¹ Contextual influence of reinforcement learning performance in Anorexia Nervosa

² Corrado Caudek[1] & Ernst-August Doelle[1,2]

³ [1] Wilhelm-Wundt-University

⁴ [2] Konstanz Business School

⁵ Author Note

⁶ Add complete departmental affiliations for each author here. Each new line herein

⁷ must be indented, like this line.

⁸ Enter author note here.

⁹ The authors made the following contributions. Corrado Caudek: Conceptualization,

¹⁰ Writing - Original Draft Preparation, Writing - Review & Editing; Ernst-August Doelle:

¹¹ Writing - Review & Editing, Supervision.

¹² Correspondence concerning this article should be addressed to Corrado Caudek, Postal

¹³ address. E-mail: my@email.com

Abstract

Objective: This study utilized a within-subject design to examine whether individuals with restrictive anorexia nervosa (R-AN; $n = 40$) perform similarly to healthy controls (HCs; $n = 45$) and healthy controls at risk of eating disorders (RI; $n = 36$) in a reinforcement learning (RL) tasks. Specifically, we aimed to determine if RL performance is comparable between groups for disorder-unrelated choices, but significantly impaired for disorder-related choices. **Method:** RL performance was assessed using a Probabilistic Reversal Learning (PRL) task, where participants were asked to perform disorder-related choices or disorder-unrelated choices. **Results:** R-AN individuals demonstrated lower learning rates for disorder-related decisions, while their performance on neutral decisions was comparable to participants with Bulimia Nervosa, Healthy Controls (HCs), and HCs at risk of eating disorders. Additionally, only AN patients exhibited reduced learning rates for outcome-irrelevant food-related decisions in reward-based learning, as opposed to food-unrelated decisions. **Discussion:** Impaired RL task performance in individuals with AN may be attributed to external factors rather than compromised learning mechanisms. These findings indicate that AN may significantly impact the cognitive processing of food-related information, even when AN patients do not show learning rate disadvantages compared to HCs in decision-making involving food-unrelated information. This study provides valuable insights into the reinforcement learning processes of individuals with AN and emphasizes the need to consider the influence of food-related information on cognitive functioning in this patient population. The findings have potential implications for the development of interventions targeting decision- making processes in individuals with AN

Keywords: keywords

Word count: X

Contextual influence of reinforcement learning performance in Anorexia Nervosa

Introduction

Anorexia Nervosa (AN) is one of the most common eating disorders characterized by distorted body perception and pathological weight loss, particularly in its restricting type (R-AN) (American Psychiatric Association, 2022). Lifetime prevalence for AN has been reported at 1.4% for women and 0.2% for men (Galmiche, Déchelotte, Lambert, & Tavolacci, 2019; Smink, Hoeken, & Hoek, 2013), with a mortality rate that can be as high as 5-20% (Qian et al., 2022). Treating AN is extremely challenging (Atwood & Friedman, 2020; Linardon, Fairburn, Fitzsimmons-Craft, Wilfley, & Brennan, 2017), highlighting the importance of gaining a deeper understanding of its underlying mechanisms (Chang, Delgadillo, & Waller, 2021).

Executive functions have gained significant attention in the research on understanding the mechanisms underlying Anorexia Nervosa (AN). Impairments in executive processes, such as cognitive inflexibility, decision-making difficulties, and inhibitory control problems, have been identified as potential risk and perpetuating factors in AN (Bartholdy, Dalton, O'Daly, Campbell, & Schmidt, 2016; Guillaume et al., 2015; Wu et al., 2014). Within this domain, Reinforcement Learning (RL) in the context of associative learning has received considerable interest. In fact, the presence of persistent maladaptive eating behaviors in individuals with AN, despite experiencing negative consequences, along with indications of altered reward and punishment sensitivity, has led to the proposal of abnormal reward responsiveness and learning in AN (Schaefer & Steinglass, 2021). While there is strong evidence supporting the presence of anomalies in reward responsiveness in individuals with AN, our current understanding of potential abnormalities in AN-related reward learning remains limited.

In relation to dysfunctions reward responsiveness among individuals with AN, research has revealed that the intense levels of dietary restriction and physical activity characteristic of AN can activate reward pathways (Keating, 2010; Keating, Tilbrook, Rossell, Enticott, &

⁶⁴ Fitzgerald, 2012; Selby & Coniglio, 2020). Additionally, individuals with AN may exhibit

⁶⁵ diminished reward responses specifically towards food (Wierenga et al., 2014). In a broader

⁶⁶ sense, research has shown that AN is associated with reduced subjective reward sensitivity

⁶⁷ and decreased neural response to rewarding stimuli. Moreover, individuals with AN may

⁶⁸ experience disruptions in processing aversive stimuli, leading to heightened harm avoidance,

⁶⁹ intolerance of uncertainty, increased anxiety, and oversensitivity to punishment (Fladung,

⁷⁰ Schulze, Schöll, Bauer, & Groen, 2013; Jappe et al., 2011; Keating et al., 2012; O'Hara,

⁷¹ Campbell, & Schmidt, 2015). These factors contribute to an altered response to negative

⁷² feedback and a tendency to avoid aversive outcomes (Jonker, Glashouwer, & Jong, 2022;

⁷³ Matton, Goossens, Braet, & Vervaet, 2013). Neuroimaging studies have further supported

⁷⁴ these findings by revealing neural dysfunction in AN's response to loss and aversive taste

⁷⁵ (Bischoff-Grethe et al., 2013; Monteleone et al., 2017; Wagner et al., 2007).

⁷⁶   Given the critical role of reinforcement learning (RL) in learning from experience,

⁷⁷ numerous attempts have been made to investigate RL deficits in individuals with AN in

⁷⁸ order to gain insights into the underlying mechanisms responsible for their maladaptive

⁷⁹ eating behaviors (Bischoff-Grethe et al., 2013; Glashouwer, Bloot, Veenstra, Franken, &

⁸⁰ Jong, 2014; Harrison, Genders, Davies, Treasure, & Tchanturia, 2011; Jappe et al., 2011;

⁸¹ Matton et al., 2013). However, results have been mixed. For example, Ritschel et al. (2017)

⁸² reported impaired RL performance in recovered AN patients compared to healthy controls

⁸³ (HCs) using a probabilistic reversal learning (PRL) task, particularly in response to negative

⁸⁴ feedback. In contrast, Bernardoni et al. (2018) observed a higher learning rate from

⁸⁵ punishment in AN patients compared to HCs. Similarly, Sarrar et al. (2015) found no

⁸⁶ differences in task performance between individuals with acute AN and HCs using the

⁸⁷ Probabilistic Object Reversal Task (pORT) with neutral stimuli. Geisler et al. (2018) also

⁸⁸ found no group differences in a PRL task with neutral stimuli and monetary feedback.

⁸⁹   To clarify RL deficits in AN, some studies incorporated food-related feedback into the

PRL paradigm. For example, Zang et al. (2014) demonstrated that patients with binge-eating disorder performed worse when exposed to food-related feedback, indicating vulnerability to food-related cues. However, attempts to replicate these findings in AN have produced conflicting results. Hildebrandt et al. (2015) reported increased rigidity in AN individuals using a Probabilistic Reversal Learning (PRL) task with food-related feedback, while Hildebrandt et al. (2018) found no differences in PRL performance between AN patients and HCs using the same paradigm.

Here, we speculate that inconsistent findings in behavioral experiments regarding RL in individuals with AN may be partially due to the predominant use of general stimuli instead of disorder-relevant stimuli (Schaefer & Steinglass, 2021). Furthermore, when disorder-related information is incorporated, it is typically limited to the feedback provided after the participant's choice, while the stimuli during the decision-making process are unrelated to the disorder. This manipulation primarily highlights the consequences of the choices rather than the contextual factors surrounding the decision-making process. This methodological choice of previous studies stems from the assumption that any potential RL impairment in AN can be attributed to deficits in a singular, underlying RL mechanism [ref].

However, an alternative perspective proposes that contextual factors, such as individual characteristics, long-term goals, and situational influences, can have a detrimental impact on RL performance, even when the underlying RL mechanisms remain intact (Haynos, Widge, Anderson, & Redish, 2022). In other words, it suggests that the deficits in RL performance observed in R-AN may be the result of contextual factors influencing the learning process, rather than a fundamental impairment in the RL mechanism itself.

To examine the proposed hypothesis, we conducted a study utilizing a modified version of the standard PRL task. In contrast to previous studies that utilized general stimuli (Schaefer & Steinglass, 2021), we incorporated two distinct contexts in our task. One context involved choices between a stimulus related to the disorder (e.g., a caloric food) and a

stimulus unrelated to the disorder (e.g., a lamp), while the other context involved choices

between two stimuli unrelated to the disorder (i.e., flower vs objects).

Considering the heightened susceptibility of individuals with R-AN to symptom-related

information, such as food, body weight, and social pressure [ref], we anticipated that their

learning abilities would diminish when making decisions between a food item and a non-food

item. Conversely, we expected no learning anomaly in individuals with R-AN when making

choices between two non-food items in a PRL task.

The hypothesis of contextuallearning in individuals with R-AN has significant

implications for treatment strategies. Current efforts aim to improve cognitive flexibility in

individuals with R-AN to address their maladaptive eating behavior. However, existing

interventions have predominantly concentrated on fostering adaptive behavioral choices in

contexts unrelated to the disorder (e.g., Tchanturia, Davies, Reeder, & Wykes, 2010).

Establishing evidence that maladaptive RL is context-dependent would necessitate a

redirection of intervention approaches.

Methods

The study, which adhered to the Declaration of Helsinki, was approved by the

University of Florence's Ethical Committee (Prot. n. 0178082). All eligible participants

provided informed consent and willingly agreed to participate in the study.

Participants

We tested the hypothesized learning asymmetry in PRL performance across three

groups: individuals diagnosed with DSM-5 restricting-type Anorexia Nervosa (AN), a control

group comprising of healthy individuals (HCs), and individuals who are at risk of developing

eating disorders (RIs). Demographic and clinical characteristics of the sample are shown in

Table 1. Details about participant selection, inclusion criteria, and sample characteristics are

given in the Supplementary Information (SI).

Procedure

During the initial session, participants underwent a clinical interview to determine their eligibility for the study. Those who met the criteria and were selected proceeded to anthropometric measurements and completed a battery of psychometric scales. In a subsequent session, participants completed the PRL task. Participants were told they were going to play a simple computer game with the objective of accumulating as many "virtual euro" as possible. During the PRL task, participants were presented with two stimuli simultaneously on a screen and were instructed to select one within a 2.5-second time limit by pressing a key. Trials were presented in an interleaved manner, with a randomly drawn inter-trial interval ranging from 0.5 to 1.5 seconds. Following each trial, a euro coin image was displayed as a reward for correct responses, while a strike-through image of a euro coin served as a punishment for incorrect responses. Feedback was provided for 2 seconds after each trial.

The PRL task consisted of two blocks, each containing 160 trials. One block included pairs of food-related and food-unrelated images, while the other block exclusively used food-unrelated images. The images were selected randomly from sets of food-related and food-unrelated categories.

All images used in the study were obtained from the International Affective Picture System (IAPS) database (Lang et al., 2005). 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). (For further details, see the SI).

Data analysis

To analyze the temporal dynamics of the two-choice decision-making in the PRL task, we employed a hierarchical reinforcement learning drift diffusion model (RLDDM), as described in Pedersen, Frank, and Biele (2017) and Pedersen and Frank (2020). Cognitive modeling analysis allows us to deconstruct decision-making task performance into its component processes. This approach enables the identification of deviations in the underlying mechanisms that may not be evident in the overall task outcome. The RLDDM consists of two key components: one describes how reward feedback is employed to update value expectations [delta learning rule; Rescorla and Wagner (1972)] and the other describes how an agent uses these expectations to arrive at a decision. In the RLDDM the traditional softmax function is replaced by Drift-Diffusion Model [DDM; Ratcliff and McKoon (2008)] which assumes a stochastic accumulation of evidence on each trial (see SI).

The RLDDM has six basic parameters: α^+, α^-, v, a, t, and z. The α parameters quantify the learning rate in the Rescorla-Wagner delta learning rule (Rescorla & Wagner, 1972); a higher learning rate results in rapid adaptation to reward expectations, while a lower learning rate results in slow adaptation. The parameter α^+ is computed from reinforcements, whereas α^+ is computed from punishments. The drift rate v is the average speed of evidence accumulation toward one decision. The decision boundary is the distance between two decision thresholds; an increase of a increases the evidence needed to make a decision. The increase of a leads to a slower but more accurate decision; a decrease in a results in a faster but error-prone decision. The non-decision time t is the time spent for stimuli encoding or motor execution (*i.e.*, time not used for evidence accumulation). The starting point parameter z captures a potential initial bias toward one or the other boundary in absence of any stimulus evidence.

To assess the presence of context-dependent learning, we conditioned the model's parameters on two specific contexts: disorder-related choices and disorder-unrelated choices.

192 This allowed us to examine how the model's parameters varied in response to these different

193 contextual conditions.

<div align="center">

Results

</div>

Models selection

196   To evaluate context-dependent learning, we compared several RLDDM models that

197 varied in their conditioning of the model's parameters on the group (R-AN, HC, RI) and

198 context (disorder-related choices and disorder-unrelated choices). We used the Deviance

199 Information Criterion (DIC) to balance model fit and complexity, selecting the model with

200 the lowest DIC as the best trade-off. The following RLDDM models were examined:

1. Model M1: Standard RLDDM without conditioning. DIC = 39879.444.

2. Model M2: Separate learning rates for positive and negative reinforcements. DIC = 39124.890

3. Model M3: Group-based α^+ and α^- parameters. DIC = 39194.763.

4. Model M4: Group and context-based α^+ and α^- parameters. DIC = 38197.467.

5. Model M5: Group and context-based α^+, α^-, and a parameters. DIC = 36427.448.

6. Model M6: Group and context-based α^+, α^-, a, and drift rate (v) parameters. DIC = 36185.146.

7. Model M7: Group and context-based α^+, α^-, a, v, and non-decision time (t) parameters. DIC = 34904.053.

8. Model M8: Group and context-based α^+, α^-, a, v, t, and starting point (z) parameters. DIC = 34917.762.

213   Among the evaluated models, Model M7 had the lowest DIC, indicating the best

214 trade-off between goodness of fit and model complexity. In Model M7, the parameters α^+,

215 α^-, a, v, and t (excluding z) were conditioned on both the group and the context.

Modelling results

Model M7 was estimated using 15,000 iterations, with a burn-in period of 5,000 iterations. Convergence of the Bayesian estimation was evaluated using the Gelman-Rubin statistic. For all parameters in Model M7, the $\hat{R}$ values were below 1.1 (maximum = 1.025, mean = 1.002), indicating no significant convergence issues. Collinearity and posterior predictive checks were also used to evaluate model validity (see SI).

To investigate the impact of disorder-related versus disorder-unrelated information on RL learning, we compared the posterior estimates of the RLDDM parameters of M7 between the two conditions (see Table 1).

Table 1
Posterior Parameter Estimates of DDMRL Model M7 by Group (R-AN, HC, RI) and Context of PRL Choice (disorder-related vs. disorder-unrelated information). The learning rates (α) are shown on a logit scale. The probability (p) describes the Bayesian test that the posterior estimate of the parameter in the disorder-related context is greater than the posterior estimate of the parameter in the disorder-unrelated context. Standard deviations are provided in parentheses.

Group	Par.	Neutral choice	Food choice	p	Cohen's d
R-AN	a	1.273 (0.039)	1.442 (0.040)	0.0013	0.802
R-AN	v	1.403 (0.320)	1.776 (0.342)	0.7907	0.190
R-AN	t	0.188 (0.011)	0.174 (0.011)	0.8311	-0.253
R-AN	α^-	1.815 (1.081)	0.738 (1.096)	0.2349	-0.432
R-AN	α^+	1.006 (0.899)	-1.786 (0.756)	0.0098	-1.206
HC	a	1.222 (0.033)	1.314 (0.034)	0.0256	0.474
HC	v	2.157 (0.265)	1.790 (0.263)	0.1606	-0.358
HC	t	0.183 (0.009)	0.172 (0.009)	0.8228	-0.280
HC	α^-	2.780 (0.874)	3.442 (0.980)	0.6993	0.298
HC	α^+	1.198 (0.680)	1.326 (0.700)	0.5544	0.071
RI	a	1.245 (0.041)	1.316 (0.039)	0.1026	0.403

Group	Par.	Neutral choice	Food choice	p	Cohen's d
RI	v	2.197 (0.322)	1.849 (0.307)	0.2133	-0.381
RI	t	0.188 (0.011)	0.186 (0.011)	0.5462	0.166
RI	α^-	2.857 (1.067)	2.904 (1.062)	0.5101	0.015
RI	α^+	1.573 (0.847)	0.739 (0.752)	0.2247	-0.438

Let consider first the evidence of context-dependent learning from within-group comparisons. We found that, on average, individuals in the R-AN group demonstrate a reduced learning rate in response to positive prediction errors (PEs) for disorder-related choices, as compared to disorder-unrelated choices (Cohen's $d = 1.206$, $p = 0.0098$). In contrast, no substantial evidence was found indicating a difference in the learning rate between disorder-related and disorder-unrelated choices in the HC ($p = 0.5544$) and RI ($p = 0.2247$) groups. We found no credible difference in the learning rate from negative prediction errors between disorder-related and disorder-unrelated choices for any of the R-AN ($p = 0.2349$), HC ($p = 0.6993$), and RI ($p = 0.5101$) groups. Moreover, we found that both the R-AN (Cohen's $d = 0.802$, $p = 0.0013$) and HC (Cohen's $d = 0.474$, $p = 0.0256$) groups showed a higher decision threshold for disorder-related choices compared to disorder-unrelated choices.

Further evidence of context-dependent learning emerges from between-groups comparisons. When making disorder-related choices, individuals with R-AN displayed a decreased learning rate following positive prediction errors (PEs) compared to both HC and RI. Specifically, the learning rate after positive PEs was lower for R-AN compared to HC, $p = 0.0009$, Cohen's $d = 1.498$. Similarly, R-AN exhibited a lower learning rate after positive PEs compared to RI ($p = 0.0085$, Cohen's $d = 1.209$). In contrast, no credible difference in the learning rate after positive PEs was found between R-AN and HC ($p = 0.4325$), as well as between R-AN and RI ($p = 0.3232$), for choices unrelated to disorder information.

245 Concerning the learning rate after negative PEs, we found that R-AN showed a lower

246 learning rate compared to HC, but only for disorder-related choices: ($p = 0.0274$, Cohen's d

247 $= 1.144$). Individuals with R-AN showed a higher decision threshold for disorder-related

248 choices compared to both HC (Cohen's $d = 0.622$, $p = 0.0068$) and RI (Cohen's $d = 0.454$, p

249 $= 0.0118$) participants. No credible group differences were found for disorder-unrelated

250 choices. Additionally, we observed that both HC (Cohen's $d = 0.520$, $p = 0.0344$) and RI

251 (Cohen's $d = 0.529$, $p = 0.0392$) participants exhibited a faster accumulation of evidence and

252 more confident decision-making, as indicated by a higher average drift rate parameter,

253 compared to individuals with R-AN. This difference was only evident for disorder-unrelated

254 choices. Finally, no credible differences were found, for both within-group and between-group

255 comparisons, regarding the non-decision time parameter (t).

Preferential choices

257   To investigate the presence of a bias against food choices in individuals with R-AN

258 during the PRL task, regardless of their past action-outcome history, we analyzed the

259 frequency of food choices in PRL blocks where a food image was paired with a neutral image.

260 Our results show that the AN-R group did not exhibit a bias against the food image, with a

261 proportion of food choices estimated at 0.49, 95% CI [0.46, 0.51]. Furthermore, there were no

262 credible differences in food choices between the R-AN group and the HC group (contrast

263 R-AN - HC = -0.007, 95% CI [-0.037, 0.024]) or between the R-AN group and the RI group

264 (contrast R-AN - RI = 0.013, 95% CI [-0.019, 0.046]).

Comorbidity

266   We conducted a further statistical analysis to investigate whether the conservative

267 learning behavior observed in individuals with R-AN could be explained by comorbid

268 conditions. Using model M7, we categorized individuals with R-AN based on the presence or

269 absence of diagnosed comorbidities. Our analysis revealed no credible differences in

270 parameters between the two groups. Specifically, for the disorder-related context, the

271 parameter differences were as follows: $\Delta\alpha^- = 2.614$, 95% CI [-3.173, 8.364]; $\Delta\alpha^+ = $ -0.635,

272 95% CI [-4.301, 2.449]; $\Delta a = $ -0.034, 95% CI [-0.188, 0.124]; $\Delta v = 0.230$, 95% CI [-1.203,

273 1.586]; $\Delta t = 0.002$, 95% CI [-0.050, 0.055]. For the disorder-unrelated context, the parameter

274 differences were: $\Delta\alpha^- = $ -0.768, 95% CI [-6.570, 4.401]; $\Delta\alpha^+ = $ -1.739, 95% CI [-6.184,

275 1.654]; $\Delta a = $ -0.126, 95% CI [-0.281, 0.025]; $\Delta v = 0.744$, 95% CI [-0.453, 1.886]; $\Delta t = $ -0.003,

276 95% CI [-0.057, 0.052].

Discussion

278   Our findings reveal a context-dependent learning asymmetry in individuals with R-AN

279 specifically in the positive learning rate. This asymmetry is observed when comparing the

280 performance in the PRL task for disorder-related choices versus disorder-unrelated choices.

281 Importantly, no similar difference is found in the two control groups.

282   The presence of context-dependent learning asymmetry is also supported by group

283 comparisons. Individuals with R-AN exhibited lower learning rates for both positive and

284 negative prediction errors compared to the HC group, and specifically for positive prediction

285 errors compared to the RI group, but these differences were observed only for

286 disorder-related choices. In contrast, no substantial difference in learning rates was found

287 between the R-AN group and the HC and RI groups for disorder-unrelated choices.

288   Support for context-dependent learning in R-AN is also provided by the DDM

289 parameters of the hDDMrl model. Specifically, we observed that the R-AN group exhibited a

290 higher decision threshold (parameter "a" in the hDDMrl model) compared to the HC and RI

291 groups, but this difference was only evident in the context of disorder-related choices. This

292 suggests that individuals with R-AN displayed a more cautious or conservative

293 decision-making behavior specifically in relation to disorder-related choices (see also Caudek,

294 Sica, Cerea, Colpizzi, & Stendardi, 2021; Schiff, Testa, Rusconi, Angeli, & Mapelli, 2021).

295   Further support of context-related learning in R-AN comes from the result which

296 indicate that both healthy control (HC) and at-risk (RI) participants exhibited a faster

297 accumulation of evidence and displayed more confident decision-making, as reflected by a

298 higher average drift rate parameter, compared to individuals with restrictive anorexia

299 nervosa (R-AN). However, this difference was specifically observed for disorder-unrelated

300 choices. It is noteworthy that individuals with R-AN displayed slower evidence accumulation

301 and less confident decision-making specifically in disorder-unrelated contexts, whereas this

302 group difference was not observed for disorder-related choices. This finding further supports

303 the notion of context-dependent learning in individuals with R-AN, particularly in the

304 context of food-related information.

305   Further evidence of context-related learning in R-AN comes from the analysis of the

306 drift rate parameter. Individuals with R-AN exhibited slower evidence accumulation and less

307 confident decision-making compared to the control groups, specifically in the context of

308 disorder-unrelated choices. Conversely, no credible group differences were observed for

309 food-related choices. These results suggest that individuals with R-AN may allocate greater

310 cognitive resources to process salient information in the disorder-related context, which leads

311 to similar evidence accumulation rates in decision-making compared to the control groups.

312 In contrast, they exhibit a slower evidence accumulation rate when faced with less salient

313 disorder-unrelated choices.

314   The analysis of preferential choices supports the conclusion that the learning

315 performance asymmetry observed in individuals with R-AN is not due to a preferential

316 selection of the disorder-unrelated image during the learning task. Additionally, our analysis

317 examining the relationship between the model's parameters and the presence of

318 comorbidities indicates that the learning performance asymmetry in individuals with R-AN

319 cannot be attributed to comorbid conditions.

Geneeral discussion

In this study, we investigated reinforcement learning using a behavioral paradigm that included two distinct learning contexts: one involving choices related to food and the other involving choices unrelated to food. We compared the performance of patients with R-AN to age-, gender-, and education-matched healthy controls, as well as healthy controls at-risk of developing eating disorders. Consistent with our hypotheses, our findings revealed a reduced learning rate in the disorder-related context for patients with R-AN, whereas both healthy participants and at-risk individuals learned equally well in both contexts.

In PRL tasks, a participant's performance can be influenced by two potential factors. First, there may be a learning impairment, where participants struggle to accurately update the value of the stimuli. Second, there may be a decision impairment, where participants may still select the wrong stimulus despite having intact learning processes. Our results provide evidence that individuals with R-AN may struggle with both accurately updating the value of disorder-related stimuli and making appropriate decisions based on this information. However, we did not observe similar impairments in decision making for disorder-unrelated choices. These findings provide evidence for context-dependent learning in individuals with R-AN, where the inclusion of disorder-related information negatively impacts their RL performance. It is important to note that this effect is specific to the disorder-related context and does not suggest a generalized RL deficit in individuals with R-AN. Thus, our results challenge the notion of a domain-general RL mechanism impairment in this population.

Previous studies have demonstrated that reward and punishment processing in individuals with AN is influenced by stimulus properties and contextual factors (Haynos, Lavender, Nelson, Crow, & Peterson, 2020). For instance, predictable and controllable behaviors such as calorie counting or purging are often perceived as rewarding, providing individuals with a sense of control and accomplishment. Conversely, unpredictable and uncontrollable situations, such as social outcomes, can be perceived as punishing, leading to

346 heightened anxiety and distress. While these previous studies have primarily focused on the

347 effect of context on the subjective value assigned to experiences in AN, our study extends

348 the investigation by examining the impact of context on the learning process itself. We find

349 that the context also affects the learning process itself (not only the subjective value),

350 providing further insights into reward and punishment processing in AN.

351   Other recent studies have focused on investigating context-specific learning in eating

352 disorders. One task specifically designed for this purpose is the two-step Markov decision

353 task, which distinguishes between automatic or habitual (model-free) learning and controlled

354 or goal-directed (model-based) learning. For instance, studies conducted by Foerde et al.

355 (2021) and Onysk and Seriès (2022) employed similar experiments using the two-step task

356 paradigm. Foerde et al. (2021) compared a monetary two-step task and a food-related

357 two-step task, while Onysk and Seriès (2022) utilized stimuli unrelated to food or body

358 images (i.e., pirate ships and treasure chests) with rewards associated with body image

359 dissatisfaction. The results of these studies consistently demonstrated that individuals with

360 AN tend to exhibit a stronger inclination towards habitual control over goal-directed control

361 across different domains compared to healthy controls. However, no significant differences

362 were observed in learning rates as a function of context, nor between AN patients and

363 healthy controls, according to these findings. In contrast, the present study reveals that, in

364 individuals with R-AN, the learning process *per se* can be influenced by contextual

365 (disorder-related) information, even when such information is not directly relevant to the

366 task outcome.

367   The implications of the present findings, if replicated by future studies, are relevant to

368 the treatment of AN. Current treatment practices address the issue of cognitive inflexibility

369 in AN, with Cognitive Remediation Therapy (CRT) being proposed as an adjunct treatment

370 targeting specific cognitive processes in AN and other eating disorders. CRT involves

371 cognitive exercises and behavioral interventions aimed at improving central coherence

372 abilities, reducing cognitive and behavioral inflexibility, and enhancing thinking style

373 comprehension (Tchanturia et al., 2010). A key aspect of CRT is to avoid focusing on

374 symptom-related themes and instead utilize neutral stimuli in cognitive and behavioral

375 exercises. This approach aims to establish a therapeutic alliance and reduce drop-out rates,

376 particularly among AN patients.

377   However, recent evidence suggests that CRT may not consistently enhance central

378 coherence abilities, cognitive flexibility, or symptoms associated with eating disorders

379 (Hagan, Christensen, & Forbush, 2020; Tchanturia, Giombini, Leppanen, & Kinnaird, 2017).

380 In response to this, Trapp, Heid, Röder, Wimmer, and Hajak (2022) have proposed

381 improvements to address practical issues encountered in the application of CRT. They

382 question the use of neutral stimuli and draw support from Beck's cognitive theory of

383 depression (Beck & Alford, 2009).

384   The proposition presented by Trapp et al. (2022) aligns with the hypothesis of our

385 study, which suggests that contextual factors play a crucial role in the maladaptive eating

386 behavior observed in individuals with R-AN, extending beyond deficits in the underlying

387 reinforcement learning mechanism alone. If abnormal reward learning is indeed identified as

388 a significant anomaly among individuals with R-AN, particularly in relation to

389 disorder-relevant choices, it would imply that treatments focused on enhancing cognitive

390 flexibility and reinforcement learning processes specific to disorder-relevant stimuli could

391 hold significant promise for this population.

392   There are few important limitations and questions for future research. 1) One aspect

393 to consider is the use of symbolic rewards and punishments in our study, represented by

394 images of a one euro coin and a barred representation of a one euro coin, respectively. These

395 rewards and punishments were merely symbolic, and it is unclear how the use of concrete,

396 non-symbolic rewards and punishments would impact the findings. Additionally, the

397 subjective value of one euro, or the loss of one euro, may vary among participants. Therefore,

future studies could aim to determine the equivalence of subjective values for rewards and punishments to enhance the understanding of the underlying processes. 2) Our study only included individuals with R-AN who were not in the most severe stage of the illness, as they were recruited from a center for voluntary medical and psychological support. We did not examine R-AN patients who require hospitalization due to the life-threatening nature of their illness. It is possible that at the later stages of the illness, associative learning abilities, which were preserved in the present sample under neutral conditions, may become impaired. Therefore, investigating the impact of illness severity on context-dependent learning in R-AN patients is an important avenue for future research. 3) While we observed no difference in the choice behavior of R-AN patients, as measured by the relative frequency of image choices, when selecting between a neutral image and a food image, we did find a slower learning rate and lower decision threshold for R-AN patients compared to healthy controls in the RLDDM model when compared to choosing between two neutral images. It is possible that the higher "salience" of food images compared to neutral images could be better captured by other measures, such as fixation length or the number of fixations, rather than solely relying on the relative frequency of image choices. This warrants further exploration in future studies. 4) It is worth noting that our study excluded women under the age of 18. However, this age range is a critical period as the onset of AN during this stage may have a more profound impact on associative learning, given the ongoing cognitive development and less-developed protective factors. Therefore, future studies should take into consideration the inclusion of participants in this age range to better understand the influence of context-dependent learning in R-AN.

References

American Psychiatric Association. (2022). *Diagnostic and Statistical Manual of Mental Disorders* (5th ed., Text Revision). Arlington, VA: American Psychiatric Publishing.

Atwood, M. E., & Friedman, A. (2020). A systematic review of enhanced cognitive behavioral therapy (CBT-e) for eating disorders. *International Journal of Eating Disorders*, *53*(3), 311–330.

Bartholdy, S., Dalton, B., O'Daly, O. G., Campbell, I. C., & Schmidt, U. (2016). A systematic review of the relationship between eating, weight and inhibitory control using the stop signal task. *Neuroscience & Biobehavioral Reviews*, *64*, 35–62.

Beck, A. T., & Alford, B. A. (2009). *Depression: Causes and treatment.* University of Pennsylvania Press.

Bischoff-Grethe, A., McCurdy, D., Grenesko-Stevens, E., Irvine, L. E. Z., Wagner, A., Yau, W.-Y. W., et al.others. (2013). Altered brain response to reward and punishment in adolescents with anorexia nervosa. *Psychiatry Research: Neuroimaging*, *214*(3), 331–340.

Caudek, C., Sica, C., Cerea, S., Colpizzi, I., & Stendardi, D. (2021). Susceptibility to eating disorders is associated with cognitive inflexibility in female university students. *Journal of Behavioral and Cognitive Therapy*, *31*(4), 317–328.

Chang, P. G., Delgadillo, J., & Waller, G. (2021). Early response to psychological treatment for eating disorders: A systematic review and meta-analysis. *Clinical Psychology Review*, *86*, 102032.

Fladung, A.-K., Schulze, U. M., Schöll, F., Bauer, K., & Groen, G. (2013). Role of the ventral striatum in developing anorexia nervosa. *Translational Psychiatry*, *3*(10), e315–e315.

Foerde, K., Daw, N. D., Rufin, T., Walsh, B. T., Shohamy, D., & Steinglass, J. E. (2021). Deficient goal-directed control in a population characterized by extreme goal pursuit. *Journal of Cognitive Neuroscience*, *33*(3), 463–481.

Galmiche, M., Déchelotte, P., Lambert, G., & Tavolacci, M. P. (2019). Prevalence of eating

disorders over the 2000–2018 period: A systematic literature review. *The American Journal of Clinical Nutrition*, *109*(5), 1402–1413.

Glashouwer, K. A., Bloot, L., Veenstra, E. M., Franken, I. H., & Jong, P. J. de. (2014). Heightened sensitivity to punishment and reward in anorexia nervosa. *Appetite*, *75*, 97–102.

Guillaume, S., Gorwood, P., Jollant, F., Van den Eynde, F., Courtet, P., & Richard-Devantoy, S. (2015). Impaired decision-making in symptomatic anorexia and bulimia nervosa patients: A meta-analysis. *Psychological Medicine*, *45*(16), 3377–3391.

Hagan, K. E., Christensen, K. A., & Forbush, K. T. (2020). A preliminary systematic review and meta-analysis of randomized-controlled trials of cognitive remediation therapy for anorexia nervosa. *Eating Behaviors*, *37*, 101391.

Harrison, A., Genders, R., Davies, H., Treasure, J., & Tchanturia, K. (2011). Experimental measurement of the regulation of anger and aggression in women with anorexia nervosa. *Clinical Psychology & Psychotherapy*, *18*(6), 445–452.

Haynos, A. F., Lavender, J. M., Nelson, J., Crow, S. J., & Peterson, C. B. (2020). Moving towards specificity: A systematic review of cue features associated with reward and punishment in anorexia nervosa. *Clinical Psychology Review*, *79*, 101872.

Haynos, A. F., Widge, A. S., Anderson, L. M., & Redish, A. D. (2022). Beyond description and deficits: How computational psychiatry can enhance an understanding of decision-making in anorexia nervosa. *Current Psychiatry Reports*, 1–11.

Jappe, L. M., Frank, G. K., Shott, M. E., Rollin, M. D., Pryor, T., Hagman, J. O., . . . Davis, E. (2011). Heightened sensitivity to reward and punishment in anorexia nervosa. *International Journal of Eating Disorders*, *44*(4), 317–324.

Jonker, N. C., Glashouwer, K. A., & Jong, P. J. de. (2022). Punishment sensitivity and the persistence of anorexia nervosa: High punishment sensitivity is related to a less favorable course of anorexia nervosa. *International Journal of Eating Disorders*, *55*(5), 697–702.

Keating, C. (2010). Theoretical perspective on anorexia nervosa: The conflict of reward.

Neuroscience & Biobehavioral Reviews, *34*(1), 73–79.

Keating, C., Tilbrook, A. J., Rossell, S. L., Enticott, P. G., & Fitzgerald, P. B. (2012). Reward processing in anorexia nervosa. *Neuropsychologia*, *50*(5), 567–575.

Linardon, J., Fairburn, C. G., Fitzsimmons-Craft, E. E., Wilfley, D. E., & Brennan, L. (2017). The empirical status of the third-wave behaviour therapies for the treatment of eating disorders: A systematic review. *Clinical Psychology Review*, *58*, 125–140.

Matton, A., Goossens, L., Braet, C., & Vervaet, M. (2013). Punishment and reward sensitivity: Are naturally occurring clusters in these traits related to eating and weight problems in adolescents? *European Eating Disorders Review*, *21*(3), 184–194.

Monteleone, A. M., Monteleone, P., Esposito, F., Prinster, A., Volpe, U., Cantone, E., et al.others. (2017). Altered processing of rewarding and aversive basic taste stimuli in symptomatic women with anorexia nervosa and bulimia nervosa: An fMRI study. *Journal of Psychiatric Research*, *90*, 94–101.

O'Hara, C. B., Campbell, I. C., & Schmidt, U. (2015). A reward-centred model of anorexia nervosa: A focussed narrative review of the neurological and psychophysiological literature. *Neuroscience & Biobehavioral Reviews*, *52*, 131–152.

Onysk, J., & Seriès, P. (2022). The effect of body image dissatisfaction on goal-directed decision making in a population marked by negative appearance beliefs and disordered eating. *Plos One*, *17*(11), e0276750.

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.

Pedersen, M. L., Frank, M. J., & Biele, G. (2017). The drift diffusion model as the choice rule in reinforcement learning. *Psychonomic Bulletin & Review*, *24*, 1234–1251.

Qian, J., Wu, Y., Liu, F., Zhu, Y., Jin, H., Zhang, H., ... Yu, D. (2022). An update on the prevalence of eating disorders in the general population: A systematic review and meta-analysis. *Eating and Weight Disorders-Studies on Anorexia, Bulimia and Obesity*,

27(2), 415–428.

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

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). New York, NY: Appleton-Century Crofts.

Schaefer, L. M., & Steinglass, J. E. (2021). Reward learning through the lens of RDoC: A review of theory, assessment, and empirical findings in the eating disorders. *Current Psychiatry Reports*, *23*, 1–11.

Schiff, S., Testa, G., Rusconi, M. L., Angeli, P., & Mapelli, D. (2021). Expectancy to eat modulates cognitive control and attention toward irrelevant food and non-food images in healthy starving individuals. A behavioral study. *Frontiers in Psychology*, *11*, 3902.

Selby, E. A., & Coniglio, K. A. (2020). Positive emotion and motivational dynamics in anorexia nervosa: A positive emotion amplification model (PE-AMP). *Psychological Review*, *127*(5), 853–890.

Smink, F. R., Hoeken, D. van, & Hoek, H. W. (2013). Epidemiology, course, and outcome of eating disorders. *Current Opinion in Psychiatry*, *26*(6), 543–548.

Tchanturia, K., Davies, H., Reeder, C., & Wykes, T. (2010). *Cognitive remediation therapy for anorexia nervosa*. London: King's College London.

Tchanturia, K., Giombini, L., Leppanen, J., & Kinnaird, E. (2017). Evidence for cognitive remediation therapy in young people with anorexia nervosa: Systematic review and meta-analysis of the literature. *European Eating Disorders Review*, *25*(4), 227–236.

Trapp, W., Heid, A., Röder, S., Wimmer, F., & Hajak, G. (2022). Cognitive remediation in psychiatric disorders: State of the evidence, future perspectives, and some bold ideas. *Brain Sciences*, *12*(6), 683.

Wagner, A., Aizenstein, H., Venkatraman, V. K., Fudge, J., May, J. C., Mazurkewicz, L., et

527   al.others. (2007). Altered reward processing in women recovered from anorexia nervosa.

528   *American Journal of Psychiatry*, *164*(12), 1842–1849.

529 Wierenga, C. E., Ely, A., Bischoff-Grethe, A., Bailer, U. F., Simmons, A. N., & Kaye, W. H.

530   (2014). Are extremes of consumption in eating disorders related to an altered balance

531   between reward and inhibition? *Frontiers in Behavioral Neuroscience*, *8*, 410.

532 Wu, M., Brockmeyer, T., Hartmann, M., Skunde, M., Herzog, W., & Friederich, H.-C. (2014).

533   Set-shifting ability across the spectrum of eating disorders and in overweight and obesity:

534   A systematic review and meta-analysis. *Psychological Medicine*, *44*(16), 3365–3385.