



Universiteit  
Leiden  
The Netherlands

## **Adapting to uncertainty: the role of anxiety and fear of negative evaluation in learning in social and non-social contexts**

Topel, S.; Ma, I.; Duijvenvoorde, A.C.K. van; van Steenbergen, H.; Bruijn, E.R.A. de

### **Citation**

Topel, S., Ma, I., Duijvenvoorde, A. C. K. van, Van Steenbergen, H., & Bruijn, E. R. A. de. (2024). Adapting to uncertainty: the role of anxiety and fear of negative evaluation in learning in social and non-social contexts. *Journal Of Affective Disorders*, 363, 310-319. doi:10.1016/j.jad.2024.07.066

Version: Publisher's Version

License: [Creative Commons CC BY 4.0 license](https://creativecommons.org/licenses/by/4.0/)

Downloaded from: <https://hdl.handle.net/1887/4054642>

**Note:** To cite this publication please use the final published version (if applicable).



Contents lists available at ScienceDirect

## Journal of Affective Disorders

journal homepage: [www.elsevier.com/locate/jad](http://www.elsevier.com/locate/jad)

Research paper

# Adapting to uncertainty: The role of anxiety and fear of negative evaluation in learning in social and non-social contexts

Selin Topel<sup>\*</sup>, Ili Ma, Anna C.K. van Duijvenvoorde, Henk van Steenbergen, Ellen R.A. de Bruijn

Leiden University, Institute of Psychology, Wassenaarseweg 52, 2333 AK Leiden, The Netherlands  
Leiden Institute for Brain and Cognition, Leiden, The Netherlands

## ARTICLE INFO

## Keywords:

Uncertainty  
Volatility  
Reinforcement learning  
Trait anxiety  
Fear of negative evaluation

## ABSTRACT

**Background:** Navigating social situations can be challenging due to uncertainty surrounding the intentions and strategies of others, which remain hidden and subject to change. Prior research suggests that individuals with anxiety-related symptoms struggle to adapt their learning in uncertain, non-social environments. Anxiety-prone individuals encounter challenges in social functioning, yet research on learning under uncertainty in social contexts is limited. In this preregistered study, we investigated whether individuals with higher levels of trait anxiety and fear of negative evaluation encounter difficulties in adjusting their learning rates in social contexts with stable or volatile outcome contingencies.

**Methods:** We implemented a modified trust game ( $N = 190$ ), where participants either retained or lost their investments based on their interactions with two players in volatile or stable environments. Participants also completed a matching non-social control task involving interactions with slot machines.

**Results:** Results from computational modeling revealed significantly higher learning rates in social compared to non-social settings. Trait anxiety did not affect the adaptability of learning rates. Individuals with heightened fear of negative evaluation were more sensitive to social compared to non-social outcomes, as reflected in their stay/switch behavior and, though less conclusive, in their learning rates.

**Limitations:** While transdiagnostic and dimensional approaches are important for investigating disturbed social functioning, the inclusion of clinical samples in future studies may contribute to a broader generalization of these findings regarding behavioral variances in uncertain social environments.

**Conclusions:** Individuals with increased fear of negative evaluation may demonstrate heightened sensitivity to learning in uncertain social contexts. This leads to heightened responsiveness to recent outcomes in their interactions with others, potentially contributing to their problems in social functioning.

## 1. Introduction

The world we live in is highly dynamic and deciding on the best course of action is a major human challenge in an uncertain and changing environment. This especially holds true in social situations where others' intentions often remain obscure and may evolve over time, thus increasing uncertainty (Beltzer et al., 2019; FeldmanHall and Shenhav, 2019). For example, an initial unpleasant interaction with a new neighbor may leave the impression that they are unfriendly. However, if we assume that the way this person feels and acts can change over time (i.e., their behavior is volatile), this may prompt us to pay attention to their recent behaviors. Learning from both distant and recent interactions is thus crucial for building and maintaining

relationships, shaping expectations, and adjusting our actions accordingly. Failure to do so might lead to missed opportunities for positive social connections or avoidance of negative ones (Lamba et al., 2020; Zabag et al., 2023). Individuals with anxiety problems also struggle with tolerating uncertainty and they often face difficulties in social functioning (Henning et al., 2007; Saris et al., 2017; Settiani and Kendall, 2013; Wood, 2006). For them, challenges in adapting their behaviors under uncertainty may exacerbate their difficulties with social functioning. Here, we aim to address this, by investigating the ability for learning under uncertainty in both non-social and social contexts and evaluate how anxiety-related traits are associated with these processes.

Learning under uncertainty through trial-and-error is formalized in reinforcement learning (RL) (Sutton and Barto, 1981; Yu and Dayan,

<sup>\*</sup> Corresponding author at: Leiden University, Wassenaarseweg 52, 2333 AK Leiden, the Netherlands.

E-mail address: [s.topel@fsw.leidenuniv.nl](mailto:s.topel@fsw.leidenuniv.nl) (S. Topel).

<https://doi.org/10.1016/j.jad.2024.07.066>

Received 8 January 2024; Received in revised form 24 June 2024; Accepted 14 July 2024

Available online 21 July 2024

0165-0327/© 2024 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

2005), applied initially to non-social and later also to social environments (for a review see Zhang et al., 2020). In simple RL models, *prediction errors*, signaling the difference between the expected and actual outcomes, guide learning. The extent to which prediction errors guide updating of our expectations is determined by the so-called *learning rate*. In stable environments, where action-outcome relationships remain consistent, having lower learning rates once an association is formed imply ideal behavior; whereas, volatile settings require constant adjustments reflected in higher learning rates (Behrens et al., 2007; Yu and Dayan, 2005).

Using a probabilistic aversive learning task, Browning et al. (2015) indeed reported higher learning rates in volatile relative to stable environments, but this volatility effect on learning rates was reduced in participants with high trait anxiety. Other studies have also linked trait anxiety (Pulcu and Browning, 2017) and negative affect (Gagne et al., 2020) to difficulties in adjusting learning rates when facing aversive or appetitive outcomes, suggesting challenges for individuals with anxiety-related problems in optimally updating information under uncertainty in non-social settings.

Yet, most of our decisions do not take place in a vacuum and involve a social context. Moreover, individuals with anxiety problems often face difficulties in social functioning. For example, studies have demonstrated that individuals with generalized anxiety disorders report impairments in social functioning (Henning et al., 2007) and behavioral and affective indicators of social functioning are impaired in patients with anxiety disorders (Saris et al., 2017). These findings stress the importance of investigating learning process in uncertain social contexts in individuals who are susceptible to anxiety problems, e.g., in relation to trait anxiety. Furthermore, in these contexts, more socially relevant transdiagnostic factors may influence how we adapt our behavior and respond to changing environments. Fear of negative evaluation (FNE) is a prominent trait in various internalizing-related problems and a key feature of social anxiety (Hong and Cheung, 2015; Poh et al., 2021), emphasizing its importance in understanding learning within uncertain social environments. However, despite the inherent uncertainty and volatility in social interactions, there is limited knowledge regarding how anxiety-related symptoms, specifically trait anxiety and FNE, are associated with adaptive learning in social contexts.

To our knowledge, only two recent studies examined anxiety-related differences in learning processes in volatile social contexts. In the first, participants played a virtual ball-tossing game with two other players (Beltzer et al., 2019). After an initial learning block (relatively stable phase), reversals occurred in the remainder of the task (volatile phase). Contrary to what the non-social tasks reported previously (Browning et al., 2015; Gagne et al., 2020), in the volatile vs. stable phases, socially anxious participants did not show difficulties adjusting their learning rates. Instead, the results showed valence-specific effects: Socially anxious participants did not adjust their learning rates (i.e. reducing their speed of updating) in response to players who steadily excluded them by not throwing the ball to them. This valence-specific effect on learning rates is suggestive of hypervigilance to a negative social stimulus. The second study employed an adapted trust game in non-social and social contexts (Lamba et al., 2020). Compared to non-anxious participants, participants with generalized anxiety symptoms invested relatively more in the social context following trust violations, leading them to be exploited when players became untrustworthy. The results of these two studies suggest that anxiety-related traits may indeed be related to different learning patterns in volatile social contexts. However, these studies primarily concentrated on volatile environments and lacked a matching stable environment for comparison. Moreover, it is difficult to directly compare the main outcomes of the two studies, given the very different paradigms used and focus on generalized versus social anxiety. Importantly, anxious participants in the trust game may actually learn the contingencies similarly, but they might be motivated to leave a positive impression or influence others by not reducing the entrusted amount. Thus, it is important to uncover the role of concerns

about being negatively evaluated by others in such anxiety-related findings.

In this study, we aimed to investigate the impact of anxiety-related traits on learning in both a social and a closely matched non-social task involving stable and volatile environments. Building on previous research (e.g., Behrens et al., 2007; Blain and Rutledge, 2020; Browning et al., 2015), our adapted task involved a sharing game framed in a social context, resembling a repeated trust game, and a matching non-social gambling game. In the social task, participants believed they were playing with two other participants in real-time, sharing points allocated at the start of the experiment. Successful performance required participants to learn which player would result in minimized losses when sharing points and consider the volatility in the environment (e.g., assessing whether an unexpected negative encounter was a random event or provided valuable information about player's character). The outcomes in both games were presented in a loss frame (loss vs. no change in endowment), creating an aversive learning environment in line with the most consistent anxiety-related findings (Browning et al., 2015; Gagne et al., 2020; Pulcu and Browning, 2017).

Following our preregistration ([https://osf.io/62mc7/?view\\_only=c5b6549a08e7422f850e3b7c7ad2426a](https://osf.io/62mc7/?view_only=c5b6549a08e7422f850e3b7c7ad2426a)) we predicted that individuals with higher trait anxiety show poorer adjustment in learning rates depending on the stability or volatility of uncertain social and non-social contexts (i.e., less increase in learning rates in volatile vs. stable environments). If confirmed these findings would indicate that anxiety predicts learning deficits across domains. We also investigated differences in social and non-social contexts in a more exploratory manner. Finally, we also explored the impact of individual differences in FNE which may impact learning in social environments in particular.

## 2. Method

### 2.1. Participants

For this online study, we recruited 250 healthy adults from Prolific.co, selecting fluent English speaking participants who had not participated in the pilot study and who did not have a history of head injury, mental health diagnoses, or neurological conditions. Technical issues prevented 30 participants from completing the task and were thus excluded from the analyses. Following preregistered criteria, nine participants were excluded for missing more than three consecutive trials and four for failing attention checks in the survey. Two participants were excluded as their behavior was not better explained by the models we fitted than by the random chance model. Two other participants were ineligible due to medication use, and five others timed out during participation. We deviated from the preregistered criterion for consecutive button presses: instead of using a fixed criterion of 10 consecutive trials, we used the data to determine extreme outliers. This led to the exclusion of eight participants who had the same response for >13 times in a row regardless of the stimulus shown. Given that there was a time limit to respond in the task, we decided not to exclude any participants or trials based on reaction times. Thus, we excluded data from 60 participants who were initially recruited for the study. The final sample included 190 participants (see Table 1 for demographics). Participants received monetary compensation and a potential £1.72 GBP bonus based on task performance. The study was approved by the Psychology Research Ethics Committee of Leiden University.

### 2.2. Materials

#### 2.2.1. Experimental tasks

Participants completed both a social sharing task and a matching non-social control task (i.e., gambling task) online, using a modified version of the probabilistic reversal learning task with stable and a volatile blocks (Behrens et al., 2007; Blain and Rutledge, 2020; Browning et al., 2015; Gagne et al., 2020).

**Table 1**

Table showing descriptive statistics of demographic and questionnaire information.

Variable	N	Mean (SD)	Median	Range
Age	190	24.64 (4.32)	24	18–36
Sex	Female: 95 Male: 95	–	–	–
Gender	Female: 93 Male: 95 Non-binary: 2	–	–	–
STAI-T	190	42.89 (11.97)	43	20–74
BFNE	190	36.72 (10.69)	37	13–60
DASS (D)	190	8.61 (8.78)	6	0–42

Note. STAI-T is the abbreviation for the trait anxiety subscale of State and Trait Anxiety Inventory. BFNE stands for Brief Fear of Negative Evaluation. DASS (D) stands for the depression subscale of the Depression Anxiety Stress Scales (Lovibond & Lovibond, 1995) for which the total scores can be between 0 and 42. Range shows the range of values observed in the current study.

To make the tasks more game-like and engaging, we let the participant choose their avatar from six options which is followed by either instructions for the social or non-social task depending on the counterbalancing. In the social task, participants were informed that they would interact with four other online players assuming roles of either a *decider* or a *responder*. As the *decider*, they chose which of the other two players they would share points with on each trial (e.g., in Fig. 1 right panel, if they choose the yellow player, they would be sharing 20 points with that player). The points shared were doubled when transferred to the chosen *responder*. According to the cover story, if assigned the role of the *responder* they decided whether to reciprocate by returning half of the points to the *decider* or keeping all the points. Responders would not know whether they were picked by the decider on a given trial and on every trial, both responders would need to decide whether to reciprocate in case they were chosen. The roles were determined through a rigged draw, which took place only once and right before the instructions for the social task were presented. Participants, unbeknownst to them, always played as the *decider* and did not interact with real players. To make the story more believable, we also added statements signaling that participants needed to wait for the other players to read the instructions and make decisions throughout the entire task.

In the non-social task, participants chose to bet on one of two slot machines, facing either a loss of the number of points assigned to their choice (i.e., in Fig. 1 left panel, a participant who bets on the green slot machine would lose 20 points) or no change in points (i.e., no loss). Each task comprised stable and volatile blocks consisting of 60 trials each (240 trials in total). Stimuli in the social task consisted of colorful avatars, while the non-social task featured colorful slot machines, with varying point assignments on each trial (see Supplementary Materials for details). Stable blocks in both tasks had one stimulus with a higher loss probability (i.e., 75 %) than the other stimulus (i.e., 25 %). In volatile blocks, the probabilities were 80 % and 20 % and these reversed every 15 trials (i.e., 3 times). In total, participants made decisions about four different stimulus pairs (two in the social and two in the non-social

task). Unique stimulus pairs were shown in each block. Self-reported affective state, uncertainty, and likeability ratings were collected per block (see Section 2.4.2 and Supplementary Materials for details).

### 2.2.2. Self-report measures

While trait anxiety was analyzed as planned, we decided not to analyze depressive symptoms due to limited symptom severity observations. Depression scores showed a highly right-skewed distribution in our healthy sample (see the summary statistics in Table 1). Fear of negative evaluation was included post-preregistration for its social relevance and was not part of our a priori analyses. Summary statistics of all self-report measures are reported in Table 1.

**2.2.2.1. Trait anxiety.** Trait anxiety was measured using the trait anxiety subscale of the State-Trait Anxiety Inventory (STAI-T; Spielberger et al., 1970; Spielberger, 1983). The trait anxiety subscale is a widely used 20-item scale that measures the non-disorder specific anxiety at a state and trait level, and it has been used in both clinical and non-clinical populations. Respondents were asked to use a four-item Likert scale ranging from 1 – *almost never* to 4 – *almost always* to indicate how they generally feel. Total scores range between 20 and 80. Example items include statements like “I worry too much over something that really doesn’t matter” or reverse-coded items like “I am calm, cool and collected.” The measure has been reported to have good reliability and validity and the reliability in the current sample is excellent ( $\alpha = 0.94$ ).

**2.2.2.2. Fear of negative evaluation.** Fear of negative evaluation (FNE) was measured using the 12-item Brief Fear of Negative Evaluation (BFNE; Leary, 1983). Respondents are asked to indicate how characteristic of them each statement would be on a five-point Likert scale (from 1 – *not at all characteristic of me* to 5 – *extremely characteristic of me*). Total scores range between 12 and 60. Example items include statements such as “I often worry that I will say or do the wrong things.” and reverse-coded items such as “I am unconcerned even if I know people are forming an unfavorable impression of me.” This scale showed excellent reliability ( $\alpha = 0.91$ ) in the current sample.

### 2.3. Procedure

Participants first provided informed consent and filled out a Qualtrics survey on inclusion criteria, demographics, and standardized questionnaires (see Supplementary Materials for details). Following the survey, participants were redirected to the experimental task, programmed using OSWeb (OpenSesame; Mathôt et al., 2012) and hosted on the Just Another Tool for Online Studies (JATOS).

Before starting the task, participants chose an avatar from the six available options with varying colors. Participants starting with the social task were informed they would be matched with other players and their role—*decider* or *responder*—would be determined by a draw. Those beginning with the non-social task directly read the instructions for the game. Both tasks included five practice trials.

Upon task completions, participants were redirected to another

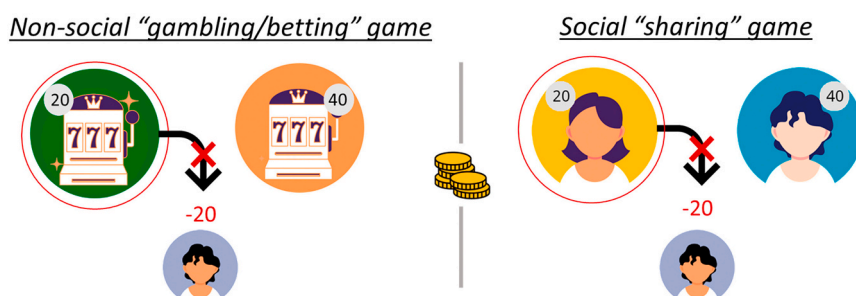


Fig. 1. Example showing how participants might be presented with a loss outcome of 20 points in both the non-social gambling task and socially-framed sharing task.

Qualtrics survey, prompting them to fill out a funnel debriefing including covering their strategy, perceptions of the experiment's purpose, and beliefs about playing with real players.

## 2.4. Analyses

The main focus of the current study was on differences in learning rate parameters estimated using a computational modeling approach. We also performed model-free analyses and investigated task-related differences in relevant self-report measures. R version 4.3.0 in RStudio 2022.12.0 was used for analyses.

We conducted model-free analyses using generalized linear mixed effects models (GLMMs) with binomial distribution and logit link using *glmer* function in the *lme4* package (Bates et al., 2015) to examine differences in binary choices people made. We used linear mixed effects models (LMMs) with *lmer* function for analyzing numeric outcome variables (i.e., learning rates, affect ratings). All models using trial-by-trial data specified the within-subject variables and their interaction as random slopes as well as fixed effects. These models were fit after setting the contrasts to sum-to-zero. All statistical analyses were also performed controlling for task and block order, but results remained consistent and we thus report only on the models without these variables. In addition, we conducted analyses on cell means using general linear models (GLM) for repeated measures using the *anova.test* function from the *rstatix* package (Kassambara, 2023). Note that we only report the outcomes of the (G)GLMMs that were replicated by the GLM approach given that effects that converge across the two approaches are considered more robust (Arnqvist, 2020). For transparency, we report all results from both sets of analyses in the supplementary materials. The significance of effects was estimated by using *mixed* function from the *afex* package (Blain and Rutledge, 2020; Singmann et al., 2023) for GLMMs. The significance of fixed effects in LMMs were estimated by fitting the models with *lmerTest* package and using the *anova* function. Model predictions were visualized using the *plot\_model* function in the *sjPlot* package (Lüdtke, 2023).

### 2.4.1. Behavioral analyses

First, we analyzed possible differences in choosing the option with lower probability of loss (i.e., the “best” option) using the factors Task (social vs. non-social) and Environment (stable vs. volatile) to describe participants' behavior under different conditions in the study. Second, we investigated stay-switch behavior following loss and no-loss outcomes using the factors Task, Environment, and Feedback (loss vs. no-loss obtained in the preceding choice). Increase in staying behavior following a positive outcome and switching following a negative outcome (more reactive responses to the most recent outcomes) can be thought of as a model-free proxy for higher learning rates within the reinforcement learning framework (for a similar approach see Blain and Rutledge, 2020). This was done by coding a binary variable for trials that were followed by a repeating choice as “1” and those followed by a switch as “0”. Last trials of each block were removed from this analysis. For the GLMMs we used the binary choices people made as dependent variables, for the GLM analyses, proportion of choices were used. In line with the preregistration, in subsequent analyses the centered and scaled continuous variables Trait Anxiety and FNE were added for the social and non-social tasks separately. In a more exploratory fashion, we also tested the specificity of possible differences more directly, by examining the role of FNE and Trait Anxiety in full-factorial models with the within-subject variables of Task, Environment, Feedback, and interactions between all variables.

### 2.4.2. Self-reported uncertainty and affective state

**2.4.2.1. Affective state.** Participants reported their affective state using the Affect Grid (Russell et al., 1989) before and after each task block.

This served as a manipulation check and gave more insight into the processes underlying behavior. The Affect Grid assessed affective states based on valence and arousal dimensions, with responses recorded via mouse clicks. Valence ranged from extremely unpleasant to extremely pleasant feelings, while arousal ranged from extremely high to extreme sleepiness. Mouse click coordinates were scaled into valence and arousal scores that ranged from 0 to 100. For the ratings of valence a score of 100 would indicate an extremely pleasant feeling and a score of 0 would indicate an extremely unpleasant feeling. For the arousal ratings, a score of 100 would correspond to extremely high arousal and a score of 0 would correspond to extreme sleepiness.

To analyze the changes in self-reported arousal and valence, as measured with the affect grid, we compared pre and post scores after each environment × task combination (i.e., social stable, social volatile, non-social stable, non-social volatile). These values were entered into LMMs that included the fixed effects Task, Environment, and Measurement Time (i.e., pre versus post as a factor). The LMMs included a random intercept for participants. Similarly, for the GLMs, we used within-subjects factors Task, Environment, and Measurement Time, and individual difference measures were entered as covariates of interest in separate analyses for the social and non-social tasks. Where we observed effects of Trait Anxiety or FNE in one of the tasks, we also carried out full-factorial analyses including Task as a variable to test for the specificity of that effect in a given context.

**2.4.2.2. Uncertainty.** After every block, participants were asked to report how uncertain they felt when making decisions on a seven-point Likert scale (not at all – very much). To test whether perceived/experienced uncertainty differed after each task and environment, we conducted a LMM with Task, Environment, and added individual difference measures separately as fixed effects for social and non-social tasks. The LMMs included a random intercept for participants. The GLMs also included Task and Environment as within-subjects factors and Trait Anxiety or FNE as covariates of interest in separate analyses for social and non-social tasks. Finally, to test the specificity of the findings related to individual differences in social vs. non-social tasks, we also tested full-factorial models including the FNE.

### 2.4.3. Computational modeling

**2.4.3.1. Main model.** Our main model is the same for the social and non-social task and is based on reinforcement learning where the decision-maker learns from experience through probability prediction errors (PPE; Sutton and Barto, 1981). Here, the PPE reflects the difference between the outcome (i.e., loss, or no-loss) and the estimated probability of loss (i.e.,  $p(\text{loss})$ ):

$$PPE_t = Outcome_t - p(\text{loss}_{stim1})_t \quad (1)$$

where *outcome* takes the value of 1 for loss and 0 for no-loss. Stimulus 1 (*stim1*) corresponds with a specific responder or a specific slot machine depending on the task.

At the start of the task, the decision-maker does not know the probability of loss for any of the stimuli and therefore marginalizes over the uniform distribution of all loss probabilities, leading to an initialization of  $p(\text{loss}) = 0.5$  for each stimulus. If stimulus 1 was chosen, then the outcome will update the estimation of  $p(\text{loss})$  for stimulus 1:

$$p(\text{loss}_{stim1})_{t+1} = p(\text{loss}_{stim1})_t + \alpha^*(PPE_t) \quad (2)$$

where  $p(\text{loss}_{stim1})_{t+1}$  is the updated belief about the probability of loss for stimulus 1. The speed of learning is modulated by  $\alpha$  which is a free parameter reflecting the learning rate that modulates the effect of the probability prediction error, such that large probability prediction errors will have a larger effect on the belief update if  $\alpha$  is higher. Note that if stimulus 2 was chosen, then the belief about the probability of loss for

stimulus 2 would be updated according to Eq. (2) for stimulus 2 instead of stimulus 1.

In the tasks, the decision-maker chooses between two stimuli (i.e., they choose between responder 1 or 2 in the trust game and in the gambling task they choose between slot machine 1 and 2). Which of the two stimuli is chosen is partially determined by the difference in loss probability associated with each stimulus at timepoint  $t$ :

$$P_{diff} = p(\text{loss}_{stim2})_t - p(\text{loss}_{stim1})_t \quad (3)$$

and partially by the normalized difference in monetary magnitude associated with each stimulus at timepoint  $t$ :

$$M_{diff} = \frac{\text{Magnitude}_{stim2t} - \text{Magnitude}_{stim1t}}{80} \quad (4)$$

The probability of choosing stimulus 1 is then given by a weighted combination of  $P_{diff}$  and  $M_{diff}$  in a Softmax function to convert the values to choice probabilities:

$$p(\text{choose}_{stim1})_t = \frac{1}{1 + e^{-\beta(\varphi P_{diff} + (1-\varphi)M_{diff})}} \quad (5)$$

where the extent to which the decision-maker's choice is guided by the  $P_{diff}$  or  $M_{diff}$  is determined by free parameter  $\varphi$  which takes a value between 0 and 1. If  $\varphi$  is larger than 0.5, then the choices are more determined by  $P_{diff}$  than  $M_{diff}$ . If  $\varphi$  is smaller than 0.5 then the choices are more guided by  $M_{diff}$  than  $P_{diff}$ . Free parameter  $\beta$  allows for decision noise.

Finally, the probability of choosing stimulus 2 is:

$$P_{choose(stim2)}_t = 1 - P_{choose(stim1)}_t \quad (6)$$

The model has six free parameters; one  $\alpha$  per condition (i.e., volatile and stable in social and non-social tasks),  $\beta$ , and  $\varphi$ .

**2.4.3.2. Alternative models.** We fitted different variations of the main model where 1) all free parameters, 2)  $\varphi$ 's or  $\eta$ 's (in the multiplicative version of the models, see Supplementary materials for details) along with the  $\alpha$ 's were systematically varied in each block per condition, 3) as well as a version of this model where the learning rates updated loss probabilities based on worse- and better-than-expected outcomes separately. While our main model combines the trade-off between  $P_{diff}$  and  $M_{diff}$  by a weighted sum based on the approach taken by Blain and Rutledge (2020), we also tested a multiplicative function as others suggested (Browning et al., 2015; Gagne et al., 2020). We furthermore tested a betrayal aversion model where loss magnitude modulated the belief updates (see Supplementary Materials for details).

**2.4.3.3. Model fitting and model comparisons.** In total, we fitted 10 models based on maximum likelihood estimation using the *nloptr* package in R. Models were fitted to individual subjects' data using the high-performance computing (HPC) facility, Academic Leiden Interdisciplinary Cluster Environment (ALICE). We then computed the Bayesian Information Criterion (BIC) to identify the best-fitting version of the additive models and the best version of multiplicative models. Next, we compared those two with each other using bootstrapped 95 % CI of the BIC differences. For parameter recovery see results Supplementary materials.

After identifying the winning model, to analyze the (log-transformed) learning rates estimated by fitting this model to each participant's choices we used an LMM with Environment and Task as fixed effects and a random intercept. Then, in line with our preregistered plan, we used LMMs for each task separately with Environment (stable vs. volatile), Trait Anxiety, and their interaction as fixed effects and a random intercept. In addition, we explored the role of FNE by replacing trait anxiety with BFNE scores. Lastly, the specificity of the link between individual differences and learning rates in social vs. non-social tasks

were tested in a more exploratory fashion by adding FNE in full-factorial models with Task and Environment.

### 3. Results

#### 3.1. Behavioral results

##### 3.1.1. Choosing the best option

There were significant main effects of Task and Environment (see Fig. 2A). Participants more often chose the best option in the social compared to the non-social task ( $B = -0.061$ ,  $OR = 0.94$ , 95 % CI [0.90, 0.98],  $p = .007$ ) as well as in the stable compared to the volatile environment ( $B = 0.325$ ,  $OR = 1.38$ , 95 % CI [1.32, 1.46],  $p < .001$ ). There was no significant interaction ( $p = .833$ ). Adding Trait Anxiety or FNE to the models for the social and non-social task did not result in any additional significant effects (all  $ps > .056$ ).

##### 3.1.2. Staying with the same stimulus on the next trial

Significant main effects of Feedback and Environment showed that participants were more likely to stay with the same option following a no-loss than a loss ( $B = -0.489$ ,  $OR = 0.61$ , 95 % CI [0.58, 0.65],  $p < .001$ ) and in the stable versus volatile environment ( $B = 0.101$ ,  $OR = 1.11$ , 95 % CI [1.07, 1.15],  $p < .001$ ; see Fig. 2B). Feedback also interacted with Task ( $B = 0.100$ ,  $OR = 1.11$ , 95 % CI [1.07, 1.14],  $p < .001$ ) such that the difference in staying behavior following loss and no-loss was more pronounced in the social ( $B = -1.178$ , 95 % CI [-1.32, -1.04],  $p < .001$ ) compared to the non-social task ( $B = -0.778$ , 95 % CI [-0.89, -0.66],  $p < .001$ ; see Fig. 2B). These findings suggest that participants in the social task adapt their behaviors following feedback which can be seen as model-free proxy of learning rate.

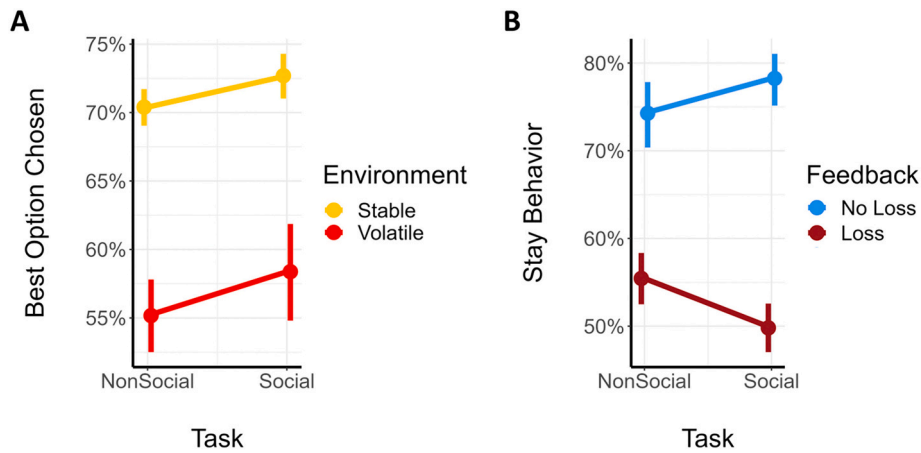
The impact of individual differences was tested for both task types separately. For the social task, we did not find converging significant findings for Trait Anxiety (for details see supplementary Tables S14–15; Fig. 3A). This means that we did not observe a significant interaction between Feedback and Trait Anxiety as opposed to what we would have predicted from the model-free indicators of our hypothesis related to trait anxiety and learning rates. Instead, when investigating the moderating effects of FNE, a significant interaction between FNE and Feedback was observed ( $B = -0.099$ ,  $OR = 0.91$ , 95 % CI [0.85, 0.94],  $p = .004$ ). Follow-up analyses showed that participants who score higher on FNE were more sensitive to feedback, such that they switched more following a loss outcome ( $B = -0.133$ , 95 % CI [-0.22, -0.05],  $p = .002$ ), whereas they tended to stay more following a no-loss although this association was not significant by itself ( $B = 0.066$ , 95 % CI [-0.08, 0.21],  $p = .361$ , Fig. 3B left panel).

In the non-social task, the model including Trait Anxiety showed a significant main effect of Trait Anxiety ( $B = -0.145$ ,  $OR = 0.86$ , 95 % CI [0.77, 0.97],  $p = .010$ ), showing that higher Trait Anxiety was associated with overall decreased stay behavior regardless of Feedback or Environment (Fig. 3A right panel). No further consistent main effects or interactions with Trait Anxiety or FNE were present (all  $ps > .09$ ).

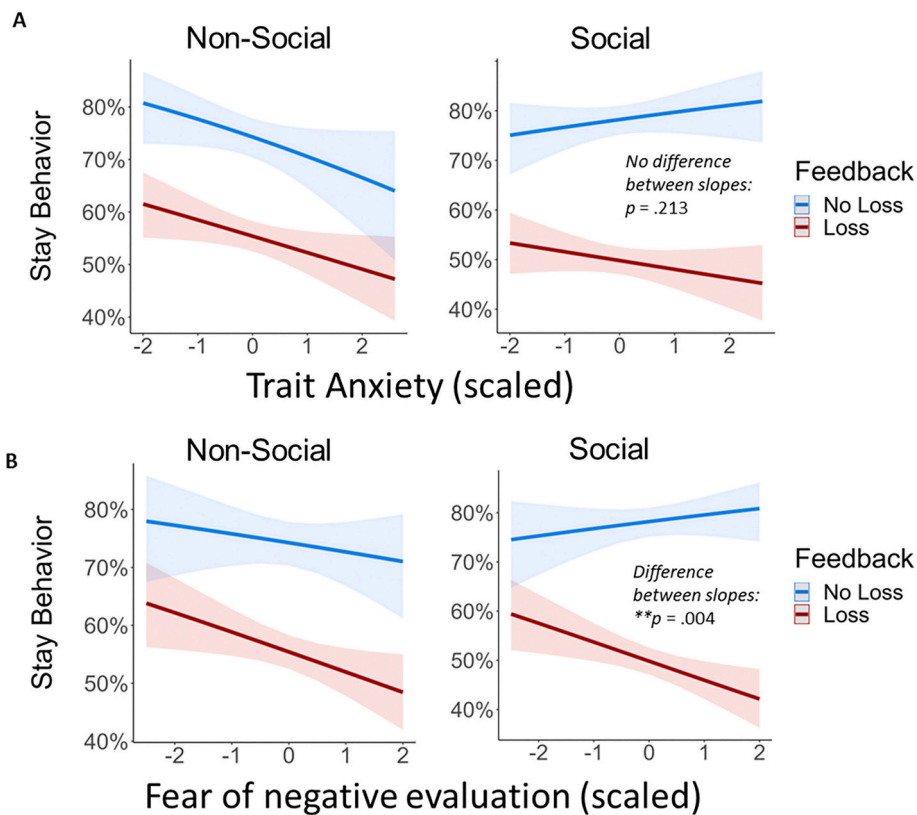
The different patterns between the social and non-social task in staying vs. switching behavior were then directly tested by running analyses with full-factorial models. There was a significant interaction between Feedback, Task, and FNE ( $B = 0.039$ ,  $OR = 1.04$ , 95 % CI [1.01, 1.08],  $p = .025$ ), supporting the pattern of results reported above. Our FNE-related findings were significantly different in the social compared to the non-social task. Moreover, we the full-factorial model with Trait Anxiety showed a significant interaction between Trait Anxiety and Task, confirming the results from our initial separate analyses ( $B = -0.061$ ,  $OR = 0.94$ , 95 % CI [0.89, 0.99],  $p = .026$ ).

#### 3.2. Computational modeling results

Our main computational model fitted best as shown by the 95 % CI of the BIC differences (see Supplementary Materials for model comparison,



**Fig. 2.** A) Plot showing the predicted proportion of choices favoring the option with lower chance of leading to a loss (“best” option) per Task and Environment by the GLMM. B) Plot showing the predicted proportion of “stay” choices in each task and following no-loss and loss by the GLMM. Error bar’s reflect confidence intervals.

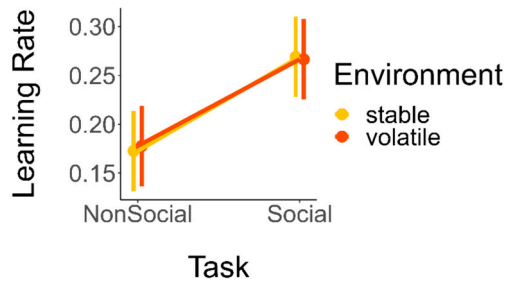


**Fig. 3.** A) Model predictions by the GLMM for differences in staying after loss vs. no loss in relation to Trait Anxiety in the non-social and social tasks on the left and right panels respectively. The slopes for Loss and No Loss feedback did not differ as a function of trait anxiety in either context. B) Model predictions by the GLMM for differences in staying after loss vs. no loss in relation to FNE in non-social and social tasks on the left and right panels respectively. The slopes for Loss and No Loss feedback only differed significantly as a function of FNE in the social context. The shaded areas reflect confidence intervals.

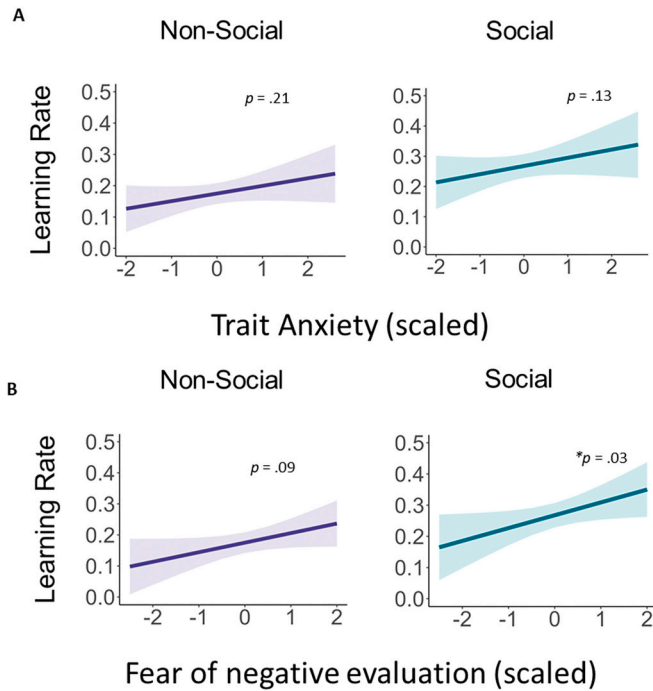
recoverability, and identifiability results). Thus, we report on the differences in learning rates estimated using this model.

Learning rates were higher in the social compared to the non-social task (Main effect of Task:  $B = -0.02$ , 95 % CI  $[-0.03, -0.02]$ ,  $p < .001$ ; Fig. 4), a finding consistent with the Feedback  $\times$  Task interaction we observed in the model-free behavioral analyses. None of the other main effects or the interaction between Environment and Task were significant (all  $ps > .59$ ). In accordance with our pre-registered analysis plan, we then tested the moderating effect of Trait anxiety separately for

the social and non-social task. Contrary to our expectations, we did not find an effect of Environment, nor did we find any interactions (or main effect) with Trait anxiety in either task (all  $ps > .13$ ; Fig. 5A). A similar analysis with FNE showed a main effect of FNE in the social task, such that participants who scored high on FNE employed higher learning rates ( $B = 0.02$ , 95 % CI  $[0.00, 0.04]$ ,  $p = .030$ ; see Fig. 5B), which dovetails with the model-free behavioral analyses revealing a FNE  $\times$  Feedback interaction for this task. None of the remaining main effects or interactions reached significance (all  $ps > .06$ ), nor were there any



**Fig. 4.** Plot showing the predicted effects by the LMM for the model-estimated learning rates in each environment and task. Error bars show the 95 % confidence intervals.



**Fig. 5.** A) Predicted effects for Trait Anxiety in the social and non-social task by the LMM. B) Predicted effects for FNE in the social and non-social task by the LMM. Shaded areas show the confidence intervals. \* =  $p < .05$ , ns.  $\geq 0.05$ .

significant effects for the non-social task ( $ps > .09$ ).

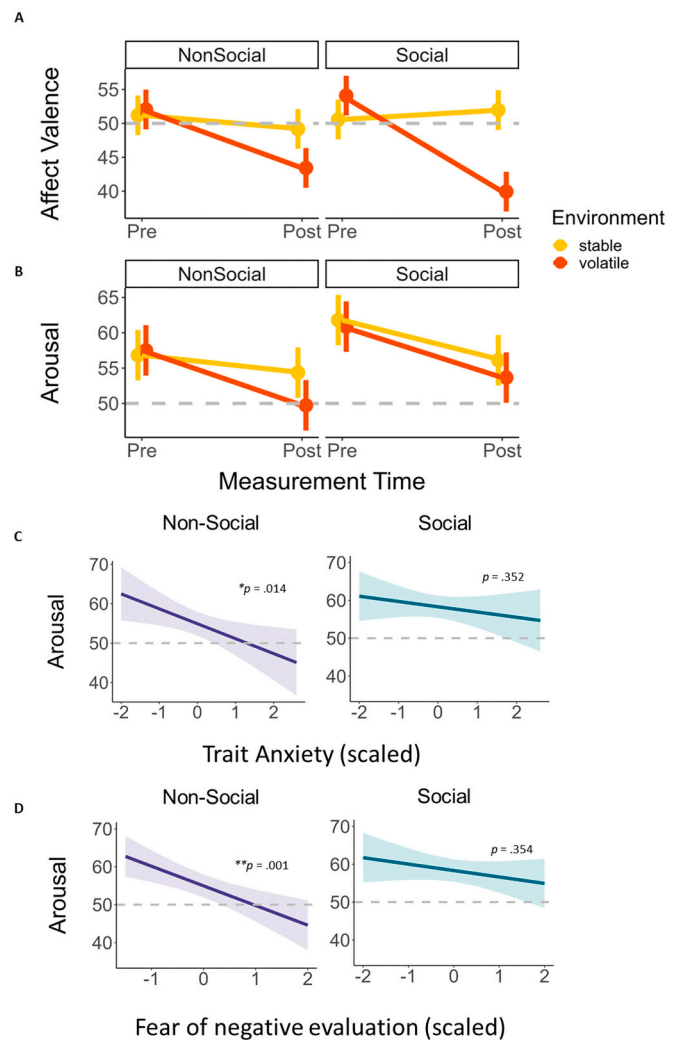
We tested the differences in the relationship between learning rates and FNE for the social and non-social task more directly in a full-factorial model (see supplementary Fig. S4B). This analysis revealed a main effect of FNE on learning rates ( $B = 0.02$ , 95 % CI [0.00, 0.03],  $p = .018$ ). There was no significant interaction between Task and FNE suggesting that the relationship between FNE and learning rates was not significantly different in the two tasks,  $B = -0.004$ , 95 % CI [-0.01, 0.00],  $p = .385$ . Instead, there was an interaction between Environment and FNE suggesting that the overall increased learning rates for FNE were more pronounced in the stable blocks compared to volatile blocks ( $B = 0.01$ , 95 % CI [0.00, 0.02],  $p = .042$ ), although this effect was not apparent when we examined the two tasks separately.

### 3.3. Changes in affect and differences in experienced uncertainty

For the self-reported affect (valence) measures, main effects of Environment ( $B = 1.68$ , 95 % CI [0.89, 2.46],  $p < .001$ ) and Measurement Time ( $B = 2.92$ , 95 % CI [2.14, 3.71],  $p < .001$ ) were present. Moreover, there was a significant interaction between Environment and Measurement Time ( $B = -2.77$ , 95 % CI [-3.55, -1.98],  $p < .001$ ), and

a 3-way interaction between Environment, Measurement Time, and Context ( $B = 1.12$ , 95 % CI [0.34, 1.90],  $p = .005$ ). Follow-up analyses showed that volatile ( $B = 11.38$ , 95 % CI [9.16, 13.60],  $p < .001$ ) but not stable environments ( $B = 0.31$ , 95 % CI [-1.90, 2.53],  $p = .781$ ) induced negative affect when comparing post-block to pre-block measures. This effect was larger in the social ( $B = -15.55$ , 95 % CI [-21.36, -9.74],  $p < .001$ ) compared to the non-social task ( $B = -6.58$ , 95 % CI [-12.40, -0.77],  $p = .019$ ; see Fig. 6A). When adding Trait Anxiety, a main effect of Trait Anxiety emerged in both the social ( $B = -3.78$ , 95 % CI [-5.95, -1.61],  $p < .001$ ) and non-social ( $B = -5.04$ , 95 % CI [-7.16, -2.93],  $p < .001$ ) tasks such that participants with higher levels of trait anxiety reported overall more negative mood. Similarly, instead of Trait Anxiety, when FNE was added, we observed that participants with higher FNE reported more negative mood in the social ( $B = -2.95$ , 95 % CI [-5.13, -0.76],  $p = .008$ ) and non-social tasks ( $B = -3.21$ , 95 % CI [-5.39, -1.02],  $p = .004$ ) alike. No other interactions with individual difference measures were found (all  $ps > .276$ ).

With regard to self-reported arousal (see Fig. 6B), we found main effects of Task ( $B = -1.73$ , 95 % CI [-2.62, -0.89],  $p < .001$ ),



**Fig. 6.** A) Plot showing the predicted effects by the LMM for the self-reported affective valence pre- and post-block in each environment and task. B) Plot showing the predicted effects by the LMM for the self-reported arousal pre- and post-block in each environment and task. C) Model predicted effects for Trait Anxiety on arousal in the social and non-social tasks. D) Model predicted effects for FNE on arousal in the social and non-social tasks. Error bars show the confidence intervals. The horizontal dashed line shows the neural point for affect valence and arousal.



Measurement Time ( $B = 2.89$ , 95 % CI [2.03, 3.75],  $p < .001$ ), and Environment ( $B = 0.92$ , 95 % CI [0.05, 1.78],  $p = .037$ ). Arousal was higher in the social vs. non-social task, in stable vs. volatile environments, and it decreased with time. No other interactions were significant (all  $p$ s  $> .052$ ). When adding Trait Anxiety and FNE separately in the models for social and non-social task, we found that neither was associated with arousal in the social task (all  $p$ s  $> .107$ ). Instead, for the non-social task Trait Anxiety ( $B = -3.78$ , 95 % CI [-6.79, -0.78],  $p = .014$ ; Fig. 6C) and FNE ( $B = -5.19$ , 95 % CI [-8.14, -2.23],  $p = .001$ ; Fig. 6D) were associated with overall lower levels of reported arousal. When tested with full-factorial analyses including the direct comparison between tasks, we observed a significant interaction of Task with both Trait Anxiety ( $B = -1.28$ , 95 % CI [-2.15, -0.41],  $p = .004$ ) and FNE ( $B = -1.71$ , 95 % CI [-2.57, -0.85],  $p < .001$ ), indicating that those with higher anxiety-related traits showed lower arousal levels in the non-social task, while there was no significant association between anxiety-related traits and arousal in the social task (see Supplementary Tables S46–49 for model outputs).

Analyses on self-reported uncertainty after each block in the experiment revealed a main effect of Environment, showing that volatile environments (Mean = 3.94, SD = 1.58) increased feelings of uncertainty relative to stable environments (Mean = 3.6, SD = 1.59;  $B = 0.17$ , 95 % CI [0.08, 0.26],  $p < .001$ ). This effect shows that the manipulation of volatility evoked feelings of uncertainty and that participants perceived the two environments, that were otherwise matched closely, differently. There were no Trait Anxiety effects in either task. However, higher FNE predicted increased overall uncertainty ( $B = 0.26$ , 95 % CI [0.07, 0.45],  $p = .008$ ), in the social task irrespective of the volatility of the environment. There was no effect of FNE on self-reported uncertainty in the non-social task. When compared directly in a full-factorial model, we found a main effect of FNE ( $B = 0.18$ , 95 % CI [0.01, 0.35],  $p = .035$ ) but we did not observe an interaction between FNE and Task on uncertainty ( $B = -0.077$ , 95 % CI [-0.16, 0.01],  $p = .079$ ) suggesting that these patterns were not significantly different in two tasks.

#### 4. Discussion

The aim of this study was to investigate the role of individual differences in trait anxiety and fear of negative evaluation (FNE) on learning to minimize or avoid negative outcomes in uncertain social and non-social contexts. We specifically tested the moderating role of trait anxiety and FNE in adjustment of learning rates depending on the volatility of the environment given prior findings demonstrating difficulties in these processes in relation to affective symptoms (Browning et al., 2015; Gagne et al., 2020; Pulcu and Browning, 2017) and its relevance for social contexts, respectively. To address these questions we employed a novel adaptation of a probabilistic learning task (Behrens et al., 2007; Blain and Rutledge, 2020; Browning et al., 2015; Gagne et al., 2020) resembling a trust game with human players and a non-social control task involving slot machines.

Behavioral, model-free, analyses showed that participants stayed more following positive feedback (i.e., no-loss) and switched more following negative feedback (i.e., loss) in the social vs. non-social context. This behavior aligned with higher learning rates in the winning RL model for the social compared to the non-social context. These results suggest heightened reactivity to immediate (positive and negative) feedback in social contexts, possibly due to increased emotional or motivational salience. Previously, it has been shown that indeed people pay more attention to and experience higher arousal when observing socially-relevant features than when they only view non-social cues (Rubo & Gamer, 2018). Consistently, our participants also reported higher arousal levels in the social task. Moreover, participants felt worse after completing the task in a highly unpredictable social context (i.e., volatile) but their mood did not decrease as much in the unpredictable non-social context. This suggests that, despite similarities in social and non-social learning mechanisms, distinctions in information processing

may arise due to the heightened motivational significance of social information.

Interestingly, FNE, rather than trait anxiety, was positively associated with learning rates in the current study. Although direct comparisons of learning rates between the social and non-social tasks did not reveal significant differences, Fig. 5B and separate analyses for each task suggest that this effect might be somewhat stronger in the social task. Notably, such an effect would largely correspond with the currently observed differences in the model-free behavior, where individuals with high FNE exhibited higher rates of staying after no-loss feedback and switching after loss feedback in the social task only, possibly indicating increased sensitivity to social outcomes. However, further studies are needed to explore a potential task-specific effect. Based on the current results, it thus seems possible that individuals with high FNE are more concerned by others' evaluations and sensitive to outcomes influenced by others' decisions, even in situations where their identity remains unknown to other "players". This pattern, however, may become more noticeable in situations where the identity of the players are known or where there is the possibility to meet the others. Future studies should focus on establishing whether in such cases, the distinction between the social and non-social contexts might be less implicit and FNE may have a more significant role. The heightened sensitivity that individuals with FNE show in social contexts could lead to increased avoidance following losses and immediate approach after positive outcomes, i.e., towards others who reciprocated their trust. Individuals with FNE are usually more sensitive to negative feedback from others, yet they may also fear that not reciprocating could result in disapproval or being disliked (e.g., 'avoidance' or 'impression-management' safety behaviors; Evans et al., 2021; Gilbert, 2014).

The win-stay lose-switch behavior observed in individuals with higher FNE may suggest a greater perception of volatility in social contexts compared to non-social ones. This model-free behavior, which is also largely reflected in learning rates, aligns with studies indicating that individuals with anxiety tend to be hypervigilant in uncertain environments, attributing random outcomes to changes (Beltzer et al., 2019; Huang et al., 2017; Piray and Daw, 2021). In the case of FNE, this would be especially noticeable within social settings due to the significance of this context for the construct. Future studies, using social tasks with conditions that are more directly relevant for FNE (e.g., participants are made aware that they would be evaluated by others) may be able to shed more light on this. Individuals with higher FNE also reported elevated uncertainty levels in the social task, highlighting the connection between subjective uncertainty and behavior in a social context. Our findings are consistent with previous studies on anxiety measures, including social anxiety and FNE, reporting increased sensitivity and avoidance of negative feedback (Harrewijn et al., 2018; Huang et al., 2017; Zhang et al., 2020). However, they differ from Lamba et al. (2020) where decreased sensitivity to recent negative outcomes was reported specifically in a social setting for those with generalized anxiety problems. Discrepancies may stem from differences in study designs, including experimental paradigms, anxiety measures used and instructions. Importantly, our FNE findings underscore the importance of context-relevant constructs in explaining variations in behavior. Note, however, that trait anxiety and FNE scores were highly correlated in our sample ( $r(188) = 0.58$ ,  $p < .001$ ) and although non-significant, trait anxiety showed a similar trend as FNE. This was also reflected in similar self-reported arousal patterns with higher trait scores associated with decreased arousal in the non-social context only.

Contrary to our pre-registered hypotheses, participants did not show differences in learning rate adjustments between stable and volatile environments in either social or non-social tasks, regardless of trait anxiety or FNE. Despite perceiving more uncertainty, participants did not employ higher learning rates in volatile environments in our task. This deviates from the findings of some previous studies (Behrens et al., 2007; Blain and Rutledge, 2020; Browning et al., 2015), although the absence of this effect was also reported in other samples (Hammond

et al., 2023). Notably, our task differed in design aspects, such as avoiding explicit instructions about stability-volatility (cf. Blain and Rutledge, 2020) to enhance ecological validity. Thus, it is possible that learning rate adjustments based on the volatility are partly dependent on task designs. However, the absence of a significant association between learning rate adjustments and trait anxiety/FNE in the current study aligns with other studies targeting learning under uncertainty and that did not find differences in learning rate adjustment in relation to affective symptoms (Beltzer et al., 2019; Blain and Rutledge, 2020; Hammond et al., 2023).

While our study has provided insights into individual differences in learning under uncertainty in social and non-social contexts, it is important to recognize and discuss its limitations. One limitation could be that we conducted an online study and used a general sample recruited on Prolific. This meant that despite including attention checks and cleaning the data based on a number of criteria we had limited control over the way participants completed the study potentially resulting noisier results. Thus, it would be informative to replicate the study in a more controlled laboratory settings. That being said, a large number of studies started making use of online platforms such as Prolific (Peer et al., 2021); and reported comparable results to laboratory studies (Gagne et al., 2020; Schidelko et al., 2021). Online platforms may also have some advantages for recruiting a preselected sample consisting of participants with good track record or administering tasks with a cover story involving other online players. As another limitation, although we initially planned to also examine differences related to depressive symptoms, we found that these analyses would not be very informative in our healthy sample where most participants scored low on depressive symptoms. Future studies may benefit from a targeted sampling approach that includes participants with a broader range of depressive symptom severity. Additionally, we adopted a dimensional approach to psychiatric symptoms, but individuals with a diagnosed anxiety disorder may not exhibit the observed patterns related to FNE or trait anxiety. Specifically, in this subclinical sample, higher FNE might be related to higher social motivation resulting from an increased concern about how other people view them and thus leading to an increased sensitivity to positive social feedback. Whereas, in individuals with a social anxiety disorder diagnosis, higher FNE might be related to higher sensitivity to negative and lower sensitivity to (or learning to a lesser degree from) positive social feedback (Harrewijn et al., 2018; Richey et al., 2019). Thus, recruiting a clinical group could reveal more complex and specific differences in behavior when learning in uncertain social contexts. Our task design did not allow us to determine performance differences related to learning about the “better” and “worse” stimuli. Future studies could modify the task to have participants independently learn about the better and worse players and non-social stimuli which may also reveal more valence-specific biases in learning in relation to anxiety and FNE (Beltzer et al., 2019; Koban et al., 2017). Lastly, our findings suggest more specific differences related to FNE but not trait anxiety. Recruiting larger samples with various related measures may help identify symptom dimensions, especially those more specific to social contexts, given that FNE is a transdiagnostic factor for multiple affective problems (i.e., Gillan et al., 2016). Relatedly, it would be interesting to test the possibility that individuals who are higher on a more general social sensitivity dimension (e.g., common to empathy, hypermentalizing, fear of positive evaluation, social anxiety), not just FNE, show more specific differences in the way they experience and learn in the social task.

To our knowledge, this is the first study to investigate anxiety-related differences in learning under uncertainty, comparing stable vs. volatile environments and social vs. non-social contexts in a factorial within-subject design. While trait anxiety did not exhibit expected patterns in adjusting learning rates to the volatility of an environment, clear context-related differences emerged, suggesting that social outcomes are more motivationally engaging and facilitate learning. Importantly, our findings support the idea that individuals with higher FNE show increased reactivity to recent social outcomes. These findings are

relevant for understanding and treating problems with social functioning, such as social anxiety. By addressing these behavioral patterns in interventions, individuals with excessive fear of being negatively evaluated by others can become aware of their hypervigilance or excessive reactivity to unexpected outcomes, fostering better adaptation to encountered uncertainty.

## Funding

ST, HvS, EdB, AvD are supported by the interdisciplinary research program ‘Social Resilience and Security’ of Leiden University.

## CRedit authorship contribution statement

**Selin Topel:** Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Ili Ma:** Writing – review & editing, Writing – original draft, Supervision, Software, Methodology, Investigation, Formal analysis. **Anna C.K. van Duijvenvoorde:** Writing – review & editing, Writing – original draft, Supervision, Methodology, Investigation, Conceptualization. **Henk van Steenbergen:** Writing – review & editing, Writing – original draft, Visualization, Supervision, Software, Methodology, Investigation, Conceptualization. **Ellen R.A. de Bruijn:** Writing – review & editing, Writing – original draft, Supervision, Project administration, Methodology, Investigation, Conceptualization.

## Declaration of competing interest

None.

## Acknowledgements

We would like to thank Bastien Blain for valuable discussions.

## Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used ChatGPT in order to paraphrase some of the statements and sentences that appeared in their protocol documents and previous drafts of this work. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the publication.

## Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.jad.2024.07.066>.

## References

- Arnqvist, G., 2020. Mixed models offer no freedom from degrees of freedom. *Trends Ecol. Evol.* 35 (4), 329–335. <https://doi.org/10.1016/j.tree.2019.12.004>.
- Bates, D., Mächler, M., Bolker, B., Walker, S., 2015. Fitting linear mixed-effects models using lme4. *J. Stat. Softw.* 67 (1) <https://doi.org/10.18637/jss.v067.i01>.
- Behrens, T.E.J., Woolrich, M.W., Walton, M.E., Rushworth, M.F.S., 2007. Learning the value of information in an uncertain world. *Nat. Neurosci.* 10 (9), 1214–1221. <https://doi.org/10.1038/nn1954>.
- Beltzer, M.L., Adams, S., Beling, P.A., Teachman, B.A., 2019. Social anxiety and dynamic social reinforcement learning in a volatile environment. *Clin. Psychol. Sci.* 7 (6), 1372–1388. <https://doi.org/10.1177/2167702619858425>.
- Blain, B., Rutledge, R.B., 2020. Momentary subjective well-being depends on learning and not reward. *eLife* 9, e57977. <https://doi.org/10.7554/eLife.57977>.
- Browning, M., Behrens, T.E., Jocham, G., O'Reilly, J.X., Bishop, S.J., 2015. Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nat. Neurosci.* 18 (4), 590–596. <https://doi.org/10.1038/nn.3961>.
- Evans, R., Chiu, K., Clark, D.M., Waite, P., Leigh, E., 2021. Safety behaviours in social anxiety: An examination across adolescence. *Behav. Res. Ther.* 144, 103931 <https://doi.org/10.1016/j.brat.2021.103931>.
- FeldmanHall, O., Shenhav, A., 2019. Resolving uncertainty in a social world. *Nat. Hum. Behav.* 3 (5), 426–435. <https://doi.org/10.1038/s41562-019-0590-x>.

- Gagne, C., Zika, O., Dayan, P., Bishop, S.J., 2020. Impaired adaptation of learning to contingency volatility in internalizing psychopathology. *eLife* 9, e61387. <https://doi.org/10.7554/eLife.61387>.
- Gilbert, P., 2014. Evolutionary models: practical and conceptual utility for the treatment and study of social anxiety disorder. In: Weeks, J.W. (Ed.), *The Wiley Blackwell Handbook of Social Anxiety Disorder*, 1st ed. Wiley, pp. 24–52. <https://doi.org/10.1002/9781118653920.ch2>.
- Gillan, C.M., Kosinski, M., Whelan, R., Phelps, E.A., Daw, N.D., 2016. Characterizing a psychiatric symptom dimension related to deficits in goal-directed control. *eLife* 5, e11305. <https://doi.org/10.7554/eLife.11305>.
- Hammond, D., Xu, P., Ai, H., Van Dam, N.T., 2023. Anxiety and depression related abnormalities in socio-affective learning. *J. Affect. Disord.* 335, 322–331. <https://doi.org/10.1016/j.jad.2023.05.021>.
- Harrewijn, A., Van Der Molen, M.J.W., Van Vliet, I.M., Tissier, R.L.M., Westenberg, P.M., 2018. Behavioral and EEG responses to social evaluation: a two-generation family study on social anxiety. *NeuroImage: Clin.* 17, 549–562. <https://doi.org/10.1016/j.nicl.2017.11.010>.
- Henning, E.R., Turk, C.L., Mennin, D.S., Fresco, D.M., Heimberg, R.G., 2007. Impairment and quality of life in individuals with generalized anxiety disorder. *Depress. Anxiety* 24 (5), 342–349. <https://doi.org/10.1002/da.20249>.
- Hong, R.Y., Cheung, M.W.-L., 2015. The structure of cognitive vulnerabilities to depression and anxiety: evidence for a common core etiologic process based on a meta-analytic review. *Clin. Psychol. Sci.* 3 (6), 892–912. <https://doi.org/10.1177/2167702614553789>.
- Huang, H., Thompson, W., Paulus, M.P., 2017. Computational dysfunctions in anxiety: failure to differentiate signal from noise. *Biol. Psychiatry* 82 (6), 440–446. <https://doi.org/10.1016/j.biopsych.2017.07.007>.
- Kassambara, A., 2023. *rstatix: Pipe-Friendly Framework for Basic Statistical Tests*. R package version 0.7.2. <https://rpkgs.datanovia.com/rstatix/>.
- Koban, L., Schneider, R., Ashar, Y.K., Andrews-Hanna, J.R., Landy, L., Moscovitch, D.A., Wager, T.D., Arch, J.J., 2017. Social anxiety is characterized by biased learning about performance and the self. *Emotion* 17 (8), 1144–1155. <https://doi.org/10.1037/emo0000296>.
- Lamba, A., Frank, M.J., FeldmanHall, O., 2020. Anxiety impedes adaptive social learning under uncertainty. *Psychol. Sci.* 31 (5), 592–603. <https://doi.org/10.1177/0956797620910993>.
- Leary, M.R., 1983. A brief version of the fear of negative evaluation scale. *Pers. Soc. Psychol. Bull.* 9 (3), 371–375. <https://doi.org/10.1177/0146167283093007>.
- Lovibond, S.H., Lovibond, P.F., 1995. *Manual for the Depression Anxiety Stress Scales*, 2nd. Ed. Psychology Foundation, Sydney.
- Lüdtke, D., 2023. *sjPlot: Data Visualization for Statistics in Social Science*. R package version 2.8.15. <https://CRAN.R-project.org/package=sjPlot>.
- Mathôt, S., Schreij, D., Theeuwes, J., 2012. OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behav. Res. Methods* 44, 314–324. <https://doi.org/10.3758/s13428-011-0168-7>.
- Peer, E., Rothschild, D., Gordon, A., Evernden, Z., Damer, E., 2021. Data quality of platforms and panels for online behavioral research. *Behav. Res. Methods* 54 (4), 1643–1662. <https://doi.org/10.3758/s13428-021-01694-3>.
- Piray, P., Daw, N.D., 2021. A model for learning based on the joint estimation of stochasticity and volatility. *Nat. Commun.* 12 (1), 6587. <https://doi.org/10.1038/s41467-021-26731-9>.
- Poh, R.Y.N., Zhuang, S., Ong, X.L., Hong, R.Y., 2021. Evaluating structural models of cognitive vulnerabilities: transdiagnostic and specific pathways to internalizing symptoms. *Assessment* 28 (6), 1635–1655. <https://doi.org/10.1177/1073191120915287>.
- Pulcu, E., Browning, M., 2017. Affective bias as a rational response to the statistics of rewards and punishments. *eLife* 6, e27879. <https://doi.org/10.7554/eLife.27879>.
- Richey, J.A., Brewer, J.A., Sullivan-Toole, H., Streve, M.V., Kim-Spoon, J., White, S.W., Ollendick, T.H., 2019. Sensitivity shift theory: a developmental model of positive affect and motivational deficits in social anxiety disorder. *Clin. Psychol. Rev.* 72, 101756. <https://doi.org/10.1016/j.cpr.2019.101756>.
- Rubo, M., Gamer, M., 2018. Social content and emotional valence modulate gaze fixations in dynamic scenes. *Sci. Rep.* 8 (1), 3804. <https://doi.org/10.1038/s41598-018-22127-w>.
- Russell, J.A., Weiss, A., Mendelsohn, G.A., 1989. Affect grid: a single-item scale of pleasure and arousal. *J. Pers. Soc. Psychol.* 57 (3), 493–502. <https://doi.org/10.1037/0022-3514.57.3.493>.
- Saris, I.M.J., Aghajani, M., Van Der Werff, S.J.A., Van Der Wee, N.J.A., Penninx, B.W.J. H., 2017. Social functioning in patients with depressive and anxiety disorders. *Acta Psychiatr. Scand.* 136 (4), 352–361. <https://doi.org/10.1111/acps.12774>.
- Schidlo, L.P., Schünemann, B., Rakoczy, H., Proft, M., 2021. Online testing yields the same results as lab testing: a validation study with the false belief task. *Front. Psychol.* 12, 703238. <https://doi.org/10.3389/fpsyg.2021.703238>.
- Settipani, C.A., Kendall, P.C., 2013. Social functioning in youth with anxiety disorders: association with anxiety severity and outcomes from cognitive-behavioral therapy. *Child Psychiatry Hum. Dev.* 44 (1), 1–18. <https://doi.org/10.1007/s10578-012-0307-0>.
- Singmann, H., Bolker, B., Westfall, J., Aust, F., Ben-Shachar, M., 2023. *afex: Analysis of Factorial Experiments*. R package version 1.3-0. <https://CRAN.R-project.org/package=afex>.
- Spielberger, C.D., 1983. *Manual for the State-Trait Inventory STAI (Form Y)*. Mind Garden, Palo Alto, CA. <https://doi.org/10.1037/t06496-000>.
- Spielberger, C.D., Gorsuch, R.L., Lushene, R.E., 1970. *Manual for the State-Trait Anxiety Inventory*. Consulting Psychologists Press, Palo Alto, CA.
- Sutton, R.S., Barto, A.G., 1981. Toward a modern theory of adaptive networks: expectation and prediction. *Psychol. Rev.* 88 (2), 135–170. <https://doi.org/10.1037/0033-295X.88.2.135>.
- Wood, J., 2006. Effect of anxiety reduction on children's school performance and social adjustment. *Dev. Psychol.* 42 (2), 345–349. <https://doi.org/10.1037/0012-1649.42.2.345>.
- Yu, A.J., Dayan, P., 2005. Uncertainty, neuromodulation, and attention. *Neuron* 46 (4), 681–692. <https://doi.org/10.1016/j.neuron.2005.04.026>.
- Zabag, R., Azoulay, R., Rinck, M., Becker, E., Levy-Gigi, E., Gilboa-Schechtman, E., 2023. You never get a chance to undo a negative first impression: social anxiety is associated with impaired positive updating of social information. *Personal. Individ. Differ.* 203, 111993. <https://doi.org/10.1016/j.paid.2022.111993>.
- Zhang, H., Shi, Y., Zhou, Z.E., Ma, H., Tang, H., 2020. Good people do bad things: how anxiety promotes unethical behavior through intuitive and automatic processing. *Curr. Psychol.* 39 (2), 720–728. <https://doi.org/10.1007/s12144-018-9789-7>.