Contextual influence of reinforcement learning performance in Anorexia Nervosa

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

[1] Wilhelm-Wundt-University

[2] Konstanz Business School

Author Note

Add complete departmental affiliations for each author here. Each new line herein must be indented, like this line.

Enter author note here.

The authors made the following contributions. Corrado Caudek: Conceptualization, Writing - Original Draft Preparation, Writing - Review & Editing; Ernst-August Doelle: Writing - Review & Editing, Supervision.

Correspondence concerning this article should be addressed to Corrado Caudek, Postal 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, despite experiencing negative consequences, along with evidence suggesting changes in responsiveness to rewards and punishments, has led researchers to propose that there may be abnormal reward processing and learning in AN (Schaefer & Steinglass, 2021). Although there is substantial evidence supporting the existence of anomalies in reward sensitivity, our understanding of potential abnormalities specifically related to the learning processes in AN is currently limited.

In relation to dysfunctions reward responsiveness among individuals with AN, research has revealed that the intense levels of dietary restriction and physical activity typically associated with AN can trigger the activation of reward pathways in the brain (Keating, 2010; Keating, Tilbrook, Rossell, Enticott, & Fitzgerald, 2012; Selby & Coniglio, 2020). Additionally, individuals with AN may exhibit diminished reward responses specifically towards food-related stimuli (Wierenga et al., 2014). In a broader sense, research has shown that AN is associated with reduced subjective reward sensitivity and decreased neural response to rewarding stimuli. Moreover, individuals with AN may experience disruptions in processing aversive stimuli, leading to heightened harm avoidance, reduced tolerance for uncertainty, increased anxiety, and heightened sensitivity to punishment (Fladung, Schulze, Schöll, Bauer, & Groen, 2013; Jappe et al., 2011; Keating et al., 2012; O'Hara, Campbell, & Schmidt, 2015). These factors collectively contribute to an altered response to negative feedback and a propensity to actively avoid aversive outcomes (Jonker, Glashouwer, & Jong, 2022; Matton, Goossens, Braet, & Vervaet, 2013). Neuroimaging studies have further supported these findings by revealing neural dysfunctions in individuals with AN regarding their response to loss and aversive taste stimuli (Bischoff-Grethe et al., 2013; Monteleone et al., 2017; Wagner et al., 2007).

Given the crucial importance of RL in acquiring knowledge from past experiences, extensive research has been conducted to examine potential deficits in RL among individuals diagnosed with AN (Bischoff-Grethe et al., 2013; Glashouwer, Bloot, Veenstra, Franken, & Jong, 2014; Harrison, Genders, Davies, Treasure, & Tchanturia, 2011; Jappe et al., 2011; Matton et al., 2013). However, the findings of these studies have been mixed. For example, Ritschel et al. (2017) found that individuals who had recovered from AN had impaired RL performance compared to healthy controls (HCs) on a Probabilistic Reversal Learning (PRL) task, particularly in response to negative feedback. In contrast, Bernardoni et al. (2018) found that AN patients had a higher learning rate from punishment than HCs. Similarly, Sarrar et al. (2015) found no differences in task performance between individuals with acute AN and HCs using the Probabilistic Object Reversal Task with neutral stimuli. Geisler et al. (2018) also found no group differences in a PRL task with neutral stimuli and monetary feedback.

To shed light on the potential reinforcement RL deficits in AN, researchers have incorporated food-related information into the PRL paradigm. For example, Zang et al. (2014) found that individuals with binge-eating disorder (BED) exhibited poorer performance when exposed to food-related feedback, indicating a vulnerability to food-related cues. However, attempts to replicate these findings in AN have yielded conflicting results. Hildebrandt et al. (2015) reported increased inflexibility in AN individuals using a PRL task with food-related feedback, while Hildebrandt et al. (2018) found no differences in PRL performance between AN patients and healthy controls (HCs) when employing the same paradigm.

Given the inconsistent findings in behavioral experiments regarding RL in AN, we propose that these discrepancies may be partly attribute to the predominant use of general stimuli instead of stimuli specifically relevant to the disorder (Schaefer & Steinglass, 2021). Furthermore, when disorder-related information has been incorporated, it has typically been limited to the feedback provided after the participant's choice, with the stimuli presented during the decision-making process unrelated to the disorder. This approach primarily emphasizes the consequences of the choices, neglecting the contextual factors surrounding the decision-making process.

From a theoretical perspective, the methodological choices made in previous studies overlook the critical role of context in learning processes. Contextual learning (Heald, Lengyel, & Wolpert, 2023) draws upon the notion, grounded in the human memory literature, that memory retrieval depends on the match between the conditions during learning and testing. When there is a mismatch, retrieval is impaired. Applying this concept to RL in AN (Rosas, Todd, & Bouton, 2013), it can be hypothesized that contextual factors, such as individual characteristics, long-term goals, and situational influences, can contribute to impaired RL specifically in decision-making related to the disorder. This impairment can occur even when RL performance remains intact for decisions unrelated to the disorder (Haynos, Widge, Anderson, & Redish, 2022). In other words, we propose that intermittent impaired RL performance in AN may arise from contextual factors that activate specific learning modes, rather than indicating a fundamental alteration in the underlying learning processes in the brain (for a discussion, see Bernardoni et al., 2021).

To examine the proposed hypothesis, we conducted a study using a modified version of the standard PRL task. Unlike previous studies that used general stimuli (Schaefer & Steinglass, 2021), our task incorporated two distinct contexts. In one context, participants made choices between a stimulus related to the disorder (e.g., a caloric food) and a stimulus unrelated to the disorder (e.g., a lamp). In the other context, participants made choices between two stimuli unrelated to the disorder (i.e., flower vs. objects).

Based on the responses to food stimuli often seen in individuals with AN, which typically exhibit reduced reward, increased aversion, and inhibition (Haynos, Lavender, Nelson, Crow, & Peterson, 2020), we hypothesized that there would be a more conservative learning rate for disorder-related choices compared to disorder-unrelated choices within a PRL task. Additionally, we anticipated a lower learning rate for disorder-related choices in individuals with AN compared to HCs. Conversely, we expected no learning abnormalities in individuals with AN for choices unrelated to the disorder.

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 restricting-type anorexia nervosa (R-AN), healthy controls (HCs), and individuals at risk of developing eating disorders (RIs). The demographic and clinical characteristics of the sample are shown in Table 1. Participant selection, inclusion criteria, and sample characteristics are further described 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. In a subsequent session, participants completed the PRL task. Participants were told they would be playing a simple computer game with the objective of accumulating as many “virtual euro” as possible. During the PRL task, they were presented with two stimuli simultaneously on a screen and instructed to select one within 2.5 seconds by pressing a key. 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.

The PRL task consisted of two blocks of 160 trials each (see Figure 1). One block included pairs of food-related and food-unrelated images, while the other block only included food-unrelated images. The stimuli shown in each trial were randomly selected from sets of food-related and food-unrelated images.

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).

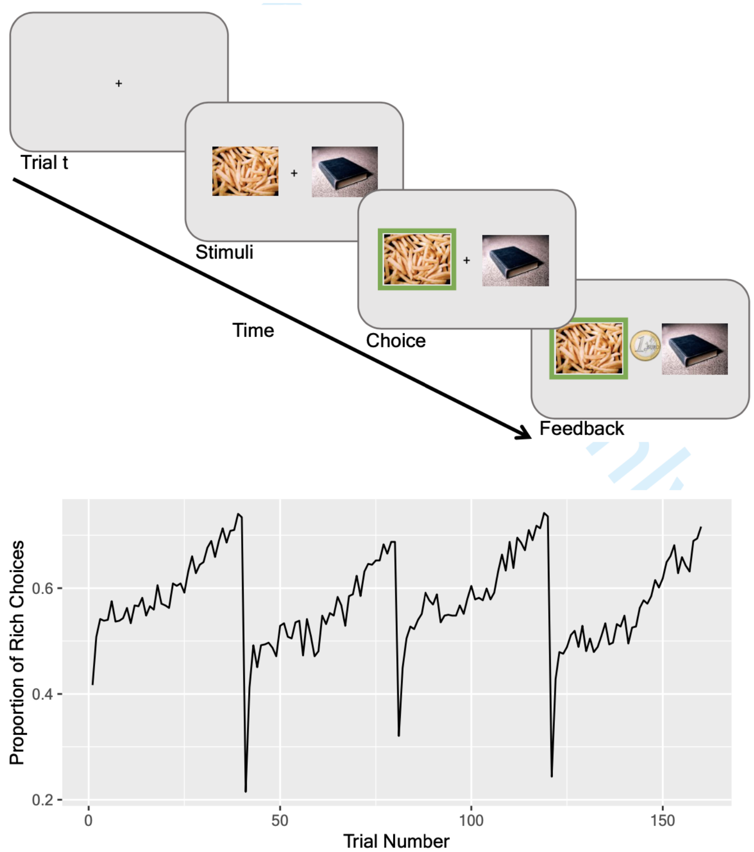

Figure 1. **Top.** The figure illustrates a single trial of the probabilistic reversal-learning task. Subjects were presented with two images and had to choose the "correct" image based on trial-and-error feedback. Feedback was provided in the form of an image of a euro coin or a crossed euro coin, depending on their choice (left or right button press). **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.

(for 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, which can help us identify deviations in the underlying mechanisms that may not be evident in the overall task outcome. The RLDDM consists of two key components. The first component describes how reward feedback is used to update value expectations, using a delta learning rule (Rescorla & Wagner, 1972). The second component describes how an agent uses these expectations to arrive at a decision, using a drift-diffusion model (Ratcliff & McKoon, 2008). The DDM assumes that evidence for each option accumulates stochastically on each trial, until a decision is made (for details, see SI).

The RLDDM has six basic parameters:

- α^+ and α^-: These parameters quantify the learning rate in the Rescorla-Wagner delta learning rule. 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.
- v: The drift rate is the average speed of evidence accumulation toward one decision.
- a: 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.
- t: The non-decision time is the time spent for stimuli encoding or motor execution (i.e., time not used for evidence accumulation).

- z: The starting point parameter 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 contexts: disorder-related choices and disorder-unrelated choices. This allowed us to examine how the model's parameters varied in response to these different contextual conditions.

Results

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 RL learning, we compared the posterior estimates of the RLDDM parameters of Model 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's first consider the evidence of context-dependent learning from within-group comparisons. We found that individuals in the R-AN group had a reduced learning rate in response to positive prediction errors (PEs) for disorder-related choices, compared to disorder-unrelated choices (Cohen's $d = 1.206$, $p = 0.0098$). However, no credible difference was found in the learning rate between disorder-related and disorder-unrelated choices in the HC ($p = 0.5544$) or 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. Concerning the learning rate after negative PEs, we found that R-AN showed a lower learning rate compared to HC, but only for disorder-related choices: (p = 0.0274, Cohen's d = 1.144). Individuals with R-AN showed a higher decision threshold for disorder-related choices compared to both HC (Cohen's d = 0.622, p = 0.0068) and RI (Cohen's d = 0.454, p = 0.0118) participants. No credible group differences were found for disorder-unrelated choices. Additionally, we observed that both HC (Cohen's d = 0.520, p = 0.0344) and RI (Cohen's d = 0.529, p = 0.0392) participants exhibited a faster accumulation of evidence and more confident decision-making, as indicated by a higher average drift rate parameter, compared to individuals with R-AN. This difference was only evident for disorder-unrelated choices. Finally, no credible differences were found, for both within-group and between-group comparisons, regarding the non-decision time parameter (t).

Preferential choices

To investigate the presence of a bias against food choices in individuals with R-AN during the PRL task, regardless of their past action-outcome history, we analyzed the frequency of food choices in PRL blocks where a food image was paired with a neutral image. Our results show that the AN-R group did not exhibit a bias against the food image, with a proportion of food choices estimated at 0.49, 95% CI [0.46, 0.51]. Furthermore, there were no 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]; Δa = -0.034, 95% CI [-0.188, 0.124];

Δv = 0.230, 95% CI [-1.203, 1.586]; Δt = 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]; Δa = -0.126, 95% CI [-0.281, 0.025]; Δv = 0.744, 95% CI [-0.453, 1.886]; Δt = -0.003, 95% CI [-0.057, 0.052]. The correlation between comorbidity and medication was 0.78.

Discussion

Our findings reveal a context-dependent learning asymmetry in individuals with R-AN specifically in the positive learning rate. This within-group asymmetry is observed when comparing the performance in the PRL task for disorder-related choices versus disorder-unrelated choices. Importantly, no similar difference is found 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 for both positive and negative prediction errors compared to the HC group, and specifically for positive prediction errors compared to the RI group, but these differences were observed only for disorder-related choices. In contrast, no credible differences in learning rates were found among the three groups for disorder-unrelated choices.

Support for context-dependent learning in R-AN is also provided by the DDM parameters of the hDDMrl model. Specifically, we observed that the R-AN group exhibited a higher decision threshold (parameter “a” in the hDDMrl model) compared to the HC and RI groups, but this difference was only evident in the context of disorder-related choices. This suggests that individuals with R-AN displayed a more cautious or conservative decision-making behavior specifically in relation to disorder-related choices (see also Caudek, Sica, Cerea, Colpizzi, & Stendardi, 2021; Schiff, Testa, Rusconi, Angeli, & Mapelli, 2021).

Further support of context-related learning in R-AN comes from the result which indicate that both healthy control (HC) and at-risk (RI) participants exhibited a faster accumulation of evidence and displayed more confident decision-making, as reflected by a higher average drift rate parameter, compared to individuals with restrictive anorexia nervosa (R-AN). However, this difference was specifically observed for disorder-unrelated choices. It is noteworthy that individuals with R-AN displayed slower evidence accumulation and less confident decision-making specifically in disorder-unrelated contexts, whereas this group difference was not observed for disorder-related choices. This finding further supports the notion of context-dependent learning in individuals with R-AN, particularly in the context of food-related information.

Further evidence of context-related learning in R-AN comes from the analysis of the drift rate parameter. Individuals with R-AN exhibited slower evidence accumulation and less confident decision-making compared to the control groups, specifically in the context of disorder-unrelated choices. Conversely, no credible group differences were observed for food-related choices. These results suggest that individuals with R-AN may allocate greater cognitive resources to process salient information in the disorder-related context, which leads to similar evidence accumulation rates in decision-making compared to the control groups. In contrast, they exhibit a slower evidence accumulation rate when faced with less salient disorder-unrelated choices.

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. 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 show 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 (see Bernardoni et al., 2021).

Previous studies have demonstrated that reward and punishment processing in individuals with AN is influenced by stimulus properties and contextual factors. 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 heightened anxiety and distress (Haynos et al., 2020). While previous studies have predominantly examined the impact of context on the subjective value attributed to experiences in AN, our study expands on this research by demonstrating that context plays a crucial role in the actual learning process itself (Heald et al., 2023). This goes beyond solely influencing subjective value and provides valuable insights into how reward and punishment processing operates in AN.

Other recent studies have focused on investigating context-specific learning in eating disorders. One task specifically designed for this purpose is the two-step Markov decision 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, no significant differences were observed in learning rates as a function of context, nor between AN patients and healthy controls, according to these findings. In contrast, the present study reveals that, in individuals with R-AN, the learning process *per se* can be influenced by contextual (disorder-related) information, even when such information is not directly relevant to the task outcome.

The hypothesis proposing that reinforcement learning (RL) anomalies in individuals with anorexia nervosa (AN) may be influenced by contextual factors carries significant implications for treatment strategies. Currently, Cognitive Remediation Therapy (CRT) is utilized to address cognitive inflexibility in AN and other eating disorders. CRT involves cognitive exercises and behavioral interventions aimed at improving central coherence abilities, reducing cognitive and behavioral inflexibility, and enhancing thinking style comprehension (Tchanturia, Davies, Reeder, & Wykes, 2010). A key aspect of CRT is to avoid addressing symptom-related themes and instead utilize neutral stimuli in cognitive and behavioral exercises. This approach aims to establish a therapeutic alliance and reduce drop-out rates, particularly among individuals with AN. However, recent evidence suggests that CRT may not consistently improve central coherence abilities, cognitive flexibility, or symptoms associated with eating disorders (Hagan, Christensen, & Forbush, 2020; Tchanturia, Giombini, Leppanen, & Kinnaird, 2017). In response to these findings, Trapp et al. (2022) propose modifications to address practical challenges encountered in the application of CRT. They question the use of neutral stimuli and draw support from Beck's cognitive theory of depression (Beck & Alford, 2009). This proposition aligns with the hypothesis of our study. If further studies consistently demonstrate that maladaptive RL is context-dependent, it would necessitate a shift in intervention approaches.

There are few important limitations and questions for future research. 1) One aspect to consider is the use of symbolic rewards and punishments in our study, represented by images of a one euro coin and a barred representation of a one euro coin, respectively. These rewards and punishments were merely symbolic, and it is unclear how the use of concrete, non-symbolic rewards and punishments would impact the findings. Additionally, the 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.

Bernardoni, F., King, J. A., Geisler, D., Ritschel, F., Schwoebel, S., Reiter, A. M., … Ehrlich, S. (2021). More by stick than by carrot: A reinforcement learning style rooted in the medial frontal cortex in anorexia nervosa. *Journal of Abnormal Psychology*, *130*(7), 736–747.

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.

Heald, J. B., Lengyel, M., & Wolpert, D. M. (2023). Contextual inference in learning and memory. *Trends in Cognitive Sciences*.

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.

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 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.

Wagner, A., Aizenstein, H., Venkatraman, V. K., Fudge, J., May, J. C., Mazurkewicz, L., et al.others. (2007). Altered reward processing in women recovered from anorexia nervosa. *American Journal of Psychiatry*, *164*(12), 1842–1849.

Wierenga, C. E., Ely, A., Bischoff-Grethe, A., Bailer, U. F., Simmons, A. N., & Kaye, W. H. (2014). Are extremes of consumption in eating disorders related to an altered balance between reward and inhibition? *Frontiers in Behavioral Neuroscience*, *8*, 410.

Wu, M., Brockmeyer, T., Hartmann, M., Skunde, M., Herzog, W., & Friederich, H.-C. (2014). Set-shifting ability across the spectrum of eating disorders and in overweight and obesity: A systematic review and meta-analysis. *Psychological Medicine*, *44*(16), 3365–3385.