Pupil dilation during negative prediction errors is related to brain choline concentration and depressive symptoms in adolescents

Mona Guath, Johan Lundin Kleberg, Jan Weis, Ebba Widegren, Matilda Frick, Stefan Möller, Lisa Klevebrant, Barry Karlsson, David Fällmar, Johanna Mårtensson, Daniel S Pine, Karin Brocki, Malin Gingnell, Andreas Frick

PII: S0166-4328(22)00328-X
DOI: https://doi.org/10.1016/j.bbr.2022.114060
Reference: BBR114060

To appear in: Behavioural Brain Research

Received date: 3 February 2022
Revised date: 10 August 2022
Accepted date: 14 August 2022

Please cite this article as: Mona Guath, Johan Lundin Kleberg, Jan Weis, Ebba Widegren, Matilda Frick, Stefan Möller, Lisa Klevebrant, Barry Karlsson, David Fällmar, Johanna Mårtensson, Daniel S Pine, Karin Brocki, Malin Gingnell and Andreas Frick, Pupil dilation during negative prediction errors is related to brain choline concentration and depressive symptoms in adolescents, Behavioural Brain Research, (2022) doi:https://doi.org/10.1016/j.bbr.2022.114060

This is a PDF file of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability, but it is not yet the definitive version of record. This version will undergo additional copyediting, typesetting and review before it is published in its final form, but we are providing this version to give early visibility of the article. Please note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2022 Published by Elsevier.
Pupil dilation during negative prediction errors is related to brain choline concentration and depressive symptoms in adolescents

Mona Guath\textsuperscript{a,∗}, Johan Lundin Kleberg\textsuperscript{b,c,∗}, Jan Weis\textsuperscript{d}, Ebba Widegren\textsuperscript{e}, Matilda Frick\textsuperscript{a,e}, Stefan Möller\textsuperscript{c,f}, Lisa Klevebrant\textsuperscript{a,e}, Barry Karlsson\textsuperscript{e}, David Fällmar\textsuperscript{g}, Johanna Mårtensson\textsuperscript{g}, Daniel S Pine\textsuperscript{h}, Karin Brocki\textsuperscript{a}, Malin Gingnell\textsuperscript{a,e,§}, Andreas Frick\textsuperscript{i}

\textsuperscript{a}Department of Psychology, Uppsala University, Uppsala, Sweden
\textsuperscript{b}Centre for Psychiatry Research, Department of Clinical Neuroscience, Karolinska Institute & Stockholm health care services, Stockholm, Sweden
\textsuperscript{c}Department of Molecular Medicine and Surgery, Karolinska Institute, Stockholm, Sweden
\textsuperscript{d}Department of Medical Physics, Uppsala University Hospital, Uppsala, Sweden
\textsuperscript{e}Department of Medical Sciences, Psychiatry, Uppsala University, Uppsala, Sweden
\textsuperscript{f}Department of Psychology, Lund University, Lund, Sweden
\textsuperscript{g}Department of Surgical Sciences, Radiology, Uppsala University Hospital, Uppsala, Sweden
\textsuperscript{h}National Institute of Mental Health (NIMH), Bethesda, Maryland, USA
\textsuperscript{i}The Beijer Laboratory, Department of Medical Sciences, Psychiatry, Uppsala University, Uppsala, Sweden

* Joint first author, § Equal contribution

Correspondence to:
Mona Guath
Department of Psychology, Uppsala University
751 42 Uppsala
Sweden
Email: mona.guath@psyk.uu.se

Abstract (217 words)
Depressive symptoms are associated with altered pupillary responses during learning and reward prediction as well as with changes in neurometabolite levels, including brain concentrations of choline, glutamate and gamma-aminobutyric acid (GABA). However, the full link between depressive symptoms, reward-learning-related pupillary responses and neurometabolites is yet to be established as these constructs have not been assessed in the same individuals. The present pilot study, investigated these relations in a sample of 24 adolescents aged 13 years. Participants completed the Revised Child Anxiety and Depression Scale (RCADS) and underwent a reward learning task while measuring pupil dilation and a single voxel dorsal anterior cingulate cortex (dACC) MEGA-PRESS magnetic resonance spectroscopy scan assessing choline, glutamate and GABA concentrations. Pupil dilation was related to prediction errors (PE) during learning, which was captured by a prediction error-weighted pupil dilation response index (PE-PDR) for each individual. Higher PE-PDR scores, indicating larger pupil dilations to negative prediction errors, were related to lower depressive symptoms and lower dACC choline concentrations. Dorsal ACC choline was positively associated with depressive symptoms, whereas glutamate and GABA were not related to PE-PDR or depressive symptoms. The findings support notions of cholinergic involvement in depressive symptoms and cholinergic influence on reward-related pupillary response, suggesting that pupillary responses to negative prediction errors may hold promise as a biomarker of depressive states.

Keywords: Magnetic resonance spectroscopy, mood disorders, reward learning, operant conditioning
Introduction

Major depressive disorder (MDD) is one of the most common mental disorders worldwide and a major cause of negative life outcome [21]. The onset of MDD is frequently seen in adolescence and typically begins with subclinical symptoms. Although not qualifying for a depressive diagnosis by themselves, subclinical symptoms are often impairing and may develop into a full-blown depressive episode [9]. Moreover, childhood and adolescent depressive symptoms are associated with recurring depressive episodes, psychiatric comorbidity and poor functional outcome [5,19]. These findings highlight the importance of applying a dimensional perspective to depressive symptoms and studying correlates of depressive symptoms in adolescents.

Both psychological and biological processes relate to depressive symptoms. Consistent with the core feature of anhedonia, depressive patients show altered reward processing, including deficient learning of reward contingencies and blunted striatal response to reward receipt [1,10]. This may be linked to failures in processing prediction errors (PE), the difference between expected and received outcomes [7,20], because deficient PE processing impairs learning. Importantly, individuals with and without depression may manifest similar behavioral markers of PEs [16] while still exhibiting perturbations in the consequences and neural representations of PEs. For example, in depressed individuals negative PEs have larger impact on memory than positive PEs [16]. This may result in a stronger neural response in brain regions associated with prediction errors and learning of rewards/punishments such as the anterior cingulate cortex (ACC), striatum, nucleus accumbens and amygdala [7,20]. Understanding the complex nature of relations between PEs and depressive symptoms may therefore be a key step to designing better treatments targeting anhedonia.

Pupillometry and magnetic resonance spectroscopy (MRS) provide tools for relating PE and depressive symptoms to neurobiology. In terms of pupillometry, Schneider et al. found that individuals with more pronounced depressive symptoms display smaller pupil dilation during anticipation of reward, possibly reflecting low arousal to upcoming rewards [17]. In terms of MRS, prior studies relate depressive symptoms to levels of brain choline, glutamate, and gamma-aminobutyric acid (GABA) [11,15]. However, no study combines assessments of reward-related pupillary responses with brain levels of neurotransmitters as the measures relate to depressive
Symptoms. Building on previous work in these areas, this pilot study examines concurrent relations in adolescents among depressive symptoms, pupillary responses during reward learning, and brain neurochemicals.

**Material and methods**

**Participants**

Twenty-four 13-year-old adolescents (14 girls, 10 boys) were recruited from a longitudinal study [6]. Of these, 22 completed the measure of depressive symptoms and 21 the MRS measures of brain metabolites, and 18 the reward-learning paradigm. The present study only used the longitudinal study as a recruitment base and we did not include any of the measures from previous visits. The study was approved by the Swedish Ethical Review Authority. All participants and their parents provided written assent and consent respectively after receiving written and verbal information about the study.

**Procedure**

Parents of participants in the longitudinal study were contacted through mail with an invitation to participate in the ongoing project Fear Extinction in Adolescence and its Regulation (FEAR) including information about the study and the address of a web site where they enrolled and answered screening questions. Of the 99 contacted, 48 volunteered to participate. Of these, 24 were excluded due to meeting exclusion criteria (n=17) or dropped out (n=7) of the study. Exclusion criteria included hearing impairment, uncorrected visual impairment; presence or history of severe psychiatric disorder, somatic or neurological (n=2) conditions or ongoing medication (n=7); atypical development (n=2); ongoing illicit drug use; magnetic safety issues (n=3) and metal braces (n=3). Individuals passing the initial screening criteria underwent the Mini International Neuropsychiatric Interview, children’s version (MINI-KID) [18] to assess presence of psychiatric disorders. Individuals not fulfilling criteria for any psychiatric disorder were included and answered questionnaires online, including the Revised Child Anxiety and Depression Scale (RCADS) [4] used in this study.

Participants underwent a neuroimaging session including T1-weighted anatomical imaging and MRS acquisition, and subsequently a reward-learning task. Participants completed the tasks as
part of a larger study including other experiments not reported here. A specialist in neuroradiology (DF) reviewed all T1 images to rule out focal lesions and other incidental findings.

**Questionnaires**

Participating children rated their level of depressive symptoms on the 11 items MDD subscale of the RCADS [4]. This scale has good psychometric properties [3] and discriminates well between MDD symptoms and symptoms of other disorders.

**Magnetic resonance imaging and spectroscopy acquisition**

T1-weighted anatomical and MRS data were acquired on a 3 Tesla MR scanner (Achieva dStream, Philips Healthcare, Best, The Netherlands) using a 32-channel head coil. T1-weighted gradient echo images were acquired with the following parameters (repetition time (TR) 8.2 ms, echo time (TE) 3.8 ms, matrix=256×256×220, acquisition voxel size=0.94×0.4×1 mm³). The spectra were acquired by a J-difference Mescher-Garwood spectral editing sequence MEGA-PRESS [13] with TR/TE 2000/68 ms, spectral bandwidth 2000 Hz, 1024 points, and phase cycling 4. Thirty groups of total 120 pairs of ON and OFF spectra were measured. Acquisition in each group began with one unsuppressed water line followed by four pairs of water suppressed ON and OFF spectra. The water lines’ positions were used for magnetic field drift correction and for updating the reference frequency of radiofrequency (RF) pulses in each group. The voxel size (volume of interest) was 40 x 40 x 20 mm³ in the left-right, anterior-posterior and feet-head directions, respectively, and positioned in the dorsal anterior cingulate cortex (dACC) (Figure 1).

**Magnetic resonance spectroscopy analyses**

GABA+ (GABA + macromolecules) to total creatine (tCr) spectral intensity ratio at 3 ppm was quantified with the software package Gannet 3.0 (Edden et al., 2014). The spectrum processing contained Gaussian line broadening (3 Hz), frequency, phase, and base line corrections. Distorted spectra were omitted in a pairwise (ON-OFF) fashion. A single Gaussian was used to fit GABA+ intensity and tCr spectral line was fitted by the Lorentzian in the summed OFF spectrum. The
percentage ratio of standard deviation (SD) of the residue to the amplitude of the GABA+ was used as a measure of the fitting error.

Concentration ratios total choline (tCho)/tCr and glutamate (Glu)/tCr were quantified from summed OFF spectrum using the LCModel spectrum processing (Provencher, 1993). The input file for LCModel was acquired in two steps: (i) the free induction decays (FIDs) of all OFF spectra were manually selected, averaged using the jMRUI software [14] and exported as a text file; (ii) this text file was then converted by an in-house program to the “RAW” format required by LCModel. No apodization was used in this spectrum preprocessing. The LCModel algorithm provides the standard errors estimate called Cramér-Rao lower bound (CRLB) expressed in percent of the estimated concentration.

**Figure 1.** (A) Magnetic resonance spectroscopy voxel position on the dorsal anterior cingulate cortex overlaid on a standard anatomical brain image. (B) Representative OFF spectrum and fitting results using magnetic resonance spectroscopy models using LCModel. (a) Residue, (b) phase and baseline corrected measured spectrum, (c) creatine, (d) choline, and (e) glutamate. In total, weighted combination of 25 spectra of different brain metabolites, macromolecules and mobile lipids were used in fitting the measured spectrum. (C) Representative MEGA-PRESS spectrum and fitting results of gamma-aminobutyric acid (GABA) using GANNET.

**Reward-learning task**

Participants completed a probabilistic reward-learning task with 60 rounds. In each round, participants chose between two stimuli with different probabilities of winning or losing one point. Choosing the
optimal stimulus resulted in a win with 85% probability and a loss with 15% probability, whereas choosing the other stimulus was associated with the reverse probabilities. The contingencies changed every 12th trial. Stimuli were two triangles with different orientations presented with a uniform gray background (see Figure 2). The orientation of the triangles was counterbalanced between participants.

**Figure 2. Overview of the reward-learning task.** Each trial started with a fixation cross presented for 1 second (A), followed by presentation of the stimuli (B). Stimuli remained on screen until participants made a choice (C). Feedback was presented one second after the key press (D), and consisted of written feedback, a stylized hand with a thumb pointing upwards or downwards, and a face with a neutral expression. For each participant, the same face was always associated with wins, and the other with losses.

**Computational modeling of behavioral data**

Participants’ choices were modelled with the Q-learning model:

\[ Q_{t+1}^k = Q_t^k + \alpha(r_t - Q_t^k) \]  \hspace{1cm} (1)

Where \( Q_t^k \) is the expected value of choosing a stimulus \( k \) on trial \( t \) after observing the outcome \( r_t \). The learning rate \( \alpha \) ranging between 0 (no updating) and 1 (complete updating) determines how much the
prediction error \( (r_t - Q^k_t) \) is updated on each trial. The probability of each choice was calculated with the Softmax rule:

\[
p^k_t = \frac{\exp(\beta q^k_t)}{\sum_{k=1}^{K} \exp(\beta q^k_t)}
\]

(2)

Where \( \beta \) is the inverse temperature, controlling the degree of exploration, ranging from 0 (random choice) to \( \infty \) (deterministic choice). Using the computational modelling parameters, \( \alpha \) and \( \beta \) estimated at the group level due to the small sample size, unique estimates of expected values (EV) prior to choice and PE after choice were calculated and used in subsequent analyses of pupil dilation data. See Appendix A for validation analyses and details of the computational model.

**Pupil dilation**

Pupil size was recorded with a Tobii Spectrum Pro corneal-reflection eye tracker at a sample rate of 1200 HZ. Gaps in the data shorter than 150 ms were replaced through linear interpolation, and the pupil signal was subsequently smoothed with a moving average filter with a window size of 150 ms. Pupil dilation responses (PDRs) were extracted during two time periods: the anticipation phase (from stimulus presentation to key press) and the feedback phase (from presentation of feedback until the end of the trial, see Figure 3). The PDR was defined as the mean percentual change in pupil size relative to a 500 ms baseline interval directly preceding the trial. The first 200 ms of each period were excluded, since task-evoked pupil dilation typically takes longer to evolve [12].

**Statistical analysis**

Visualizations of the behavioral data showed that EVs followed a bimodal distribution, and trials were therefore coded as explorative (negative EVs based on the learning history) or exploitative (positive EVs). Reward prediction errors followed a trimodal distribution, and were therefore coded as negative (PE < -0.5), medium (PE > -0.5 – 0.5), or positive (PE > 0.5).

Statistical analyses were conducted in MATLAB, JASP and R. Generalized linear models (GLMM) were used to analyze relations between PEs, expected outcomes and pupil dilation. These
analyses were conducted at the level of individual trials, including random intercepts for each individual to account for the fact that these were repeated measures. Statistical significance of the GLMMs was tested using analyses of variance (ANOVAs) with Satterthwaite approximated degrees of freedom. Relations between pupil dilation, neurometabolite concentrations and depressive symptoms were analyzed using Kendall’s tau (τb) due to the small sample size and non-normal data distribution.

Results

No relation was found between anticipatory PDRs and expected values (b = 0.12, F = 0.07, p = .782). Negative PEs elicited larger PDRs in the feedback phase than positive PEs (b = 1.45, F = 3.88, p = .049) and marginally larger PDRs than medium PEs (b = 1.10, F = 3.54, p = .061), whereas no difference was found between medium and positive PEs (b = 0.32, F = 0.25, p = .617) (see Figure 3).

To examine how individual differences in the strength of this effect related to neurometabolite concentrations and depressive symptoms, we calculated a difference score for each individual. First, the PEs were divided into the 25% most negative and the 25% most positive. Next, the mean value of the PDRs for the most positive PEs was subtracted from the mean value of the PDRs for the most negative PEs, henceforth, prediction-error-weighted pupil dilation response (PE-PDR). Higher PE-PDRs indicate relatively larger PDRs to negative PEs.

Figure 3. Pupil dilation as a function of time for trials with negative, medium, and positive prediction errors as a function of time. Following a one second presentation of stimuli (0-1 seconds), participants made a choice and
awaited the outcome (within 1-3.25 seconds in 90% of trials, see dotted line), and then viewed the outcome for 2 seconds. Pupil dilation is calculated using a 500 ms baseline interval directly preceding each trial. Shaded areas cover means ±1 SEM.

Depressive symptom scores were in the subclinical range (M±SD=4.5±2.5, range: 0-11).

Higher PE-PDRs were associated with lower levels of depressive symptoms ($\tau_b = -.44, p = .023, 95\% CI: -.10$ to $-.83$) (Figure 4A) and lower choline concentrations in the dACC, ($\tau_b = -.43, p = .017, 95\% CI: -.11$ to $-.75$) (Figure 4B). That is, more pronounced depressive symptoms and higher choline concentrations were related to smaller pupil dilations to negative vs. positive prediction errors. Higher levels of depressive symptoms were related to higher dACC choline concentrations ($\tau_b = .39, p = .026, 95\% CI: .02$ to $.76$) (Figure 4C). GABA and glutamate concentrations were not related to PE-PDR (GABA: $\tau_b = .10, p = .598, 95\% CI: -.31$ to $.51$; glutamate: $\tau_b = -.11, p = .536, 95\% CI: -.54$ to $.31$) or to depressive symptoms (GABA: $\tau_b = -.02, p = .915, 95\% CI: -.39$ to $.35$); glutamate, $\tau_b = .14, p = .431, 95\% CI: -.19$ to $.46$).

**Figure 4.** Scatterplots showing relations between prediction error-weighted pupil dilation response (PE-PDR), depressive symptoms measured using the Revised Child Anxiety and Depression Scale (RCADS), and choline ($t\text{Cho}$) to creatine ($t\text{Cr}$) concentration in the dorsal anterior cingulate cortex (dACC). Linear fit lines with shaded areas covering means +/- 1 SEM.

To disentangle the contribution of negative and positive prediction errors to these associations, we performed separate analyses for pupil dilations to negative and positive prediction errors. These post-hoc analyses could not detect any correlation between depressive symptoms or choline concentration and pupillary responses to negative or positive prediction errors ($\tau_b < -.26, ps > .17$).
Discussion

In this exploratory study, we combined self-rated depressive symptoms and magnetic resonance spectroscopy measures of brain metabolites with assessment of pupil dilation during behavioral tests of reward learning in a sample of adolescents. Smaller pupillary responses to negative vs positive prediction errors were related to higher depressive symptom scores as well as higher brain choline concentrations, with higher choline concentrations also predicting more pronounced depressive symptoms.

Previous studies have reported that more pronounced depressive symptoms are related to smaller pupillary response during reward anticipation [17]. Taking the anhedonic features of depression into account, it could be interpreted as if upcoming rewards are less arousing to individuals with more pronounced depressive symptoms. We only found changes during the feedback phase and specifically that depressive symptoms were related to smaller pupillary responses to negative prediction errors. That is, when the outcome is worse than expected, as compared to better than expected, individuals with more depressive symptoms are not as reactive in their autonomous response, possibly indicating less surprise consistent with mood-congruent cognitions and expectations. This points to a reduced engagement of locus coeruleus (LC) activity (i.e., norepinephrine signaling) modulating pupil size in response to negative prediction errors and stands in contrast to previous findings that depressive symptoms are related to increased LC activity during negative prediction errors [16]. One reason for the contradictory results is tied to the population under study. Previous work has targeted adults whereas our study only included 13-year-old adolescents. LC activity and modulations of pupil size during prediction errors may be driven by connections to the LC from the ACC. Indeed, the dACC plays an important role in prediction error signaling [2,8], which is in line with our findings of a negative correlation between dACC choline concentration and negative prediction error-related pupil dilation and increased dACC choline concentrations in individuals with higher depression scores.

Dorsal ACC choline concentration was related to depressive symptoms, which is in line with reports of higher frontal choline concentrations in patients with depression [15]. Here, we extend
previous findings of cholinergic alterations in patients by also showing a relation in adolescents with relatively low levels of depressive symptom scores. In contrast to previous findings of glutamatergic and GABA-ergic involvement in depressive states [11], we could not detect a relation between depressive symptoms and dACC glutamate and GABA concentrations. It is possible that this is related to the adolescent sample or that we did not contrast clinically depressed individuals with healthy comparison subjects. The findings may also suggest that changes in choline levels are associated with subclinical depressive symptoms or occur earlier in the development of depressive episodes than changes in GABA and glutamate.

Limitations of the study include the small sample size, the relatively low levels of depressive symptoms, and the restricted age range of study participants. Hence, the findings should be regarded as preliminary. Further, the MEGA-PRESS MRS sequence which was used cannot disentangle GABA from macromolecules. Additionally, modeling $\alpha$ and $\beta$ at a group level precluded analyses of relations between learning and depressive symptoms and neurometabolite concentrations.

In conclusion, these preliminary findings lend further support to notions of cholinergic involvement in depressive symptoms and in the difference between pupillary response to negative and positive prediction errors. Findings also suggest that the pupillary difference score may hold promise as a biomarker of depressive symptoms.

**Acknowledgements**

The study was funded by a grant from Riksbankens Jubileumsfond (The Swedish Foundation for Humanities and Social Sciences) to Dr. Frick and Dr. Gingnell (P17-0256:1). Dr. Frick was supported by Kjell och Märta Beijers stiftelse. Funding sources had no role in study design, in the collection, analysis and interpretation of the data, in writing the report, or in the decision to submit the article for publication. Dr. Pine was supported by NIMH Intramural Research Grant MH002781. Dr. Kleberg was supported by Åke Wiberg Foundation (M20-0152), Magnus Bergwall Foundation, Fonden för psykisk hälsa (The Foundation for Mental Health), The Centre for Psychiatry Research and The Strategic Research Area Neuroscience (StratNEURO).

**Declarations of interest: none.**
References

Appendix A: Computational modeling of behavioral data

Modelling details

Due to the small sample size, reinforcement learning parameters were modelled at the group level. We modelled behavioural responses with the Q-learning model [22], a variant of Rescorla-Wagner model that take each history of previous outcomes into account, using these to make a decision about the next trial. In the Q-learning model the expected value of choosing a stimulus $k$, $Q_k^t$, is updated at each trial $t$ after observing the outcome of that choice, $r_t$, according to:

$$Q_{t+1}^k = Q_t^k + \alpha (r_t - Q_t^k) \quad (Eq.1)$$

where $(r_t - Q_t^k)$ is the prediction error (PE) or mismatch between expected and received outcomes for the chosen stimulus. The PE is weighed according to the parameter $\alpha$, which is a learning rate that can vary between 0 and 1. An $\alpha$ of 1 means that the expected value of choosing stimulus $k$ is fully updated at each trial (i.e., that only the most recent outcome is considered). In contrast, an $\alpha$ of zero means that no learning takes place. Following previous studies, we used the Softmax choice rule (Equation 2) to compute the probabilities for each choice. The Softmax choice rule assumes that the participant chooses the alternative with the highest probability for reward (exploiting), but occasionally selects the hitherto less rewarding stimulus (exploring).

$$p_t^k = \frac{\exp(\beta q_t^k)}{\sum_{i=1}^{k} \exp(\beta q_t^i)} \quad (Eq.2)$$

$\beta$ is the inverse temperature, controlling the degree of exploration, ranging from $\beta = 0$ for complete random choices and $\beta = \infty$ for completely deterministic choices. Using the computational modeling parameters, $\alpha$ and $\beta$ estimated at the group level, unique estimates of expected values (EV) prior to choice and PE after choice were calculated and used in subsequent analyses of pupil dilation data.
Raw data were processed using custom scripts written in Matlab (Mathworks, Inc.). Maximum-likelihood estimation using the Matlab function `fminsearchbound` was used to optimise the free parameters (see below) in the model.

**Model validation in the observed data**

Following previous studies, we compared the final model to two alternative models. *Alternative model 1* was an adaptation of the final Q-learning model, but with separate learning rates $\alpha$ for positive and negative outcomes. *Alternative model 2* assumes that participants are not engaged in the task, and that their choices are therefore not influenced by their learning history. Instead, the probability of choosing stimulus $k_1$ would be $b$ and the probability of choosing $k_2$ would be $1 - b$, where $b$ is a parameter capturing a potential bias for either stimulus ranging between 0 (never choosing $k_1$) to 1 (always choosing $k_1$). In other words, alternative model 2 assumes random choices except for a potential bias for either stimulus that is not affected by learning history. We compared the fit of the final model and the two alternative models using the Bayesian information criterion (BIC). These comparisons favoured the final model (BIC= 1242), over alternative model 1 (BIC= 1266) and alternative model 2 (BIC= 1579). At the individual level, the final model produced the lowest BIC values for 15/19 individuals, and for four individuals, the lowest BIC values were produced by alternative model 2.

**Parameter recovery from simulated data**

In a second step, we tested the ability of the final model to recover parameters of simulated data sets generated from the model. We simulated 100 data sets with the same number of participants as the actual experiment, and $\alpha$ and $\beta$ parameters drawn from Gaussian distributions with means and standard deviations in the approximate range found in the observed data.

The final model and alternative models were then fit to the data and we extracted the proportion of simulated data sets in which the final model produced better fit (lower BIC values) than the alternative models. This analysis showed that the final model had better fit than alternative model 1 in 87% of the simulated data sets and alternative model 2 in 98% of cases. Strong correlations were
observed between simulated and recovered $\alpha (r = 0.93)$ and $\beta (r = 1)$, demonstrating successful parameter recovery.

**Credit author statement**

Mona Guath: Software, formal analysis, writing - original draft., writing - review & editing.
Johan Lundin Kleberg: Conceptualization, methodology, software, formal analysis, data curation, visualization, writing - original draft, writing - review & editing.
Jan Weis: Software, methodology, formal analysis, resources, visualization, writing - review & editing.
Ebba Widegren: Investigation, writing - review & editing.
Matilda Frick: Investigation, writing - review & editing.
Stefan Möller: Investigation, writing - review & editing.
Lisa Klevebrant: Investigation, writing - review & editing.
Barry Karlsson: Investigation, writing - review & editing.
David Fällmar: Resources, writing - review & editing.
Johanna Mårtensson: Methodology, resources, writing - review & editing.
Daniel S Pine: Supervision, writing - review & editing.
Karin Brocki: Resources, writing - review & editing.
Malin Gignell: Conceptualization, methodology, investigation, resources, writing - review and editing, supervision, project administration, funding acquisition.
Andreas Frick: Conceptualization, methodology, formal analysis, investigation, resources, data curation, writing - review & editing, supervision, project administration, funding acquisition.