, 458–471.

Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*(4), 873–922.

Raven, J. (2000). The raven's progressive matrices: Change and stability over culture and time. *Cognitive Psychology*, *41*(1), 1–48.

Rescorla, R. A., & Wagner, A. R. (1972). A theory of pavlovian conditioning: Variations in the effectiveness of reinforcement and nonreinforcement. In A. H. Black & W. F. Prokasy (Eds.), *Classical conditioning II: Current research and theory* (pp. 64–69). Appleton-Century Crofts.

Rosenberg, M. (1965). *Society and the adolescent self-image*. Princeton University Press.

Shahar, N., Moran, R., Hauser, T. U., Kievit, R. A., McNamee, D., Moutoussis, M., Consortium, N., & Dolan, R. J. (2019). Credit assignment to state-independent task representations and its relationship with model-based decision making. *Proceedings of the National Academy of Sciences*, *116*(32), 15871–15876.

Sica, C., Musoni, I., Bisi, B., Lolli, V., & Sighinolfi, C. (2007). Social phobia scale e social interaction anxiety scale: Traduzione e adattamento italiano. *Bollettino Di Psicologia Applicata*, *252*, 59–71.

Stöber, J. (1998). The frost multidimensional perfectionism scale revisited: More perfect with four (instead of six) dimensions. *Personality and Individual Differences*, *24*(4), 481–491.

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction*. MIT Press.

Wiecki, T. V., Sofer, I., & Frank, M. J. (2013). HDDM: Hierarchical bayesian estimation of the drift-diffusion model in python. *Frontiers in Neuroinformatics*, 14.

Woodside, B. D., & Staab, R. (2006). Management of psychiatric comorbidity in anorexia nervosa and bulimia nervosa. *CNS Drugs*, *20*, 655–663.