¹ Impact of Contextual Learning on 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

Abstract

Objective: This study compared individuals with Restrictive Anorexia Nervosa (R-AN; n = 40) to Healthy Controls (HCs; n = 45) and healthy controls at RIsk of eating disorders (RI; n = 36) in a Probabilistic Reversal Learning (PRL) task. The aim was to investigate whether R-AN individuals perform similarly to HCs and RIs in neutral contexts but show significant impairments in food-related contexts. **Method:** Participants completed a PRL task, making choices related to their disorder or unrelated to it. **Results:** R-AN individuals showed lower learning rates for disorder-related decisions, but their performance on neutral decisions was similar to the HC and RI groups. Additionally, only the R-AN patients exhibited reduced learning rates for food-related decisions compared to food-unrelated decisions. **Discussion:** These findings suggest that contextual cues, like food images, negatively impact Reinforcement Learning (RL) processes in individuals with R-AN. This raises questions about whether the impaired RL performance should be solely attributed to compromised learning mechanisms, especially when RL abilities appear intact in neutral contexts. The study's insights may have implications for developing interventions that target decision-making processes in individuals with R-AN.

Keywords: anorexia nervosa, reinforcement learning, contextual learning, probabilistic reversal learning

Word count: X

Impact of Contextual Learning on 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 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). In particular, cognitive inflexibility in AN refers to difficulties adapting thoughts and actions to new information. This rigidity can lead to repetitive behaviors. On the other hand, reinforcement learning is the brain's way of optimizing decision-making based on action outcomes, where positive outcomes encourage repetition and negative outcomes discourage it. In AN, cognitive inflexibility can hinder updating of action-outcome associations, affecting adaptive behavior learning and leading to impaired reinforcement learning processes. This may contribute to difficulties recognizing the negative consequences of restrictive eating behaviors, perpetuating the disorder. While studies have explored reward processing abnormalities, our understanding of learning-related anomalies in AN remains limited (Schaefer & Steinglass, 2021).

In relation to dysfunctions reward responsiveness among individuals with AN, research

59 has revealed that the intense levels of dietary restriction and physical activity typically

60 associated with AN can trigger the activation of reward pathways in the brain (Keating,

61 2010; Keating, Tilbrook, Rossell, Enticott, & Fitzgerald, 2012; Selby & Coniglio, 2020).

62 Additionally, individuals with AN may exhibit diminished reward responses specifically

63 towards food-related stimuli (Wierenga et al., 2014). In a broader sense, research has shown

64 that AN is associated with reduced subjective reward sensitivity and decreased neural

65 response to rewarding stimuli. Moreover, individuals with AN may experience disruptions in

66 processing aversive stimuli, leading to heightened harm avoidance, reduced tolerance for

67 uncertainty, increased anxiety, and heightened sensitivity to punishment (Fladung, Schulze,

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

69 Schmidt, 2015). These factors collectively contribute to an altered response to negative

70 feedback and a propensity to actively avoid aversive outcomes (Jonker, Glashouwer, & Jong,

71 2022; Matton, Goossens, Braet, & Vervaet, 2013). Neuroimaging studies have further

72 supported these findings by revealing neural dysfunctions in individuals with AN regarding

73 their response to loss and aversive taste stimuli (Bischoff-Grethe et al., 2013; Monteleone et

74 al., 2017; Wagner et al., 2007).

75   Given the crucial importance of RL in acquiring knowledge from past experiences,

76 extensive research has been conducted to examine potential deficits in RL among individuals

77 diagnosed with AN (Bischoff-Grethe et al., 2013; Glashouwer, Bloot, Veenstra, Franken, &

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

79 Matton et al., 2013). However, the findings of these studies have been mixed. For example,

80 Ritschel et al. (2017) found that individuals who had recovered from AN had impaired RL

81 performance compared to healthy controls (HCs) on a Probabilistic Reversal Learning (PRL)

82 task, particularly in response to negative feedback. In contrast, Bernardoni et al. (2018)

83 found that AN patients had a higher learning rate from punishment than HCs. Similarly,

84 Sarrar et al. (2015) found no differences in task performance between individuals with acute

85 AN and HCs using the Probabilistic Object Reversal Task with neutral stimuli. Geisler et

86 al. (2018) also found no group differences in a PRL task with neutral stimuli and monetary

87 feedback.

88   To shed light on the potential reinforcement RL deficits in AN, researchers have

89 incorporated food-related information into the PRL paradigm. For example, Zang et

90 al. (2014) found that individuals with binge-eating disorder (BED) exhibited poorer

91 performance when exposed to food-related feedback, indicating a vulnerability to

92 food-related cues. However, attempts to replicate these findings in AN have yielded

93 conflicting results. Hildebrandt et al. (2015) reported increased inflexibility in AN

94 individuals using a PRL task with food-related feedback, while Hildebrandt et al. (2018)

95 found no differences in PRL performance between AN patients and healthy controls (HCs)

96 when employing the same paradigm.

97   Given the inconsistent findings in behavioral experiments regarding RL in AN, we

98 propose that these discrepancies may be partly attribute to the predominant use of general

99 stimuli instead of stimuli specifically relevant to the disorder (Schaefer & Steinglass, 2021).

100 Furthermore, when disorder-related information has been incorporated, it has typically been

101 limited to the feedback provided after the participant's choice, with the stimuli presented

102 during the decision-making process unrelated to the disorder. This approach primarily

103 emphasizes the consequences of the choices, neglecting the contextual factors surrounding

104 the decision-making process.

105   From a theoretical perspective, the methodological choices made in previous studies

106 overlook the critical role of context in learning processes. Contextual learning (Heald,

107 Lengyel, & Wolpert, 2023) draws upon the notion, grounded in the human memory

108 literature, that memory retrieval depends on the match between the conditions during

109 learning and testing. When there is a mismatch, retrieval is impaired. Applying this concept

110 to RL in AN (Rosas, Todd, & Bouton, 2013), it can be hypothesized that contextual factors,

111 such as individual characteristics, long-term goals, and situational influences, can contribute

¹¹² to impaired RL specifically in decision-making related to the disorder. This impairment can

¹¹³ occur even when RL performance remains intact for decisions unrelated to the disorder

¹¹⁴ (Haynos, Widge, Anderson, & Redish, 2022). In other words, we propose that intermittent

¹¹⁵ impaired RL performance in AN may arise from contextual factors that activate specific

¹¹⁶ learning modes, rather than indicating a fundamental alteration in the underlying learning

¹¹⁷ processes in the brain (for a discussion, see Bernardoni et al., 2021).

¹¹⁸   To examine the proposed hypothesis, we conducted a study using a modified version of

¹¹⁹ the standard PRL task. Unlike previous studies that used general stimuli (Schaefer &

¹²⁰ Steinglass, 2021), our task incorporated two distinct contexts. In one context, participants

¹²¹ made choices between a stimulus related to the disorder (e.g., a caloric food) and a stimulus

¹²² unrelated to the disorder (e.g., a lamp). In the other context, participants made choices

¹²³ between two stimuli unrelated to the disorder (i.e., flower vs. objects).

¹²⁴   Based on the responses to food stimuli often seen in individuals with AN, which

¹²⁵ typically exhibit reduced reward, increased aversion, and inhibition (Haynos, Lavender,

¹²⁶ Nelson, Crow, & Peterson, 2020), we hypothesized that there would be a more conservative

¹²⁷ learning rate for disorder-related choices compared to disorder-unrelated choices within a

¹²⁸ PRL task. Additionally, we anticipated a lower learning rate for disorder-related choices in

¹²⁹ individuals with AN compared to HCs. Conversely, we expected no learning abnormalities in

¹³⁰ individuals with AN for choices unrelated to the disorder.

Methods

¹³² The study, which adhered to the Declaration of Helsinki, was approved by the

¹³³ University of Florence's Ethical Committee (Prot. n. 0178082). All eligible participants

¹³⁴ provided informed consent and willingly agreed to participate in the study.

Participants

The study included 40 individuals diagnosed with Restricting-Type Anorexia Nervosa

(R-AN) according to the DSM-5 criteria and 45 healthy volunteers (Table 2). The

participants with R-AN were recruited from three facilities in Italy: the Specchidacqua

Institute in Montecatini (Pisa), the Villa dei Pini Institute in Firenze, and the Gruber

Center, Outpatient Clinic in Bologna. The treatment approach consisted of Cognitive

Behavioral Therapy and family-based treatment, with patients receiving treatment for 2 to 6

hours per day, 2 days per week. The treatment program included various components such

as individual therapy, family therapy, group therapy, nutritional counseling, psychiatric care,

and medical monitoring. The diagnosis of AN was made by specialized psychiatrists and

psychologists through a semi-structured interview based on the DSM-5 criteria, at the start

of treatment.

The study recruited additional 310 female adolescents or young adults through social

media or university advertisements, and they completed the Eating Attitudes Test-26

(EAT-26) screening tool. Those scoring above 20 on the EAT-26 and not receiving current

treatment for eating disorders were classified as "at-risk" [RI; Dotti and Lazzari (1998)],

resulting in 36 individuals. From the remaining participants scoring below 20 on the EAT-26

and not receiving current treatment for eating disorders, 45 females were randomly selected

and assigned to the healthy control (HC) group. Both the RI and HC groups had

participants with a normal Body Mass Index (BMI). More details about eligibility criteria

and sample characteristics can be found in the Supplementary Information (SI).

Procedure

In the initial session, participants underwent a clinical interview to determine their

eligibility for the study. Those who met the criteria proceeded to anthropometric

measurements and completed a battery of psychometric scales: the Eating Attitude Test

(EAT-26;), the Body Shape Questionnaire-14 (BSQ-14; Dowson & Henderson, 2001), the

161 Social Interaction Anxiety Scale (SIAS; Mattick & Clarke, 1998), the Depression Anxiety

162 Stress Scale-21 (DASS-21; Lovibond & Lovibond, 1995), the Rosenberg Self-Esteem Scale

163 (RSES; Rosenberg, 1965), and the Multidimensional Perfectionism Scale (MPS-F; Frost et

164 al., 1990). In a subsequent session, participants completed the PRL task. Participants were

165 told they would be playing a simple computer game with the objective of accumulating as

166 many "virtual euro" as possible (see Figure 1).

167   The stimuli shown in each trial were randomly selected from sets of food-related and

168 food-unrelated images. All images used in the study were obtained from the International

169 Affective Picture System (IAPS) database (Lang et al., 2005) – for details, see the SI.

Data analysis

171   To analyze the temporal dynamics of the two-choice decision-making in the PRL task,

172 we employed a hierarchical reinforcement learning drift diffusion model (RLDDM), as

173 described in Pedersen, Frank, and Biele (2017) and Pedersen and Frank (2020). Cognitive

174 modeling analysis allows us to deconstruct decision-making task performance into its

175 component processes, which can help us identify deviations in the underlying mechanisms

176 that may not be evident in the overall task outcome.

177   We used a Bayesian approach in our study because estimating RLDDM models is

178 currently only possible using Markov chain Monte Carlo (MCMC) procedures. Additionally,

179 the Bayesian approach prioritizes estimation over hypothesis testing, which helps to

180 overcome the binary nature of decision-making inherent in null hypothesis significance

181 testing (NHST) (Kruschke & Liddell, 2018). We determined credible effects by examining

182 95% credible intervals or assessing the proportion of posterior samples (97.5%) indicating the

183 direction of the effect.

184   The RLDDM consists of two key components. The first component describes how

185 reward feedback is used to update value expectations, using a delta learning rule (Rescorla &
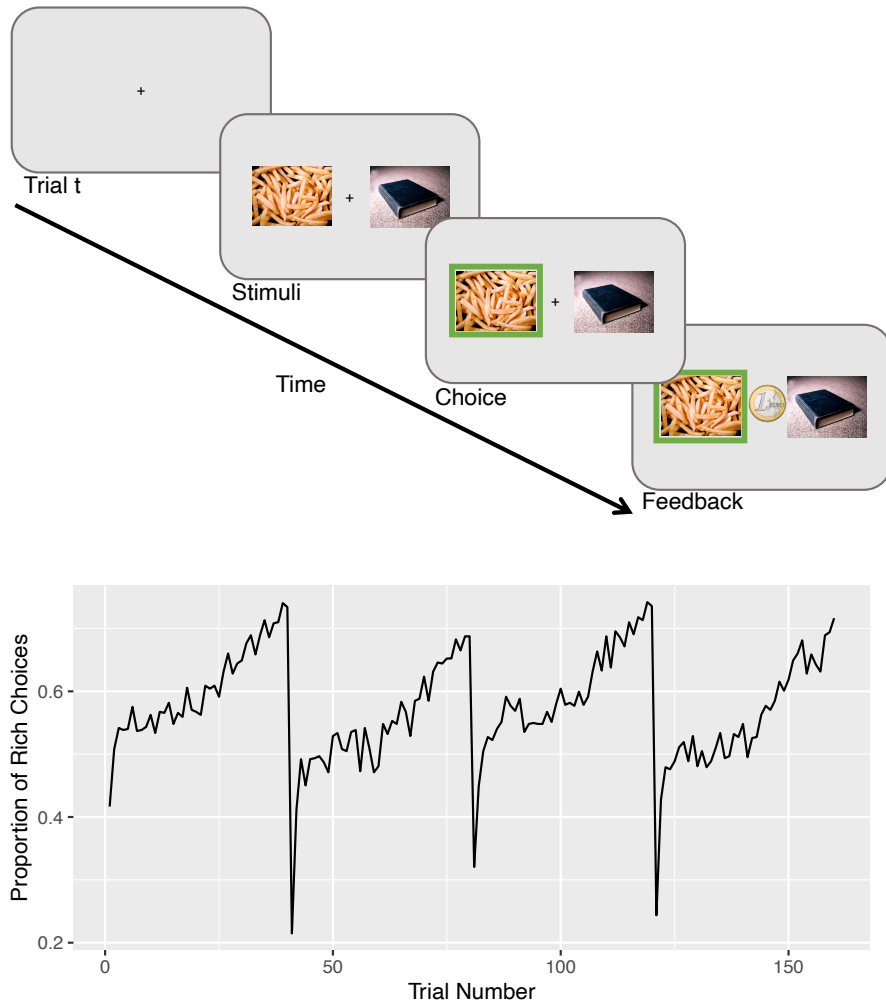
Figure 1. **Top.** During the PRL task, participants were presented with two images and had to choose the "correct" image based on trial-and-error feedback within 2.5 seconds. A euro coin image was displayed as a reward for correct responses, while a strike-through image of a euro coin served as a punishment for incorrect responses (left or right button press). The PRL task consisted of two blocks of 160 trials each. One block included pairs of food-related and food-unrelated images, while the other block only included food-unrelated images. The figure illustrates a single trial of the PRL task. **Bottom.** The trial-by-trial proportion of choosing the image with the highest probability of reward in the first epoch is shown for all participants. Initially, when participants have limited knowledge of the contingencies, the proportion of positive rewards is at chance level. However, as they engage in trial-and-error learning, their action selection improves, leading to enhanced performance and increased rewards. Following a context switch, there is an abrupt decline in performance, which is then followed by a gradual recovery that eventually levels off. It is worth noting that in this particular experiment, "correct choices" were rewarded positively only 70% of the time, thus establishing an upper limit of 0.7 on the plateau. Each session comprised 160 trials and included three context switches.

186   Wagner, 1972). The second component describes how an agent uses these expectations to

187   arrive at a decision, using a drift-diffusion model (Ratcliff & McKoon, 2008). The DDM

188   assumes that evidence for each option accumulates stochastically on each trial, until a

189   decision is made (for details, see SI).

190   The RLDDM has six basic parameters:

191   • α^+ and α^-: These parameters quantify the learning rate in the Rescorla-Wagner delta

192   learning rule. A higher learning rate results in rapid adaptation to reward expectations,

193   while a lower learning rate results in slow adaptation. The parameter α^+ is computed

194   from reinforcements, whereas α^- is computed from punishments.

195   • v: The drift rate is the average speed of evidence accumulation toward one decision.

196   • a: The decision boundary is the distance between two decision thresholds. An increase

197   of a increases the evidence needed to make a decision. The increase of a leads to a

198   slower but more accurate decision; a decrease in a results in a faster but error-prone

199   decision.

200   • t: The non-decision time is the time spent for stimuli encoding or motor execution (i.e.,

201   time not used for evidence accumulation).

202   • z: The starting point parameter captures a potential initial bias toward one or the

203   other boundary in absence of any stimulus evidence.

204   To assess the presence of context-dependent learning, we conditioned the model's

205   parameters on two contexts: disorder-related choices and disorder-unrelated choices. This

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

207   contextual conditions.

<div align="center">

Results

</div>

Models selection

We evaluated context-dependent learning by comparing several RLDDM models that differed in how they conditioned the model's parameters on group (R-AN, HC, RI) and context (disorder-related choices and disorder-unrelated choices). We used the Deviance Information Criterion (DIC) to balance model fit and complexity, and selected the model with the lowest DIC as the best trade-off. The following RLDDM models were examined. Model M1: Standard RLDDM without conditioning. DIC = 39879.444. Model M2: Separate learning rates for positive and negative reinforcements. DIC = 39124.890 Model M3: Group-based α^+ and α^- parameters. DIC = 39194.763. Model M4: Group and context-based α^+ and α^- parameters. DIC = 38197.467. Model M5: Group and context-based α^+, α^-, and a parameters. DIC = 36427.448. Model M6: Group and context-based α^+, α^-, a, and drift rate (v) parameters. DIC = 36185.146. Model M7: Group and context-based α^+, α^-, a, v, and non-decision time (t) parameters. DIC = 34904.053. Model M8: Group and context-based α^+, α^-, a, v, t, and starting point (z) parameters. DIC = 34917.762. Among the evaluated models, Model M7 had the lowest DIC, indicating the best trade-off between goodness of fit and model complexity. In Model M7, the parameters α^+, 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. The $\hat{R}$ values for all parameters in Model M7 were below 1.1, indicating that the model converged well. 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

233 RL learning, we compared the posterior estimates of the RLDDM parameters of Model M7

234 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
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

235 Let's first consider the evidence of context-dependent learning from within-group

236 comparisons. We found that individuals in the R-AN group had a reduced learning rate from

237 rewards for disorder-related choices, compared to disorder-unrelated choices (Cohen's $d =$

238 -1.206, $p = 0.0098$). However, no credible difference was found in the learning rate between

239 disorder-related and disorder-unrelated choices in the HC ($p = 0.5544$) or RI ($p = 0.2247$)

240 groups. We found no credible difference in the learning rate from punishments between

241 disorder-related and disorder-unrelated choices for any of the R-AN ($p = 0.2349$), HC ($p =$

242 0.6993), and RI ($p = 0.5101$) groups. Moreover, we found that individuals with R-AN

243 showed a higher decision threshold for disorder-related choices compared to

244 disorder-unrelated choices (Cohen's $d = 0.802$, $p = 0.0013$).

245   Further evidence of context-dependent learning emerges from between-groups

246 comparisons. When making disorder-related choices, individuals with R-AN displayed a

247 decreased learning rate from rewards compared to both HC ($p = 0.0009$, Cohen's $d = 1.498$)

248 and RI ($p = 0.0085$, Cohen's $d = 1.209$). In contrast, no credible difference in the learning

249 rate from rewards was found between R-AN and HC ($p = 0.4325$), as well as between R-AN

250 and RI ($p = 0.3232$), for choices unrelated to disorder information. Individuals with R-AN

251 showed a higher decision threshold for disorder-related choices compared to both HC

252 (Cohen's $d = -0.622$, $p = 0.0068$) and RI (Cohen's $d = -0.454$, $p = 0.0118$) participants,

253 whereas no credible group differences were found for disorder-unrelated choices. Finally, no

254 credible differences were found, for both within-group and between-group comparisons,

255 regarding the drift rate (ν) and the non-decision time parameters (t).

256 **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

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

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

Comorbidity

To examine the potential impact of comorbidity and medication status on the present

results, we compared R-AN participants who had diagnosed comorbidities (45% of the

sample) and those without any comorbid conditions using Model M7. We did not find any

credible differences in parameters between the two groups. Specifically, when considering the

disorder-related context, the parameter differences were as follows: $\Delta\alpha^- = 2.614$, 95% CI

[-3.173, 8.364]; $\Delta\alpha^+ = -0.635$, 95% CI [-4.301, 2.449]; $\Delta a = -0.034$, 95% CI [-0.188, 0.124];

$\Delta v = 0.230$, 95% CI [-1.203, 1.586]; $\Delta t = 0.002$, 95% CI [-0.050, 0.055]. Similarly, for the

disorder-unrelated context, the parameter differences were: $\Delta\alpha^- = -0.768$, 95% CI [-6.570,

4.401]; $\Delta\alpha^+ = -1.739$, 95% CI [-6.184, 1.654]; $\Delta a = -0.126$, 95% CI [-0.281, 0.025]; $\Delta v =$

0.744, 95% CI [-0.453, 1.886]; $\Delta t = -0.003$, 95% CI [-0.057, 0.052]. The correlation between

comorbidity and medication was 0.78.

Discussion

Our study reveal a context-dependent learning asymmetry in individuals with R-AN,

specifically in the positive learning rate (α^+). This asymmetry is evident when comparing

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

choices. Notably, this difference is not observed in the two control groups.

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

between-group comparisons. Individuals with R-AN exhibited lower learning rates (α^+)

compared to the HC and RI group. However, these differences were observed only for

disorder-related choices. In contrast, no credible differences in learning rates were found

among the three groups for choices unrelated to the disorder.

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

parameters of the RLDDM model. Specifically, we observed that the R-AN group exhibited a higher decision threshold (parameter "a" in the RLDDM model) compared to the HC and RI groups, but this difference was only evident in the context of disorder-related choices. Additionally, individuals with R-AN exhibited a higher decision threshold for disorder-related choices compared to disorder-unrelated choices, which was not seen in the HC and RI groups. Overall, these results suggest that individuals with R-AN tended to display a more cautious or conservative decision-making behavior in relation to choices related to the disorder (see also Caudek, Sica, Cerea, Colpizzi, & Stendardi, 2021; Schiff, Testa, Rusconi, Angeli, & Mapelli, 2021). We did not find any credible contextual effects or group differences for the parameters ν and t.

The analysis of preferential choices supports the conclusion that the learning performance asymmetry observed in individuals with R-AN is not due to a preferential selection of the disorder-unrelated image during the learning task. Additionally, our analysis examining the relationship between the model's parameters and the presence of comorbidities indicates that the learning performance asymmetry in individuals with R-AN cannot be attributed to comorbid conditions.

General 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 lower learning rate in the disorder-related context for individuals 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.

313 First, there may be a learning impairment, where participants struggle to accurately update

314 the value of the stimuli. Second, there may be a decision impairment, where participants

315 may still select the wrong stimulus despite having intact learning processes. Our results show

316 that individuals with R-AN may struggle with both accurately updating the value of

317 disorder-related stimuli and making appropriate decisions based on this information.

318 However, we did not observe similar impairments in decision making for disorder-unrelated

319 choices. These findings provide evidence for context-dependent learning in individuals with

320 R-AN, where the inclusion of disorder-related information negatively impacts their RL

321 performance. It is important to note that this effect is specific to the disorder-related context

322 and does not suggest a generalized RL deficit in individuals with R-AN. Thus, our results

323 challenge the notion of a domain-general RL mechanism impairment in this population (see

324 Bernardoni et al., 2021).

325   Previous studies have demonstrated that reward and punishment processing in

326 individuals with AN is influenced by stimulus properties and contextual factors. For

327 instance, predictable and controllable behaviors such as calorie counting or purging are often

328 perceived as rewarding, providing individuals with a sense of control and accomplishment.

329 Conversely, unpredictable and uncontrollable situations, such as social outcomes, can be

330 perceived as punishing, leading to heightened anxiety and distress (Haynos et al., 2020).

331 While previous studies have predominantly examined the impact of context on the subjective

332 value attributed to experiences in AN, our study expands on this research by demonstrating

333 that context plays a crucial role in the actual learning process itself (Heald et al., 2023).

334 This goes beyond solely influencing subjective value and provides valuable insights into how

335 reward and punishment processing operates in AN.

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

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

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

or goal-directed (model-based) learning. For instance, studies conducted by Foerde et al. (2021) and Onysk and Seriès (2022) employed similar experiments using the two-step task paradigm. Foerde et al. (2021) compared a monetary two-step task and a food-related two-step task, while Onysk and Seriès (2022) utilized stimuli unrelated to food or body images (i.e., pirate ships and treasure chests) with rewards associated with body image dissatisfaction. The results of these studies consistently demonstrated that individuals with AN tend to exhibit a stronger inclination towards habitual control over goal-directed control across different domains compared to healthy controls. However, these studies did not identify significant differences in learning rates based on context or between AN patients and healthy controls. In contrast, the present study reveals that, in individuals with R-AN, the learning process *per se* (i.e., the learning rates) can be influenced by contextual (disorder-related) information, even when contextual information is not directly relevant to the task outcome.

The finding that individuals with R-AN demonstrated a conservative learning rate specifically for rewards in response to disorder-related context, but not for punishments, supports a neurobiological model suggesting that AN symptoms are maintained by diminished responses to rewarding stimuli and exaggerated responses to aversive stimuli, potentially due to an overreliance on cognitive control brain circuitry (Wierenga et al., 2014). A study by Bernardoni et al. (2018), which used a PRL task with disorder-unrelated stimuli, revealed increased learning rates specifically after punishments in AN. This study suggested that AN individuals, when faced with negative feedback, tend to critically reassess their beliefs more quickly than healthy controls (HCs), leading them to base decisions on more recent experiences. Such inclination may be linked to pre-existing harm-avoidant traits in AN patients, characterized by a reduced tolerance for negative outcomes. Considering these factors, we speculate that the lack of contextual effect on negative learning rates in our study may arise from a trade-off between the conservative learning induced by disorder-related cues and the increased learning rate observed after punishment.

366   The hypothesis proposing that reinforcement learning (RL) anomalies in individuals

367   with anorexia nervosa (AN) may be influenced by contextual factors carries significant

368   implications for treatment strategies. Currently, Cognitive Remediation Therapy (CRT) is

369   utilized to address cognitive inflexibility in AN and other eating disorders. CRT involves

370   cognitive exercises and behavioral interventions aimed at improving central coherence

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

372   comprehension (Tchanturia, Davies, Reeder, & Wykes, 2010). A key aspect of CRT is to

373   avoid addressing symptom-related themes and instead utilize neutral stimuli in cognitive and

374   behavioral exercises. This approach aims to establish a therapeutic alliance and reduce

375   drop-out rates, particularly among individuals with AN. However, recent evidence suggests

376   that CRT may not consistently improve central coherence abilities, cognitive flexibility, or

377   symptoms associated with eating disorders (Hagan, Christensen, & Forbush, 2020;

378   Tchanturia, Giombini, Leppanen, & Kinnaird, 2017). In response to these findings, Trapp et

379   al. (2022) propose modifications to address practical challenges encountered in the

380   application of CRT. They question the use of neutral stimuli and draw support from Beck's

381   cognitive theory of depression (Beck & Alford, 2009). This proposition aligns with the

382   hypothesis of our study. If further studies consistently demonstrate that maladaptive RL is

383   context-dependent, it would necessitate a shift in intervention approaches.

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

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

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

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

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

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

390   future studies could aim to determine the equivalence of subjective values for rewards and

391   punishments to enhance the understanding of the underlying processes. 2) Our study only

392   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

411

American Psychiatric Association. (2022). *Diagnostic and Statistical Manual of Mental*
412

Disorders (5th ed., Text Revision). Arlington, VA: American Psychiatric Publishing.
413

Atwood, M. E., & Friedman, A. (2020). A systematic review of enhanced cognitive
414

behavioral therapy (CBT-e) for eating disorders. *International Journal of Eating*
415

Disorders, *53*(3), 311–330.
416

Bartholdy, S., Dalton, B., O'Daly, O. G., Campbell, I. C., & Schmidt, U. (2016). A
417

systematic review of the relationship between eating, weight and inhibitory control using
418

the stop signal task. *Neuroscience & Biobehavioral Reviews*, *64*, 35–62.
419

Beck, A. T., & Alford, B. A. (2009). *Depression: Causes and treatment.* University of
420

Pennsylvania Press.
421

Bernardoni, F., Geisler, D., King, J. A., Javadi, A.-H., Ritschel, F., Murr, J., . . . Ehrlich, S.
422

(2018). Altered medial frontal feedback learning signals in Anorexia Nervosa. *Biological*
423

Psychiatry, *83*(3), 235–243.
424

Bernardoni, F., King, J. A., Geisler, D., Ritschel, F., Schwoebel, S., Reiter, A. M., . . .
425

Ehrlich, S. (2021). More by stick than by carrot: A reinforcement learning style rooted in
426

the medial frontal cortex in anorexia nervosa. *Journal of Abnormal Psychology*, *130*(7),
427

736–747.
428

Bischoff-Grethe, A., McCurdy, D., Grenesko-Stevens, E., Irvine, L. E. Z., Wagner, A., Yau,
429

W.-Y. W., et al.others. (2013). Altered brain response to reward and punishment in
430

adolescents with anorexia nervosa. *Psychiatry Research: Neuroimaging*, *214*(3), 331–340.
431

Caudek, C., Sica, C., Cerea, S., Colpizzi, I., & Stendardi, D. (2021). Susceptibility to eating
432

disorders is associated with cognitive inflexibility in female university students. *Journal*
433

of Behavioral and Cognitive Therapy, *31*(4), 317–328.
434

Chang, P. G., Delgadillo, J., & Waller, G. (2021). Early response to psychological treatment
435

for eating disorders: A systematic review and meta-analysis. *Clinical Psychology Review*,
436

86, 102032.
437

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.

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

465   and deficits: How computational psychiatry can enhance an understanding of

466   decision-making in anorexia nervosa. *Current Psychiatry Reports*, 1–11.

467   Heald, J. B., Lengyel, M., & Wolpert, D. M. (2023). Contextual inference in learning and

468   memory. *Trends in Cognitive Sciences*.

469   Jappe, L. M., Frank, G. K., Shott, M. E., Rollin, M. D., Pryor, T., Hagman, J. O., . . .

470   Davis, E. (2011). Heightened sensitivity to reward and punishment in anorexia nervosa.

471   *International Journal of Eating Disorders*, *44*(4), 317–324.

472   Jonker, N. C., Glashouwer, K. A., & Jong, P. J. de. (2022). Punishment sensitivity and the

473   persistence of anorexia nervosa: High punishment sensitivity is related to a less favorable

474   course of anorexia nervosa. *International Journal of Eating Disorders*, *55*(5), 697–702.

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

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

477   Keating, C., Tilbrook, A. J., Rossell, S. L., Enticott, P. G., & Fitzgerald, P. B. (2012).

478   Reward processing in anorexia nervosa. *Neuropsychologia*, *50*(5), 567–575.

479   Kruschke, J. K., & Liddell, T. M. (2018). Bayesian data analysis for newcomers.

480   *Psychonomic Bulletin & Review*, *25*(1), 155–177.

481   Linardon, J., Fairburn, C. G., Fitzsimmons-Craft, E. E., Wilfley, D. E., & Brennan, L.

482   (2017). The empirical status of the third-wave behaviour therapies for the treatment of

483   eating disorders: A systematic review. *Clinical Psychology Review*, *58*, 125–140.

484   Matton, A., Goossens, L., Braet, C., & Vervaet, M. (2013). Punishment and reward

485   sensitivity: Are naturally occurring clusters in these traits related to eating and weight

486   problems in adolescents? *European Eating Disorders Review*, *21*(3), 184–194.

487   Monteleone, A. M., Monteleone, P., Esposito, F., Prinster, A., Volpe, U., Cantone, E., et

488   al.others. (2017). Altered processing of rewarding and aversive basic taste stimuli in

489   symptomatic women with anorexia nervosa and bulimia nervosa: An fMRI study.

490   *Journal of Psychiatric Research*, *90*, 94–101.

491   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.

Rosas, J. M., Todd, T. P., & Bouton, M. E. (2013). Context change and associative learning. *Wiley Interdisciplinary Reviews: Cognitive Science*, *4*(3), 237–244.

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

519   healthy starving individuals. A behavioral study. *Frontiers in Psychology, 11*, 3902.

520 Selby, E. A., & Coniglio, K. A. (2020). Positive emotion and motivational dynamics in

521   anorexia nervosa: A positive emotion amplification model (PE-AMP). *Psychological*

522   *Review, 127*(5), 853–890.

523 Smink, F. R., Hoeken, D. van, & Hoek, H. W. (2013). Epidemiology, course, and outcome of

524   eating disorders. *Current Opinion in Psychiatry, 26*(6), 543–548.

525 Tchanturia, K., Davies, H., Reeder, C., & Wykes, T. (2010). *Cognitive remediation therapy*

526   *for anorexia nervosa.* London: King's College London.

527 Tchanturia, K., Giombini, L., Leppanen, J., & Kinnaird, E. (2017). Evidence for cognitive

528   remediation therapy in young people with anorexia nervosa: Systematic review and

529   meta-analysis of the literature. *European Eating Disorders Review, 25*(4), 227–236.

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

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

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

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

534   (2014). Are extremes of consumption in eating disorders related to an altered balance

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

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

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

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

Table 2
Summary of Sample's Demographic and Clinical Characteristics.

	R-AN (n = 40)	HC (n = 45)	RI (n = 36)	AN - HC	AN - RI	AN - HC	AN - RI
	Mean (SD)	Mean (SD)	Mean (SD)	PE (95% CI)	PE (95% CI)	Cohen's d	Cohen's d
Age (years)	21.11 (4.33)	19.49 (2.32)	21.31 (4.82)	-0.00 (-0.22, 0.18)	0.01 (-0.21, 0.25)	0.09	-0.18
Education (years)	14.53 (1.11)	14.11 (0.98)	13.89 (0.78)	0.08 (-0.14, 0.33)	0.13 (-0.11, 0.41)	-0.08	-0.13
BMI (kg/m^2)	17.79 (2.85)	21.78 (3.53)	21.64 (4.11)	-2.74 (-3.67, -1.79)	-3.22 (-4.32, -2.07)	2.67	3.15
EAT-26 Total score	35.89 (19.46)	5.09 (5.10)	25.86 (12.44)	1.76 (1.32, 2.20)	0.18 (-0.25, 0.62)	-1.84	-0.18
EAT-26 Dieting	19.11 (11.24)	3.27 (3.96)	16.06 (8.00)	53.96 (32.78, 80.9)	4.38 (-2.46, 12.10)	-4.54	-0.36
EAT-26 Bulimia	7.17 (3.95)	0.76 (1.48)	5.78 (3.86)	1.36 (0.99, 1.72)	0.11 (-0.17, 0.41)	-2.33	-0.18
EAT-26 Oral control	9.61 (6.23)	1.07 (1.67)	4.03 (4.05)	1.53 (1.12, 1.96)	0.81 (0.42, 1.20)	-1.95	-1.03
BSQ-14 Total score	139.78 (35.26)	97.47 (32.37)	147.94 (37.13)	17.50 (11.30, 23.91)	-3.40 (-10.10, 3.15)	-1.21	0.24
RSES Total score	22.69 (5.29)	28.33 (5.76)	22.53 (5.76)	-5.601 (-8.12, -3.15)	0.192 (-2.31, 2.90)	-1.22	0.23
DASS-21 Stress	12.86 (4.67)	9.13 (3.55)	12.17 (3.74)	3.73 (1.89, 5.38)	0.69 (-1.12, 2.59)	-0.94	-0.18
DASS-21 Depression	10.61 (5.49)	6.82 (4.33)	11.22 (4.99)	3.78 (1.76, 6.14)	-0.62 (-2.91, 1.77)	-0.77	0.12
DASS-21 Anxiety	8.25 (4.51)	5.76 (4.26)	7.56 (4.47)	2.48 (0.50, 4.33)	0.70 (-1.39, 2.64)	-0.56	-0.16
SIAS Total score	37.31 (15.45)	27.69 (13.01)	39.03 (14.87)	9.63 (3.12, 15.84)	-1.68 (-8.28, 5.07)	-0.66	0.12
MPS Cmd	45.47 (8.21)	40.02 (7.24)	49.25 (8.17)	5.46 (2.05, 9.02)	-3.73 (-7.47, -0.15)	-0.69	0.48
MPS Ps	25.33 (5.74)	22.00 (4.86)	25.67 (6.33)	3.33 (0.89, 5.86)	-0.32 (-3.07, 2.19)	-0.60	0.06
MPS Pepc	20.78 (6.64)	21.02 (5.84)	25.22 (7.92)	-0.63 (-3.49, 2.42)	-3.88 (-7.09, -0.90)	0.09	0.57
MPS Or	23.94 (5.11)	23.07 (5.16)	22.17 (5.55)	0.84 (-1.49, 3.17)	1.77 (-0.66, 4.26)	-0.17	-0.34