Investigation of reward learning and feedback sensitivity in non-clinical participants with a history of early life stress.

Matthew P Wilkinson[1], Jack R Mellor[1], Emma S J Robinson[1*]

1. University of Bristol, School of Physiology, Pharmacology and Neuroscience, Biomedical Sciences Building, University Walk, Bristol, BS8 1TD

*Author for correspondence:

Email: Emma.S.J.Robinson@bristol.ac.uk

Tel: +44 (0)117 331 1449, Fax: +44 (0)117 3312288

Abstract

Background: Early life stress (ELS) is an important risk factor for the development of depression. Impairments in reward learning and feedback sensitivity have been suggested. to be an intermediate phenotype in depression aetiology. We therefore hypothesised that healthy adults with a history of ELS would have impairments in reward learning and feedback sensitivity.

Methods: We recruited 64 adult participants with high levels of ELS and no diagnosis of a current mental health disorder in addition to 65 controls. Participants completed two online reward learning tasks: the probabilistic reversal learning task (PRLT) and probabilistic reward task (PRT). Participants also completed depression, anhedonia, social status and stress scales with PRLT data being additionally analysed utilising a reinforcement learning model.

Results: Participants with high levels of ELS showed decreased positive feedback sensitivity (PFS) in the PRLT compared to controls. High ELS participants also tended towards possessing a decreased model-free learning rate which strengthened in subsequent analysis. This was coupled with a decreased learning ability in the acquisition phase of block 1 following the practice session. Neither groups of participants showed a reward induced response bias in the PLT however high ELS participants exhibited decreased discrimination ability between stimuli; this was however accounted for by depression symptomology in further analysis.

Conclusions: These data suggest that healthy participants without a mental health diagnosis and high levels of ELS show deficits in PFS and reward learning in the PRLT that are distinct from depressed patients. These deficits may be relevant to an increased vulnerability to depression.

1. Introduction

Early life stress (ELS) is a major known risk factor for the development of depression (Agid et al., 1999; Green et al., 2010; Lemoult et al., 2019; McCauley et al., 1997). Elevated levels of childhood stress lead to widespread functional and morphological alterations in the adult brain with the hippocampus, amygdala and prefrontal cortex being most impacted (Cohodes, Kitt, Baskin-Sommers, & Gee, 2020; Tottenham, 2009). This not only renders those with a history of ELS vulnerable to depression but may also lower the threshold of stress required to precipitate depression (Hammen, Henry, & Daley, 2000). However, how ELS influences the developing brain to predispose individuals to psychiatric illness is not yet understood.

Reward learning deficits have been proposed to be an intermediate phenotype in the aetiology and maintenance of depression (Halahakoon et al., 2020; Pizzagalli, Losifescu, Hallet, Ratner, & Fava, 2008; Vrieze et al., 2013; Whitton, Treadway, & Pizzagalli, 2015). Depressed patients show decreased reward sensitivity in the probabilistic reward task (PRT, Pizzagalli et al., 2008), a test of reward learning. These deficits have been observed to both predict the risk of disease development (Bress, Foti, Kotov, Klein, & Hajcak, 2013) and persistence (Pechtel, Dutra, Goetz, & Pizzagalli, 2013; Vrieze et al., 2013). Utilising a different reward learning assay, the probabilistic reversal learning task (PRLT), depressed patients show impaired accuracy following probabilistic rule reversal and increased sensitivity to probabilistic negative feedback (Murphy, Michael, Robbins, & Sahakian, 2003; Taylor Tavares et al., 2008). Acute stress has also been observed to impair reward learning (Berghorst,

Bogdan, Frank, & Pizzagalli, 2013; Bogdan & Pizzagalli, 2006) suggesting a potential link between stress, reward processing deficits and depression aetiology.

Previous studies have investigated reward processing deficits in people who have experienced ELS. Hanson et al., 2017 recruited adolescents with a history of physical abuse who then completed a probabilistic learning task where they showed lower associative learning compared to controls. Changes in reward learning have also been reported within another probabilistic reward task, the probabilistic stimulus selection task (PSST), by Pechtel and Pizzagalli, 2013. Women with a history of childhood sexual abuse and a diagnosis of MDD showed decreased performance on trials requiring learning of previously rewarded information compared to MDD only and control groups. Although these studies provide valuable insights, they use different tasks to those previously used to study depressed populations making direct comparisons difficult. Additionally, studies are needed in adults without a current mental health diagnosis to understand if any reward processing changes are present prior to the development of mental health disorders.

In this study it was hypothesised that ELS leads to alterations in reward processing and feedback sensitivity in an otherwise healthy adult population. Two groups of adult participants that reported no diagnosis of a mental health condition or Parkinson's disease were recruited online and completed a survey of adverse childhood experiences (Cohen et al., 2006) before being split into high and no ELS groups. Participants completed the PRT and PRLT with PRLT data additionally being analysed using a Q-learning model to probe reward learning parameter changes. Participants were asked about stress exposure to enable

exploratory analysis investigating if life stress interacts with ELS to cause reward processing deficits. By understanding the links between ELS and reward processing deficits as a hypothesised intermediate phenotype in depression this aims to provide insights into how a person with a history of ELS is rendered at higher risk for depression.

2. Methods

All procedures were approved by the Faculty of Life Sciences and Faculty of Science Research Ethics Committee at the University of Bristol and the study protocol was pre-registered (www.osf.io/538yk). All participants provided full written consent for the collection, analysis and publication of their data which is available open access and were reimbursed at a rate of £6.00 per hour.

2.1 Participants

586 participants were recruited using the Prolific (www.prolific.co) online platform to complete an online screening questionnaire (see supplementary figure 1 for study overview). These participants were 25 - 65 years of age, fluent in English, resident in the UK and had no mild cognitive impairments or dementia. Participants completed the early life stress questionnaire (ELSQ, Cohen et al., 2006) while also being asked to self-report if they had a diagnosis of a mental health condition or Parkinson's disease.

Participants who met the inclusion criteria for high ELS or no ELS and did not report a diagnosis of a mental health disorder or Parkinson's were then invited to take part in a second phase of the experiment online within a week of screening and were allocated into two groups. A no ELS group (n = 65) contained people scoring 0 on the ELSQ while a high ELS group (n = 64) consisted of those who scored ≥3 (estimated to be the top tercile of the population from Cohen et al., 2006). In this second phase of the experiment participants entered demographic information before completing the MacArthur Scale of Subjective Social Status (Adler, Epel, Castellazzo, & Ickovics, 2000), Beck's depression inventory II (BDI-II, Beck et al., 1996), the Snaith Hamilton pleasure scale (SHAPS, Snaith et al., 1995) and the Holmes and Rahe stress scale (Holmes & Rahe, 1967). The SHAPS was additionally scored using the SHAPS-C criteria (Ameli et al., 2014) while for the stress scale participants were asked if each event occurred in either their adult life or the last year. For all stages of the experiment participants were instructed to use a desktop or laptop only and that they should be in a quiet place with minimal distractions. Sample size was estimated for a medium effect size (Cohen's d = 0.5) and 80% power for a t-test at 64 participants per group.

2.2 Behavioural testing

Following completion of self-report measures, participants completed the Probabilistic reversal learning task (Cools, Clark, Owen, & Robbins, 2002; Waegeman, Declerck, Boone, Seurinck, & Parizel, 2014) followed by the Probabilistic reward task (Pizzagalli, Jahn, & O'Shea,

2005). To complete the tasks participants were required to download and install the Millisecond Inquisit web player (Millisecond, US) which ran both tasks using Millisecond Inquisit v6.2.1. Participants were instructed they were able to earn an additional £2.00 for high performance on the behavioural tasks.

2.2.1 Probabilistic Reversal Learning task

The PRLT was conducted as previously described (Cools et al., 2002; Waegeman et al., 2014) using the task from the Millisecond test library (Millisecond, 2020a). Participants were instructed to choose between a "lucky" (rich) and "unlucky" (lean) pattern to maximise points. Selection of the rich stimulus enabled participants to gain a point 80% of the time and lose a point 20% of the time with the lean stimulus having opposite contingencies. If no stimulus was chosen within 2s then this was classed as incorrect and participants lost a point. After meeting the reversal criterion, the contingencies reverse such that the rich stimuli becomes lean and vice versa. This criterion was set randomly between 10 to 15 consecutive correct rich choices to stop participants counting to the criterion. Participants first completed a practise phase where they had to achieve the criterion for a single reversal before proceeding to the main task which was completed in three blocks each limited to 9 minutes. Participants who did not pass the practice phase were excluded from analysis. Data was analysed as previously described (Wilkinson, Grogan, Mellor, & Robinson, 2020) with win-stay and lose-shift probabilities being calculated as measures of positive and negative feedback sensitivity respectively. These were subdivided into either true, feedback that matches with the underlying task rules, or misleading feedback, that which is opposite to the underlying

task rule. The number of rule changes, accuracy and response latency per block were additionally analysed. A Qlearn reinforcement learning model was applied to data as previously described (Grogan et al., 2017; Wilkinson et al., 2020) to give estimates of learning rate, accuracy compared to a model predicted perfect strategy (subjective accuracy) and beta, a measure of choice variability. Additionally, data per phase (practice, acquisition of the first rule in block 1 and the following two reversals) was analysed consisting of participant accuracy, errors to criterion and win-stay / lose-shift probability.

2.2.1 Probabilistic Reward Task

The PRT was conducted as previously described (Pizzagalli et al., 2005) using the task from the Millisecond test library (Millisecond, 2020b). Participants were instructed to identify whether the mouth of a presented cartoon face was long or short to win points over 3 blocks of 100 trials. Participants were shown a face before a mouth was rapidly presented for 100ms with participants given up to 1750ms to respond. Feedback was not provided on every trial but unknown to participants one mouth was rewarded with points three times more often than the other (rich = 60%, lean = 20%). Response key and rich/lean stimuli assignments were counterbalanced across participants and responses that were quicker than 150ms or slower than 1750ms were excluded from analysis. Additional responses that differed by more than 3 standard deviations from the mean following natural log transformation of latencies for each participant were excluded from analysis. Response bias (logB), a measure of reward learning, and discriminability (logD), a measure of task difficulty, were calculated as described previously (Pizzagalli et al., 2005).

2.3 Data Analysis

Demographic and self-report measures were compared between groups using either X^2, t-tests or Mann-Whitney U tests where appropriate. The primary analysis for each measure was a direct comparison between no ELS and high ELS groups. Where data was not normally distributed then efforts were first made to transform data to normality and where this was not possible Mann-Whitney U tests were completed. Win-stay by block data was transformed using the bestNormalize package in R (Peterson & Cavanaugh, 2019). Where measures were split by a within subject factor such as block or feedback type these were analysed with repeated measures ANOVAs. Where Mauchly's test identified a violation of the Sphericity assumption then this was corrected using the Huynh-Feldt correction. T-tests were used for direct group comparisons.

Due to differences in social status, BDI-II score and SHAPS score between the no ELS and high ELS groups, principle component analysis (PCA) was conducted to reduce the dimensionality of these variables to account for depression symptomology as an analysis stage (see supplementary tables 1 and 2). Because only principle component 1 (PC1) differed between groups and explained 94.6% of variance this was used in ANCOVAs (analysis of covariance) to analyse whether parameter changes were due to ELS or due to changes in depression symptomology accounted for by the PC1 component. To understand if stress and gender interacted with ELS to modify reward learning, exploratory analysis was also undertaken using generalised linear mixed models (GLMMs) containing the factors: gender, ELS, lifetime stress, last year stress and age. GLMMs were fit using the glmmTMB package in R 4.0 (Brooks et al.,

2017; R Core Team, 2020) with model refinement conducted utilising stepwise deletion based upon Akaike information criterion before being compared with a null model to protect against overfitting. PC1 was also added to each model following final model selection to assess the effects of depression symptomology.

Statistical analysis was conducted in SPSS v26 (IBM, US), MATLAB 2018a (Mathworks, USA) and R 4.0 (R Core Team, 2020) with output graphics constructed in GraphPad Prism 8 (GraphPad, US). All data is shown as mean ± SE with a bar and stars showing a main effect of ELS in the primary analysis. * ≤ 0.05, ** < 0.01, *** < 0.001, **** < 0.0001.

3. Results

Early life stress was highly prevalent in the study population with only 21.0% of participants having no adverse childhood experiences (ACEs) and 44.4% of the population suffering three or more ACEs in their childhood (see supplementary figure 2). 16.0% of respondents self-reported a diagnosis of a mental health disorder or Parkinson's with this being associated with a higher ELSQ score (Mann-Whitney, $U = 15725$, $p < 0.0001$).

The two study groups were well matched with respect to gender, age, education, ethnicity, relationship status, employment status and the presence of monetary worries (see table 1). However, high ELS participants had a self-reported lower social status coupled with higher

depression scores in the BDI-II and elevated anhedonia scores in the SHAPS questionnaires. There was no difference between groups when participants were asked about stress they encountered in both the last year and their adult lives. When the BDI-II scores were classified into either minimal, mild, moderate or severe depression (see supplementary figure 3, Beck et al., 1996) participants from the high ELS group were more likely to be in greater severity depression groupings (chi^2, $X^2(3) = 12.9$, $p = 0.005$). Similarly when SHAPS scores were classified into either normal (≤2) or abnormal (≥3) hedonic responses (Snaith et al., 1995) members of the high ELS group were more likely to have abnormal scores (see supplementary figure 3, chi^2, $X^2(1) = 6.3$, $p = 0.012$).

Measure	No ELS (n = 65)	High ELS (n = 64)	Test statistic	p
Gender (% male)	44.6	37.5	$x^2(2) = 2.5$	0.28
Age (years)	37.3 ± 1.30	38.0 ± 1.24	U = 1936.0	0.50
Education (% graduates)	64.6	65.6	$x^2(5) = 4.9$	0.43
Ethnicity (% white)	95.4	82.8	$x^2(4) = 8.7$	0.070
Relationship status (% single)	18.5	28.1	$x^2(3) = 1.9$	0.60
Employment status (% full time)	64.6	60.9	$x^2(5) = 3.5$	0.61
Monetary concerns (% agree / strongly agree)	36.9	56.3	$x^2(3) = 4.4$	0.22
ELSQ	0 ± 0	4.36 ± 0.17	-	-
Social status	6.2 ± 0.17	5.2 ± 0.21	U = 1397.5	**0.001**
BDI-II	9.4 ± 1.0	15.2 ± 1.22	U = 1315.5	**0.0003**
SHAPS	1.4 ± 0.25	2.56 ± 0.32	U = 1496.5	**0.004**
SHAPS-C	24.3 ± 0.67	26.4 ± 0.86	$t_{119.4} = -1.92$	0.057
Lifetime stress	472.8 ± 22.4	529.2 ± 23.9	$t_{127} = -1.72$	0.088
Last year stress	111.4 ± 12.3	139.8 ± 17.0	U = 1939.5	0.51

Table 1. Demographic and self-report measures in the study population. Values are shown for each group as mean ± standard error with significant p values indicated in bold.

3.1 Probabilistic reversal learning task

There was no difference between groups in either the number of rule changes participants were able to complete (Figure 1A) or accuracy (Figure 1B). However participants with a history of high ELS did have a slower average response latency (Figure 1C, RM-ANOVA, $F_{1,126} = 5.03$, $p = 0.027$) with both groups getting equally faster over the course of the three blocks (RM-ANOVA, $F_{1.88,236.7} = 16.1$, $p < 0.0001$). Secondary analysis revealed little effect of depression symptomology (RM-ANCOVA, PCA1: $p > 0.05$) with the main effect of ELS persisting (RM-ANCOVA, ELS: $F_{1,125} = 4.9$, $p = 0.028$). Exploratory analysis on overall reaction times did not replicate a main effect of group but did observe older participants having slower reaction times (GLMM, $Z = 2.8$, $p = 0.005$). This analysis also indicated a trend towards an interaction between group and lifetime stress (GLMM, $Z = 1.55$, $p = 0.065$) but further investigation did not reveal an effect of lifetime stress in either group.

When data was analysed using the Q-learning reinforcement learning model a trend emerged towards high ELS participants having a lower learning rate compared to the no ELS study population (Figure 1D, t-test, $t_{127} = 1.78$, $p = 0.077$). Secondary analysis revealed no effect of PCA component 1 upon learning rate but abolished any effect of ELS. In exploratory analysis a main effect of ELS was observed (GLMM, $Z = 2.1$, $p = 0.037$) with the addition of PC1 impairing model fit ($\Delta AIC = 1.69$, $X^2(1) = 0.31$, $p = 0.57$). Additionally, a relationship between stress in the last year and learning rate was observed whereby increased stress in the last year decreased learning rate (GLMM, $Z = -2.3$, $p = 0.024$). There was no difference in choice

variability (Figure 1E) or accuracy compared to a model predicted perfect strategy (Figure 1F) between groups.

Participants with a history of high ELS exhibited reduced positive feedback sensitivity (PFS, Figure 2A, RM-ANOVA, $F_{1,122}$ = 10.4, p = 0.002) which persisted once depression symptomology was accounted for using PCA component 1 (RM-ANOVA, $F_{1,121}$ = 6.6, p = 0.01). Exploratory analysis revealed an interaction between ELS and both lifetime stress (GLMM, Z = -2.15, p = 0.031) and last year stress (GLMM, Z = -1.99, p = 0.047). Further investigation revealed effects of both stress types upon PFS in the low ELS group only (GLMM, lifetime stress: Z = -2.35, p = 0.019, last year stress: Z = -2.2, p = 0.026) whereby higher lifetime stress led to greater PFS but higher stress in the last year was associated with decreased PFS. However it should be noted that although all suggested terms were removed from the model the overall model was a poorer fit than the null when measured by AIC (Δ AIC = 7.3, $X^2(13)$ = 18.7, p = 0.13).

The effect of ELS upon PFS was consistent across feedback that matched (true feedback) or clashed (misleading feedback) with the underlying task rules (Figure 2B, Mann-Whitney U, true: U = 1443, p = 0.03; misleading: U = 1337, p = 0.005). This effect appeared to be constrained to PFS with no corresponding changes in lose-shift probability between no ELS and high ELS groups (Figures 2C and D).

When initial learning in the PRLT task was assessed it was apparent that although ELS and control participants performed similarly during the practice phase there was a learning deficit during acquisition of the first reversal criterion in block 1 as evidenced by increased errors to criterion (Figure 3A, Mann-Whitney U, $U = 1580$, $p = 0.045$) and decreased accuracy (Figure 3B, Mann-Whitney U, $U = 1584$, $p = 0.036$). Both groups of participants however performed equally well at achieving criterion for a second and third reversal. Unlike the overall measures there was no difference in win-stay probability between groups (Figure 3C), however there was a trend for high ELS participants to show increased negative feedback sensitivity (NFS) in the practice phase (Figure 3D, Mann-Whitney U, $U = 1532$, $p = 0.052$).

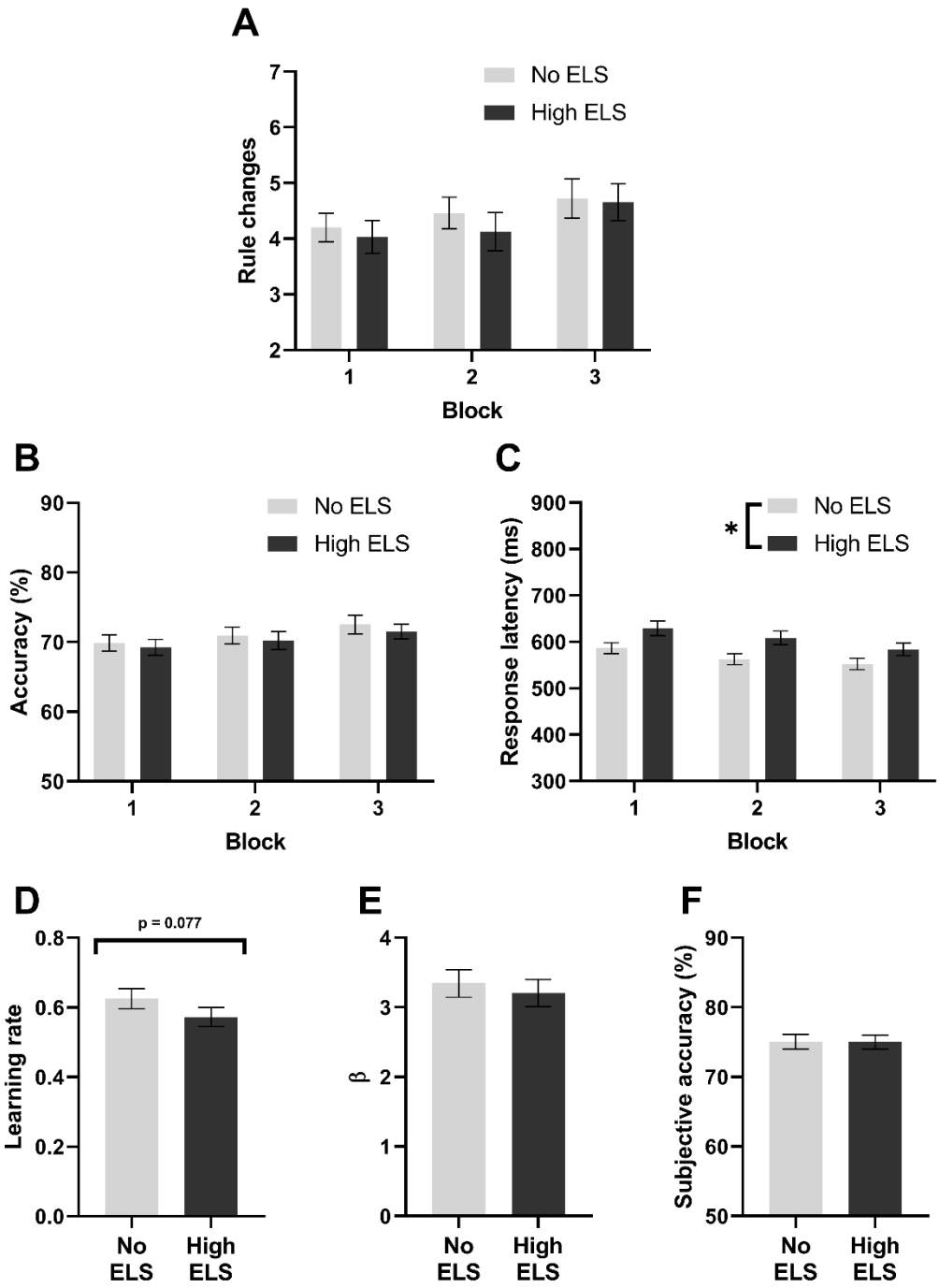

Figure 1. Overall reward learning and reinforcement learning in the PRLT. (A) Rule changes within each block, **(B)** accuracy by block and **(C)** average response latency per block. From the Q-learn reinforcement learning model: **(D)** learning rate, **(E)** β, the inverse of the softmax temperature and a measure of choice variability and **(F)** subjective accuracy, participant accuracy compared to a model predicted perfect strategy.

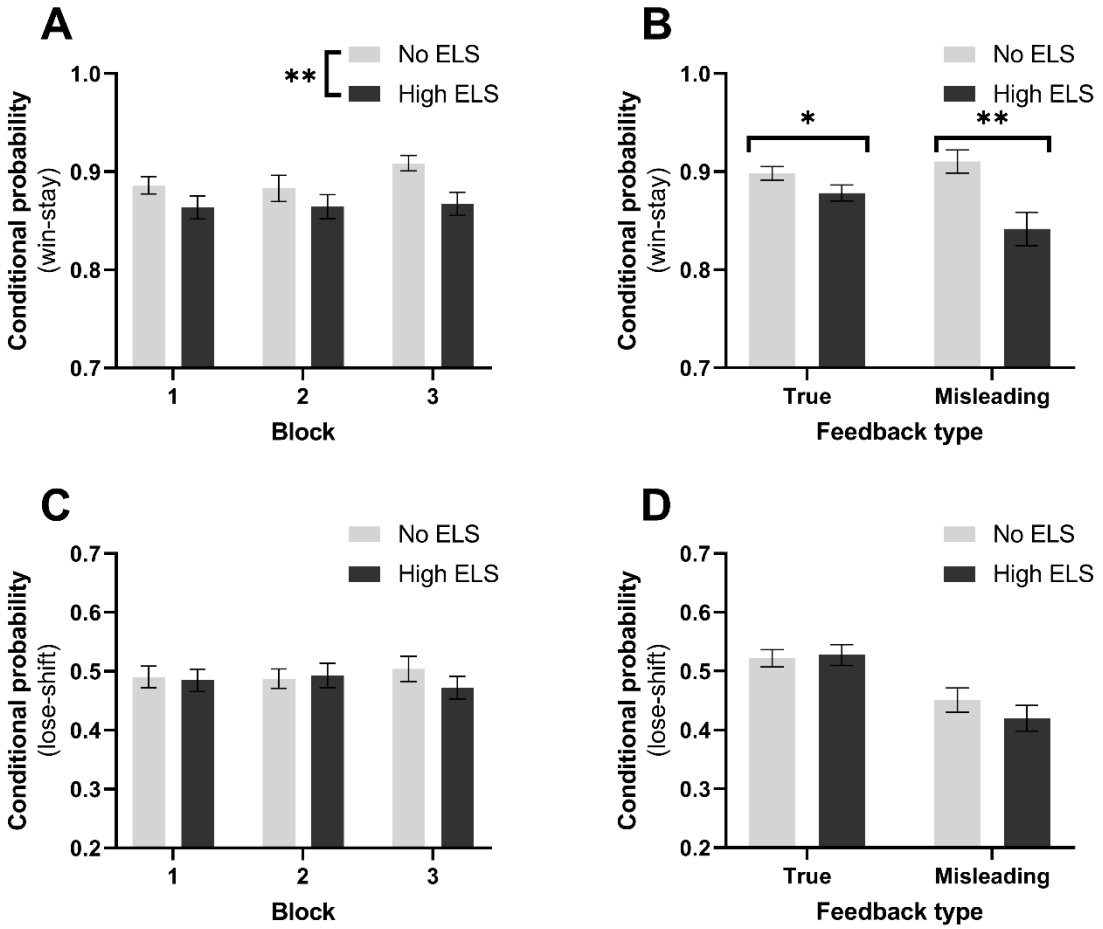

Figure 2. High ELS participants exhibited lower positive feedback sensitivity than those without a history of ELS. (A) Win-stay probability, **(B)** Lose-shift probability, **(C** and **D)** win-stay and lose-shift probability respectively subdivided into true and misleading feedback.

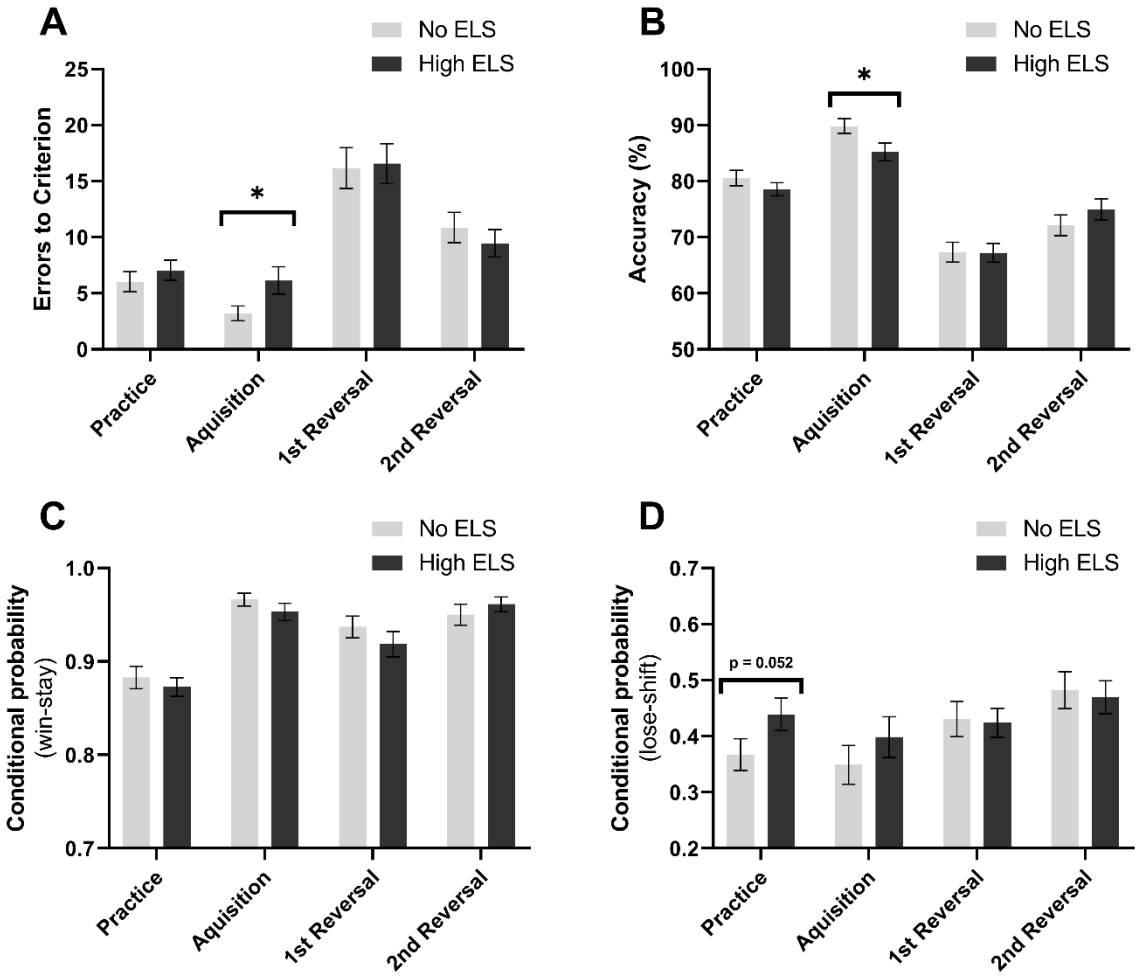

Figure 3. High ELS participants show impaired learning in the acquisition phase of block 1. **(A)** Errors made while reaching criterion for each phase, **(B)** accuracy within each phase, **(C** and **D)** win-stay and lose-shift probabilities for each phase of block 1 and practice respectively.

3.2 Probabilistic reward task

Neither group of participants developed a response bias towards the more highly rewarded stimulus in any block (Figure 4A) nor was there evidence for a response bias developing between blocks (Figure 4B). However, participants with a history of high ELS did show an impaired ability to discriminate between stimuli (Figure 4C, ANOVA, $F_{1,127} = 4.8$, $p = 0.030$). Secondary analysis revealed that this difference between groups appeared to be driven by differences in depression symptomology with the effect of ELS disappearing when PCA component 1 was included in the analysis (ANCOVA, PCA1: $F_{1,126} = 6.08$, $p = 0.015$; ELS: $F_{1,126} = 1.7$, $p = 0.19$). Exploratory analysis further revealed a main effect of lifetime stress with higher lifetime stress corresponding to increased discrimination ability (GLMM, $Z = 2.6$, $p = 0.007$). An effect of gender was also revealed (GLMM, $Z = 2.04$, $p = 0.04$) with males showing increased discrimination ability. Finally, there was no difference between groups in response latencies (Figure 4D).

Consistent with Pizzagalli et al., 2008 the probability of misclassifying a stimulus based upon the preceding trial outcome was also analysed (supplementary table 3). Participants with a history of high levels of ELS were more likely to misclassify rich stimuli if either the previous trial was a not rewarded rich trial or a lean not rewarded trial with these measures roughly corresponding with rich lose-shift and lean lose-stay probability in the PRLT respectively.

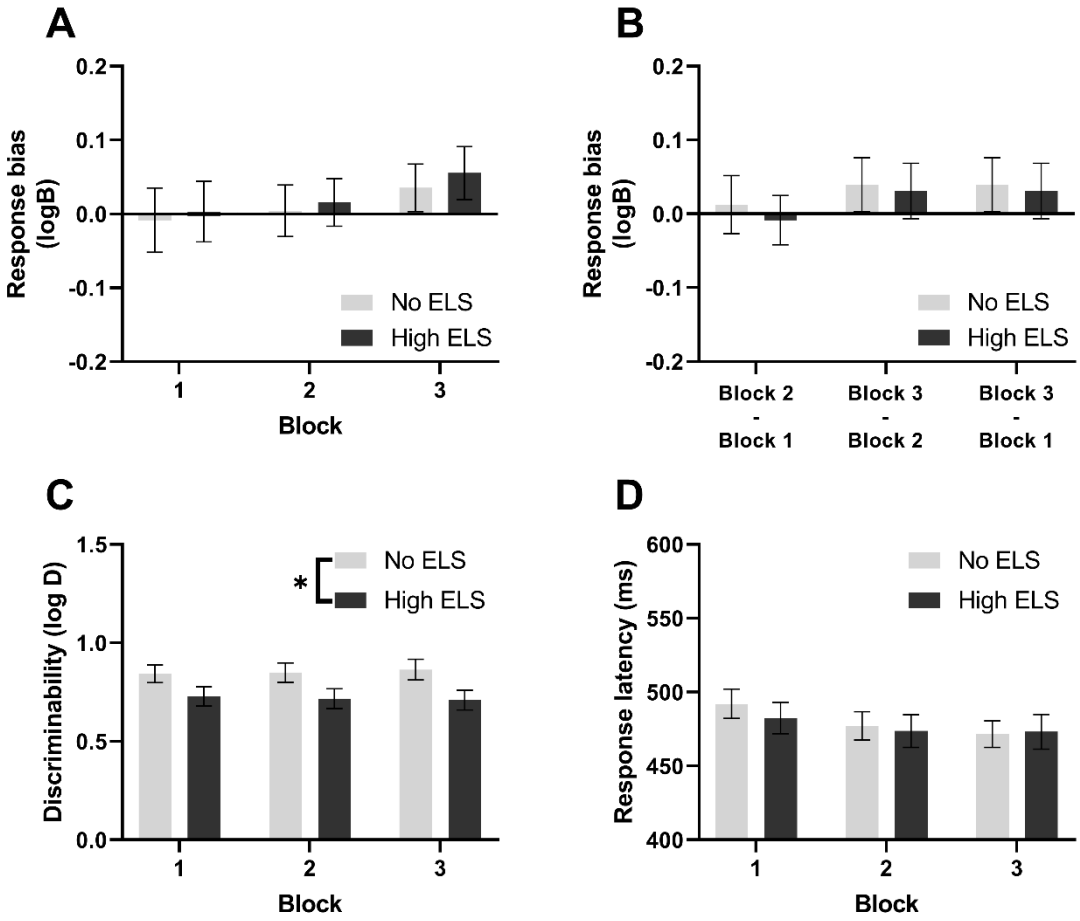

Figure 4 Participants with a history of ELS show decreased discriminability in the PRT. (A) Response bias to the more highly rewarded stimulus, **(B)** response bias development between blocks, **(C)** discriminability between long and short face lengths and **(D)** average response latency.

4. Discussion

This study was designed to investigate whether healthy adults with a history of ELS show alterations in reward processing and feedback sensitivity. Nearly 600 participants were screened; ELS was highly prevalent in the population with 79.0% of participants experiencing one or more ACE and 44.4% experiencing three or more.

Participants with a history of high ELS had higher self-report depression and anhedonia symptoms. Although participants stated they did not have a diagnosis of depression, 54.7% of high ELS and 26.2% of no ELS participants showed at least mild symptoms based upon the BDI-II questionnaire. BDI-II scores in no ELS (mean: 9.4 ± 1.0) and high ELS (mean: 15.2 ± 1.2) participants were higher than controls for similar studies (range 1.3-3.62, Pechtel & Pizzagalli, 2013; Pizzagalli, Bogdan, Ratner, & Jahn, 2007; Pizzagalli et al., 2008) but lower than depressed patients (mean: 32.1 ± 8.6, Pizzagalli et al., 2008) or participants described as having a high BDI (>16, Pizzagalli et al., 2007). These data consistent with a large societal burden of un-diagnosed depression (Lewis et al., 2019; Li et al., 2009; Lotfaliany et al., 2018). It should be noted that the present study was undertaken during the Covid-19 global pandemic with it being estimated that levels of depression had doubled during this period (Office for National Statistics, 2020). The high level of undiagnosed depression is a major limitation of this study as depression and anhedonia are well known to reduce reward learning in both the PRLT (Murphy et al., 2003) and PRT (Pizzagalli et al., 2005, 2008; Vrieze et al., 2013). It is also worth considering that 75% of adults with mental health conditions experience the onset of symptoms before aged 24 (Kessler et al., 2005). This means that the

study population, all 25 years of age or greater, is potentially biased towards those more protected from mental health disorders.

4.1 Probabilistic reversal learning task

In the PRLT participants with high ELS displayed decreased positive feedback sensitivity compared to controls as measured by win-stay probability. This finding was independent of depression symptomology and specific to PFS with no changes observed in lose-shift probability. Blunted striatal responses to reward in participants with a history of ELS have been previously reported (Hanson et al., 2016, 2015) which may underly the decreased PFS observed in the present study. Consistent with the present study, women with MDD and a history of childhood sexual abuse have also been found to have impaired performance in the PSST but only for trials requiring use of previously rewarded information and not those requiring use of previously punished information (Pechtel & Pizzagalli, 2013). Within the PRLT depressed patients have been observed to show increased sensitivity to misleading negative feedback (Murphy et al., 2003; Taylor Tavares et al., 2008). This was not observed in the high ELS cohort in the present study. In other tasks depressed patients have also been reported to show increased NFS alongside attenuated PFS (Elliott, Sahakian, Herrod, Robbins, & Paykel, 1997; Foti & Hajcak, 2009; Herzallah et al., 2013; Mueller, Pechtel, Cohen, Douglas, & Pizzagalli, 2015; Webb et al., 2017). These findings suggest that ELS influences feedback sensitivity in the PRLT differently to depression with ELS decreasing PFS but not effecting NFS while depression has an opposite effect.

The PRLT also allows for assessment of reinforcement learning through the analysis of rule changes and accuracy in addition to parameters calculated through use of the reinforcement learning model. In contrast with our hypothesis, ELS did not affect rule changes which is surprising considering evidence that both depression and ELS can impair cognitive flexibility (Murphy, Michael, & Sahakian, 2012; Zhou et al., 2020). Although rule changes was used as the main behavioural reward learning output, when data was analysed with the Q-learn model a trend towards decreased learning rate was observed in high ELS participants. This was became significant in exploratory analysis and decreased associative learning has been previously observed in juveniles previously exposed to physical abuse (Hanson et al., 2017). There is a lack of consistent evidence in depression studies as to whether model free learning rate differs between patients and controls (Chen, Takahashi, Nakagawa, Inoue, & Kusumi, 2015; Robinson & Chase, 2017). These findings warrant future investigation due to this study being only powered to detect group differences between two groups meaning that ANCOVA and exploratory analysis is likely to be underpowered.

A slower response latency was also observed in high ELS participants which was specific to the PRLT with no congruent changes seen in the PRT. This discrepancy may be related to differing cognitive demands with the PRLT potentially requiring greater working memory.

No directly comparable studies have been carried out in humans. However, maternally separated marmosets, an animal model of ELS, showed no change in simple visual discrimination compared to controls but showed impairments when the contingencies reversed (Pryce, Dettling, Spengler, Schnell, & Feldon, 2004). This is similar to that seen in

both depressed and bipolar patients in the human PRLT (Gorrindo et al., 2005; Murphy et al., 2003) who acquire the initial rule successfully but then are impaired following reversal. This compares to ELS participants in the present study who performed equally well in the practice phase and reversal phases but showed a deficit in acquisition of the first rule in block 1. This suggests a potential impairment in the ability to generalise the task rules between the practice and acquisition phase. However previous probabilistic learning studies did not include a practice phase meaning that this likely changed the way participants processed the start of block 1. This might explain the contrast with Pechtel and Pizzagalli, 2013 who reported that women with remitted MDD and ELS learnt acquisition in the PSST at the same rate as controls.

One of the hypotheses of this study was that stress in adult life would modulate the relationship between reward processing deficits and ELS. There was little evidence that this was the case except for an observed interaction between PFS and ELS whereby stress only influenced PFS in participants without a history of ELS. Higher lifetime stress led to greater PFS but higher stress in the last year was associated with decreased win-stay probability. There are few previous studies investigating similar constructs but Berghorst et al., 2013 reported that after stress induction those who had higher cortisol reactivity and self-reported negative affect had lower reward but not punishment sensitivity. Additionally, it is worth noting that due to the relatively poor model fit for this exploratory analysis that these findings should be taken as preliminary due to the risk of data overfitting.

4.2 Probabilistic reward task

In contrast to previous studies employing the PRT neither groups showed a response bias toward the more highly rewarded stimulus (Pizzagalli et al., 2005, 2008) suggesting a general failure of all participants to modulate their responses as a function of reward. There are no previous studies carrying out the PRT online but in this study we failed to replicate the main outcome measure. All aspects of the task were similar between the lab and online version except for participants being informed high performance would lead to a bonus payment with the actual reward in the task being points. Previous studies instead used direct monetary compensation in the task (Pizzagalli et al., 2005). It should also be noted that the online testing platform limits the ability to ensure that participants are completing the tasks in as controlled an environment as would be possible by laboratory testing providing another explanation for the high data variability. The lack of response bias indicates that participants solved the task in a different manner using potentially different cognitive processes making comparison to previous literature challenging. Nevertheless, participants with high levels of ELS did show impairments in discrimination, a measure of task difficulty which appeared to be driven by changes in depression symptomology as opposed to ELS specifically.

4.3 Conclusions

These data suggest that participants without a formal diagnosis of a mental health condition but a history of ELS show impairments in positive feedback sensitivity and reward learning in

the PRLT compared to controls. These impairments may be important in understanding how ELS predisposes to depression with reduced reward learning being a key feature in MDD patients (Halahakoon et al., 2020). However, high levels of undiagnosed depression are a potential confound and highlight a potential wider issue in terms of the number of people who meet criteria for MDD but are not formally diagnosed or receiving care. Future studies are needed to replicate these findings, investigate the neural circuit changes underlying these reward learning impairments and investigate whether these findings are directly related to psychiatric risk.

5. Acknowledgements

The authors would like to thank Chloe Slaney for helpful discussions regarding experimental design, ethical approval and data analysis in addition to Michelle Taylor for advice on statistical analysis.

6. Declarations

- The funding for this study was provided by the BBSRC SWBio DTP PhD program (Grant numbers: BB/J014400/1 and BB/M009122/1) awarded to MPW.
- The authors declare no conflict of interest.

- All procedures were approved by the Faculty of Life Sciences and Faculty of Science Research Ethics Committee at the University of Bristol
- All participants consented to participation both prior to and following completion of the study
- All participants consented to publication of anonymised data
- Data will be made available open access at www.osf.io/63e8j and by contacting the corresponding author
- Code will be made available open access at www.osf.io/63e8j and by contacting the corresponding author

7. Open Practices statement

This study was pre-registered and available at www.osf.io/538yk. Data and code will also be made available open access at www.osf.io/63e8j.

8. References

Adler, N. E., Epel, E. S., Castellazzo, G., & Ickovics, J. R. (2000). Relationship of subjective and objective social status with psychological and physiological functioning: preliminary data in healthy white women. *Health Psychology : Official Journal of the Division of Health Psychology, American Psychological Association*, *19*(6), 586–592. https://doi.org/10.1037//0278-6133.19.6.586

Agid, O., Shapira, B., Zislin, J., Ritsner, M., Hanin, B., Murad, H., … Lerer, B. (1999).

Environment and vulnerability to major psychiatric illness: a case control study of early parental loss in major depression, bipolar disorder and schizophrenia. *Molecular Psychiatry*, *4*(2), 163. https://doi.org/10.1038/sj.mp.4000473

Ameli, R., Luckenbaugh, D. A., Gould, N. F., Holmes, M. K., Lally, N., Ballard, E. D., & Zarate Jr, C. A. (2014). SHAPS-C: the Snaith-Hamilton pleasure scale modified for clinician administration. *PeerJ*, *2*, e429–e429. https://doi.org/10.7717/peerj.429

Beck, A., Steer, R., & Brown, G. (1996). *Beck Depression Inventory. Second Edition*. San Antonio, TX.

Berghorst, L. H., Bogdan, R., Frank, M. J., & Pizzagalli, D. A. (2013). Acute stress selectively reduces reward sensitivity. *Frontiers in Human Neuroscience*, *7*(April), 1–15. https://doi.org/10.3389/fnhum.2013.00133

Bogdan, R., & Pizzagalli, D. A. (2006). Acute Stress Reduces Reward Responsiveness: Implications for Depression. *Biological Psychiatry*, *60*(10), 1147–1154.

Bress, J. N., Foti, D., Kotov, R., Klein, D. N., & Hajcak, G. (2013). Blunted neural response to rewards prospectively predicts depression in adolescent girls. *Psychophysiology*, *50*(1), 74–81. https://doi.org/10.1111/j.1469-8986.2012.01485.x

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., … Bolker, B. M. (2017). glmmTMB balances speed and flexibility among packages for zero-inflated generalized linear mixed modeling. *R Journal*, *9*(2), 378–400. https://doi.org/10.32614/rj-2017-066

Chen, C., Takahashi, T., Nakagawa, S., Inoue, T., & Kusumi, I. (2015). Reinforcement learning in depression: A review of computational research. *Neuroscience and Biobehavioral*

Reviews, *55*, 247–267. https://doi.org/10.1016/j.neubiorev.2015.05.005

Cohen, R. A., Hitsman, B. L., Paul, R. H., McCaffery, J., Stroud, L., Sweet, L., … Gordon, E. (2006). Early life stress and adult emotional experience: An international perspective. *International Journal of Psychiatry in Medicine*, *36*(1), 35–52. https://doi.org/10.2190/5R62-9PQY-0NEL-TLPA

Cohodes, E. M., Kitt, E. R., Baskin-Sommers, A., & Gee, D. G. (2020). Influences of early-life stress on frontolimbic circuitry: Harnessing a dimensional approach to elucidate the effects of heterogeneity in stress exposure. *Developmental Psychobiology*, (September 2019), 1–20. https://doi.org/10.1002/dev.21969

Cools, R., Clark, L., Owen, A. M., & Robbins, T. W. (2002). Defining the Neural Mechanisms of Probabilistic Reversal Learning Using Event-Related Functional Magnetic Resonance Imaging. *Journal of Neuroscience*, *22*(11), 4563–4567. https://doi.org/10.1523/jneurosci.22-11-04563.2002

Elliott, R., Sahakian, B. J., Herrod, J. J., Robbins, T. W., & Paykel, E. S. (1997). Abnormal response to negative feedback in unipolar depression: evidence for a diagnosis specific impairment. *Journal of Neurology, Neurosurgery & Psychiatry*, *63*(1), 74–82. https://doi.org/10.1136/jnnp.63.1.74

Foti, D., & Hajcak, G. (2009). Depression and reduced sensitivity to non-rewards versus rewards : Evidence from event-related potentials. *Biological*, *81*, 1–8. https://doi.org/10.1016/j.biopsycho.2008.12.004

Gorrindo, T., Blair, R. J. R., Budhani, S., Dickstein, D. P., Pine, D. S., & Leibenluft, E. (2005). Deficits on a probabilistic response-reversal task in patients with pediatric bipolar

disorder. *American Journal of Psychiatry*, *162*(10), 1975–1977. https://doi.org/10.1176/appi.ajp.162.10.1975

Green, J. G., McLaughlin, K. A., Berglund, P. A., Gruber, M. J., Sampson, N. A., Zaslavsky, A. M., & Kessler, R. C. (2010). Childhood adversities and adult psychopathology in the National Comorbidity Survey Replication (NCS-R) I: Associations with first onset of DSM-IV disorders. *Archives of General Psychiatry*, *67*(2), 113. https://doi.org/10.1016/j.pestbp.2011.02.012.Investigations

Grogan, J. P., Tsivos, D., Smith, L., Knight, B. E., Bogacz, R., Whone, A., & Coulthard, E. J. (2017). Effects of dopamine on reinforcement learning and consolidation in Parkinson ' s disease. *ELife*, *6*, 1–23. https://doi.org/10.7554/eLife.26801

Halahakoon, D. C., Kieslich, K., O'Driscoll, C., Nair, A., Lewis, G., & Roiser, J. P. (2020). Reward Processing Behavior in Depressed Participants Relative to Healthy Volunteers: A Systematic Review and Meta-analysis. *JAMA Psychiatry*. https://doi.org/10.1001/jamapsychiatry.2020.2139

Hammen, C., Henry, R., & Daley, S. E. (2000). Depression and sensitization to stressors among young women as a function of childhood adversity. *Journal of Consulting and Clinical Psychology*, *68*(5), 782–787. https://doi.org/10.1037//0022-006X.68.5.782

Hanson, J. L., Albert, D., Iselin, A.-M. R., Carré, J. M., Dodge, K. A., & Hariri, A. R. (2016). Cumulative stress in childhood is associated with blunted reward-related brain activity in adulthood. *Social Cognitive and Affective Neuroscience*, *11*(3), 405–412. https://doi.org/10.1093/scan/nsv124

Hanson, J. L., Nacewicz, B. M., Sutterer, M. J., Cayo, A. A., Schaefer, S. M., Rudolph, K. D., …

Davidson, R. J. (2015). Behavioral problems after early life stress: Contributions of the hippocampus and amygdala. *Biological Psychiatry*, *77*(4), 314–323. https://doi.org/10.1016/j.biopsych.2014.04.020

Hanson, J. L., van den Bos, W., Roeber, B. J., Rudolph, K. D., Davidson, R. J., & Pollak, S. D. (2017). Early adversity and learning: implications for typical and atypical behavioral development. *Journal of Child Psychology and Psychiatry and Allied Disciplines*, *58*(7), 770–778. https://doi.org/10.1111/jcpp.12694

Herzallah, M., Moustafa, A., Natsheh, J., Abdellatif, S., Taha, M., Tayem, Y., … Gluck, M. (2013). Learning from negative feedback in patients with major depressive disorder is attenuated by SSRI antidepressants . *Frontiers in Integrative Neuroscience* , Vol. 7, p. 67.

Holmes, T. H., & Rahe, R. H. (1967). The social readjustment rating scale. *Journal of Psychosomatic Research*, *11*(2), 213–218. https://doi.org/10.1016/0022-3999(67)90010-4

Kessler, R. C., Berglund, P., Demler, O., Jin, R., Merikangas, K. R., & Walters, E. E. (2005). Lifetime Prevalence and Age-of-Onset Distributions of DSM-IV Disorders in the National Comorbidity Survey Replication. *Archives of General Psychiatry*, *62*(6), 593–602. https://doi.org/10.1001/archpsyc.62.6.593

Lemoult, J., Humphreys, K. L., Tracy, A., Hoffmeister, J., Ip, E., & Gotlib, I. H. (2019). Meta-Analysis: Exposure to Early Life Stress and Risk for Depression in Childhood and Adolescence. *Journal of the American Academy of Child & Adolescent Psychiatry*. https://doi.org/10.1016/j.jaac.2019.10.011

Lewis, K., Marrie, R. A., Bernstein, C. N., Graff, L. A., Patten, S. B., Sareen, J., … Bolton, J. M. (2019). The Prevalence and Risk Factors of Undiagnosed Depression and Anxiety Disorders Among Patients With Inflammatory Bowel Disease. *Inflammatory Bowel Diseases*, *25*(10), 1674–1680. https://doi.org/10.1093/ibd/izz045

Li, C., Ford, E. S., Zhao, G., Ahluwalia, I. B., Pearson, W. S., & Mokdad, A. H. (2009). Prevalence and correlates of undiagnosed depression among U.S. adults with diabetes: The Behavioral Risk Factor Surveillance System, 2006. *Diabetes Research and Clinical Practice*, *83*(2), 268–279. https://doi.org/10.1016/j.diabres.2008.11.006

Lotfaliany, M., Bowe, S. J., Kowal, P., Orellana, L., Berk, M., & Mohebbi, M. (2018). Depression and chronic diseases: Co-occurrence and communality of risk factors. *Journal of Affective Disorders*, *241*, 461–468. https://doi.org/10.1016/j.jad.2018.08.011

McCauley, J., Kern, D. E., Kolodner, K., Dill, L., Schroeder, A. F., DeChant, H. K., … Bass, E. B. (1997). Clinical characteristics of women with a history of childhood abuse: unhealed wounds. *Jama*, *277*(17), 1362–1368. https://doi.org/10.1001/jama.277.17.1362

Millisecond. (2020a). *Probabilistic Reversal Learning Task*.

Millisecond. (2020b). *Probabilistic Reward Task*.

Mueller, E. M., Pechtel, P., Cohen, A. L., Douglas, S. R., & Pizzagalli, D. A. (2015). Potentiated processing of negative feedback in depression is attenuated by anhedonia. *Depression and Anxiety*, *32*(4), 296–305. https://doi.org/10.1002/da.22338

Murphy, F. C., Michael, A., Robbins, T. W., & Sahakian, B. J. (2003). Neuropsychological impairment in patients with major depressive disorder: The effects of feedback on task performance. *Psychological Medicine*, *33*(3), 455–467.

https://doi.org/10.1017/S0033291702007018

Murphy, F. C., Michael, A., & Sahakian, B. J. (2012). Emotion modulates cognitive flexibility in patients with major depression. *Psychological Medicine*, *42*(7), 1373–1382. https://doi.org/10.1017/S0033291711002418

Office for National Statistics. (2020). *Coronavirus and depression in adults, Great Britain: June 2020*.

Pechtel, P., Dutra, S. J., Goetz, E. L., & Pizzagalli, D. A. (2013). Blunted reward responsiveness in remitted depression. *Journal of Psychiatric Research*, *47*(12), 1864–1869. https://doi.org/10.1016/j.jpsychires.2013.08.011

Pechtel, P., & Pizzagalli, D. A. (2013). Disrupted reinforcement learning and maladaptive behavior in women with a history of childhood sexual abuse: A high-density event-related potential study. *JAMA Psychiatry*, *70*(5), 499–507. https://doi.org/10.1001/jamapsychiatry.2013.728

Peterson, R. A., & Cavanaugh, J. E. (2019). Ordered quantile normalization: a semiparametric transformation built for the cross-validation era. *Journal of Applied Statistics*, 1–16. https://doi.org/10.1080/02664763.2019.1630372

Pizzagalli, D. A., Bogdan, R., Ratner, K. G., & Jahn, A. L. (2007). Increased Perceived Stress is Associated with Blunted Hedonic Capacity: Potential Implications for Depression Research. *Behav Res Ther*, *45*(11), 2742–2753.

Pizzagalli, D. A., Jahn, A. L., & O'Shea, J. P. (2005). Toward an objective characterization of an anhedonic phenotype: A signal-detection approach. *Biological Psychiatry*, *57*(4), 319–327. https://doi.org/10.1016/j.biopsych.2004.11.026

Pizzagalli, D. A., Losifescu, D., Hallet, L. A., Ratner, K. G., & Fava, M. (2008). Reduced Hedonic Capacity in Major Depressive Disorder: Evidence from a Probabilistic Reward Task. *J Psychiatr Res*, *43*(1), 76–87. https://doi.org/10.1016/j.jpsychires.2008.03.001.Reduced

Pryce, C. R., Dettling, A. C., Spengler, M., Schnell, C. R., & Feldon, J. (2004). Deprivation of Parenting Disrupts Development of Homeostatic and Reward Systems in Marmoset Monkey Offspring. *Biological Psychiatry*, *56*, 72–79. https://doi.org/10.1016/j.biopsych.2004.05.002

R Core Team. (2020). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing.

Robinson, O. J., & Chase, H. W. (2017). Learning and Choice in Mood Disorders: Searching for the Computational Parameters of Anhedonia. *Computational Psychiatry (Cambridge, Mass.)*, *1*(1), 208–233. https://doi.org/10.1162/CPSY_a_00009

Snaith, R. P., Hamilton, M., Morley, S., Humayan, A., Hargreaves, D., & Trigwell, P. (1995). A scale for the assessment of hedonic tone the Snaith-Hamilton Pleasure Scale. *The British Journal of Psychiatry : The Journal of Mental Science*, *167*(1), 99–103. https://doi.org/10.1192/bjp.167.1.99

Taylor Tavares, J. V., Clark, L., Furey, M. L., Williams, G. B., Sahakian, B. J., & Drevets, W. C. (2008). Neural basis of abnormal response to negative feedback in unmedicated mood disorders. *NeuroImage*, *42*(3), 1118–1126. https://doi.org/10.1016/j.neuroimage.2008.05.049

Tottenham, N. (2009). A review of adversity, the amygdala and the hippocampus: a consideration of developmental timing. *Frontiers in Human Neuroscience*, *3*(January),

68. https://doi.org/10.3389/neuro.09.068.2009

Vrieze, E., Pizzagalli, D. A., Demyttenaere, K., Hompes, T., Sienaert, P., De Boer, P., … Claes, S. (2013). Reduced reward learning predicts outcome in major depressive disorder. *Biological Psychiatry*, *73*(7), 639–645. https://doi.org/10.1016/j.biopsych.2012.10.014

Waegeman, A., Declerck, C. H., Boone, C., Seurinck, R., & Parizel, P. M. (2014). Individual differences in behavioral flexibility in a probabilistic reversal learning task: An fMRI study. *Journal of Neuroscience, Psychology, and Economics*, Vol. 7, pp. 203–218. https://doi.org/10.1037/npe0000026

Webb, C. A., Auerbach, R. P., Bondy, E., Stanton, C. H., Foti, D., & Pizzagalli, D. A. (2017). Abnormal neural responses to feedback in depressed adolescents. *Journal of Abnormal Psychology*, *126*(1), 19–31. https://doi.org/10.1037/abn0000228

Whitton, A. E., Treadway, M. T., & Pizzagalli, D. A. (2015). Reward processing dysfunction in major depression, bipolar disorder and schizophrenia. *Current Opinion in Psychiatry*, *28*(1), 7–12. https://doi.org/10.1097/YCO.0000000000000122

Wilkinson, M. P., Grogan, J. P., Mellor, J. R., & Robinson, E. S. J. (2020). Comparison of conventional and rapid-acting antidepressants in a rodent probabilistic reversal learning task. *Brain and Neuroscience Advances*, *4*, 1–11. https://doi.org/10.1177/2398212820907177

Zhou, X., Meng, Y., Schmitt, H. S., Montag, C., Kendrick, K. M., & Becker, B. (2020). Cognitive flexibility mediates the association between early life stress and habitual behavior. *Personality and Individual Differences*, *167*, 110231. https://doi.org/https://doi.org/10.1016/j.paid.2020.110231