Acute stress effects on probabilistic reversal learning in healthy participants

Lara Wieland[1], Claudia Ebrahimi[1], Teresa Katthagen[1], Martin Panitz[2], Lennart Luettgau[3], Andreas Heinz[1], Florian Schlagenhauf[1]*, Zsuzsika Sjoerds[4]* (*shared last author ship)

Affiliations:

[1]Charité-Universitätsmedizin Berlin, Department of Psychiatry and Neurosciences | CCM, Berlin, Germany

[2]Max-Planck-Institute for Human Cognitive and Brain Sciences, Leipzig, Germany

[3]Max Planck University College London Centre for Computational Psychiatry and Ageing Research, London, UK

[4]Cognitive Psychology Unit, Institute of Psychology & Leiden Institute for Brain and Cognition, Leiden University, Leiden, the Netherlands

Corresponding author:

Dr. Zsuzsika Sjoerds

Leiden University, Institute of Psychology

Cognitive Psychology Unit

Wassenaarseweg 52

2333 AK Leiden, The Netherlands

E-mail: z.sjoerds@fsw.leidenuniv.nl

Keywords: Stress, fMRI, cognitive flexibility, reversal learning, decision-making, modeling

Text no. of words: 7036

no. of references: 48

Abstract

Behavioral adaptation is a fundamental cognitive ability, ensuring an organism's survival by allowing for flexible adjustment to changing environmental conditions. These adaptive abilities can be measured using reversal learning paradigms requiring agents to adjust their reward learning to sudden changes in stimulus-action-outcome contingencies. Stressful situations have been found to alter flexibility of reward learning, but directionality of effects has been mixed across studies. Here, we used functional MRI (fMRI) informed by computational modeling in a within-subjects design with healthy male human volunteers to investigate the effect of acute psychosocial stress on flexible behavioral adaptation. Participants (n=28) underwent fMRI during a reversal learning task, once after the Trier Social Stress Test (TSST), a validated psychosocial stress induction method, and once after a control condition, in two separate sessions. Effects of stress on choice behavior were investigated using multilevel generalized linear models and a set of computational models describing different learning processes that might have generated the data. Computational models were fitted using a hierarchical Bayesian approach, and model-derived reward prediction errors (RPE) were used as regressors for fMRI analyses. We found that acute psychosocial stress only slightly increased correct response rates in our participants. Model comparison revealed that double-update learning with stress-specific scaling of the inverse decision temperature parameter best explained the observed behavior under stress. On the neural level, RPE signals were represented in striatum and ventromedial prefrontal cortex (vmPFC). No whole-brain correctable effects of stress on RPE representations were found. Our study suggests that acute psychosocial stress does not alter neural representation of RPE and that interindividual variability on the behavioral level might be more related to use of choice values, expressed by the temperature parameters.

1. Background

Humans and other agents are routinely confronted with decision-making situations under stress, for example when choosing an efficient and cheap way of commuting to work, despite running late. Different choice options, such as taking the car, bike or train, are associated with relatively stable and predictable levels of cost and reward. In contrast, the weather forecast of the day, a congestion on the preferred route or a train delay, are more uncertain, less predictable factors. Both, stable and uncertain factors interact, in that cycling to work may be rewarding in sunny weather but not on a rainy day. Stress impacts individuals' emotions, mood, physiological responses and may affect their cognitive processing resources, influencing their decision-making strategies (Lupien et al., 2007). This might be especially relevant in situations that afford high behavioral flexibility, for instance in constantly changing environments. Stress is also an important factor in causing and maintaining psychiatric conditions (McEwen, 2004) and strongly influences health-related behavior in general (Cohen et al., 2016). Therefore, the development of a model of how stress affects choice behavior in healthy individuals is pivotal for a mechanistic understanding of maladaptive behavior in a spectrum from daily mistakes to psychiatric disorders.

Flexible decision-making requires one to learn what is most rewarding in the current environment and adapt one's decision-making to that. Studies have found mixed results for the influence of stress on decision-making, ranging from beneficial to detrimental effects across paradigms (Goldfarb et al., 2015; Plessow et al., 2012, 2011). In a meta-analysis, acute stress showed a small negative impact for tasks in which reward seeking and risk taking is disadvantageous (d = .26 and d = .44), but showed no effect if this was not the case (Starcke and Brand, 2016). Similarly, a meta-analysis over fewer studies investigating the effects of acute stress on cognitive flexibility concluded that stress had a small impairing effect (Hedges' g= -.30) (Shields et al., 2016). Different processes involved in decision-making are presumably differentially prone to interruption by stress (Schwabe and Wolf, 2011, 2009). Whereas habitual decision-making relies on simple stimulus-related associations, goal-directed decision-making associates actions with a motivational value and is therefore more flexible but also computationally more costly. It has been found that acute and chronic stress disrupt goal-directed decision-making, while habitual decision-making appears unaffected at the behavioral as well as neural level (Schwabe et al., 2013, 2008). One possible explanation for the variable findings are different types of standardized stressors, which are commonly used in behavioral experiments. They can be physiological as in the Cold Pressor Task, psychosocial as in the Trier Social Stress Test (TSST) or both as in the Socially Evaluated Cold Pressor Test (Starcke and Brand, 2016). The physiological paradigms lead to more immediate stress during learning, whereas the psychosocial paradigms release their full physiological effect 10-20 mins after stress induction. Another explanation for the inconsistent meta-analytical findings could lie in how cognitive flexibility was measured. Both meta-analyses predominantly focused on classical paradigms such as the Wisconsin card sorting test or task-switching tests. While providing valuable insight into overall cognitive flexibility, these paradigms mostly rely on averaged outcome measures. In contrast, tasks designed for computational modeling may provide a more fine-grained measure of behavioral adaptation.

An understudied subject remains how the brain adapts to learning from rewards in a changing environment under stress. Probabilistic reversal learning requires participants to choose between stimuli with varying reward contingencies. In these paradigms contingencies are reversed several times throughout the task unannounced and therefore demand behavioral adaptation to a changing environment. A computational mechanism underlying the putative learning process can be formalized by the reward prediction error (RPE), a computational quantity derived from the reinforcement learning (RL) framework. RPE signal the difference between an observed and expected reward (Dolan and Dayan, 2013) and are used to update the value of a stimulus, a state, or an action. The neural signature of RPE during reversal learning is reliably found in the human ventral frontostriatal circuitry (Doherty et al., 2003).

So far, small sample sizes, heterogenous subdomains in the operationalization of decision-making, and methodological considerations regarding the type of stressor have complicated the picture (Porcelli and Delgado, 2017). Most previous studies on stress effects on decision-making have employed between-subject designs – but subjects vary drastically in both individual stress responses, choice behavior and how stress affects performance. In the previously used between-subject designs it thus remains unclear, how much of stress-related changes to the neural correlates of probabilistic reversal learning can be attributed to the stressor and how much may be related to interindividual differences in stress reactivity. The few studies using within-subjects designs to investigate learning are either purely behavioral (Radenbach et al., 2015a) or employ electroencephalography (Cavanagh et al., 2011), lacking the possibility of precise spatial signal localization and anatomical specificity with respect to the neural representation of RPE signals. Additionally, few studies in the realm of cognitive flexibility use computational modeling to elucidate underlying cognitive mechanisms. Here, we used fMRI and a psychosocial stress intervention to study probabilistic reversal learning in healthy male participants using a within-subjects design. Applying a state-of-the-art hierarchical Bayesian modeling approach (Piray et al., 2019) allowed us to model the impact of stress on behavioral adaptation.

2. Methods

2.1. Study Design:

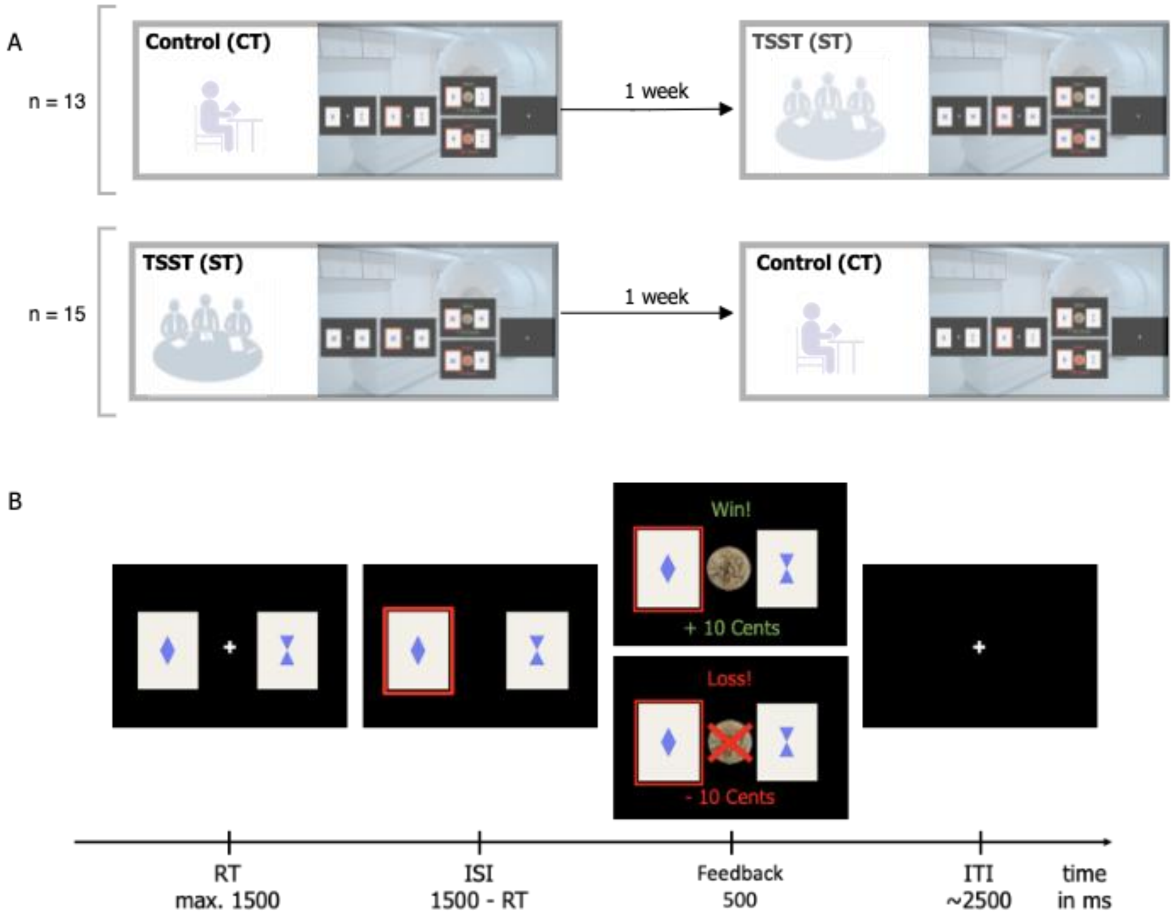


Figure 1. Study design (A) and task design (B)

Employing a within-subjects design, 38 healthy male adult participants (n=28 in the final analyzed sample) performed a probabilistic reversal learning task during fMRI in two separate sessions seven days apart. Procedures and materials are identical with a previous study from our laboratory using another paradigm (Luettgau et al., 2018). During the stress condition, participants were exposed to a mock interview and calculus in front of a socially unresponsive committee in white lab coats, following the standardized Trier Social Stress Test (TSST) protocol (Kirschbaum et al., 1993). During the control condition, participants read a neutral text in absence of the committee (see Supplement). Order of session type (stress vs. control) was counter-balanced across participants. In order to prevent confounding effects of circadian rhythm on cortisol levels (Kudielka et al., 2004), both experimental sessions were scheduled at the same time of the day. Acute stress responses were assessed at physiological (cortisol) and subjective (self-report) levels at six time points throughout the session (Figure 3).

2.2. Physiological stress response:

We assessed physiological stress response via salivary cortisol, measured six times throughout the experiment at the following time points relative to the start of intervention (stress or control): t1: -30 minutes; t2: -2 minutes; t3: +10 minutes; t4: +15 minutes; t5: +30 minutes; t6: +45 minutes (Luettgau et al., 2018). For collection and extraction of saliva we used Salivette saliva sampling tubes (SalivetteCortisol®, Sarstedt, Nuembrecht, Germany) (see Supplement). Individual cortisol reactivity was determined by calculating the area under the curve (AUC) with respect to ground (AUCg-stress and AUCg-control, see Pruessner et al., 2003) separately for both conditions and subtracting AUCg-control from AUCg-stress. The AUC was calculated based on individual subject-wise time points, taking into account slight temporal dispersion in the testing protocol. For an additional analysis to confirm stress reactivity please refer to the supplement (subsection: physiological stress response).

2.3. Subjective stress response:

Three different visual analogue scales (VAS) ranging from 0 to 100 were used to assess subjective arousal, valence and stress at all time points (T1-T6). Participants were asked to rate how they felt, regarding arousal on a scale "Please rate your current state" from 0 (sleepy) to 100 (active), valence on a scale from 0 (unhappy) to 100 (happy) and stress on a scale from 0 (not stressed) to 100 (stressed). Analogue to cortisol values this was determined by calculating the area under the curve with respect to ground (AUCg-stress and AUCg-control; Pruessner et al., 2003) separately for both conditions and subtracting AUCg-control from AUCg-stress.

2.4. Working memory capacity:

Participants also performed the digit span backwards task from the test battery Hamburg-Wechsler-Intelligenztest HAWIK (Tewes and Wechsler, 1991) to assess working memory capacity.

2.5. Task Design

Participants performed a probabilistic reversal learning task, which included 160 trials and comprised around 15 minutes. The task (Boehme et al., 2015; Reiter et al., 2016a) was programmed in Matlab (The MathWorks, Natick, MA) with Psychtoolbox (Brainard, 1997). On every trial, participants chose between two cards, each depicting a different geometric figure. The underlying reward structure was not explicitly instructed but had to be inferred: reward probabilities associated with the two choice options were anti-correlated (i.e. when card A had a reward probability of 80% and therefore a punishment probability of 20%, card B had a reward probability of 20% and a punishment probability of 80% and vice versa). Furthermore, participants were informed of the probabilistic nature of the task but not on the actual probabilities: the currently "better" card was only rewarded in 80% of all trials with 10 Cent. Right-side versus left-side location of the stimulus was randomized on each trial. After a fixed number of 55 trials, contingencies reversed and these reversals repeated several times over the middle experimental phase, followed by another stable phase in the end starting at trial 126 (see Figure 2). Participants were instructed to win as much money as possible and received the winnings at the end of the experiment.

Because feedback was drawn probabilistically on each trial, and we wanted to ensure that the number of probabilistic events was matched between the control and the stress condition six participants had to be excluded from the final sample to avoid confounds due to different task environments. Additionally, two participants had to be excluded due to technical failure and two additional participants had to be excluded because they performed the task below chance level, leaving a total of 28 participants for final analyses.

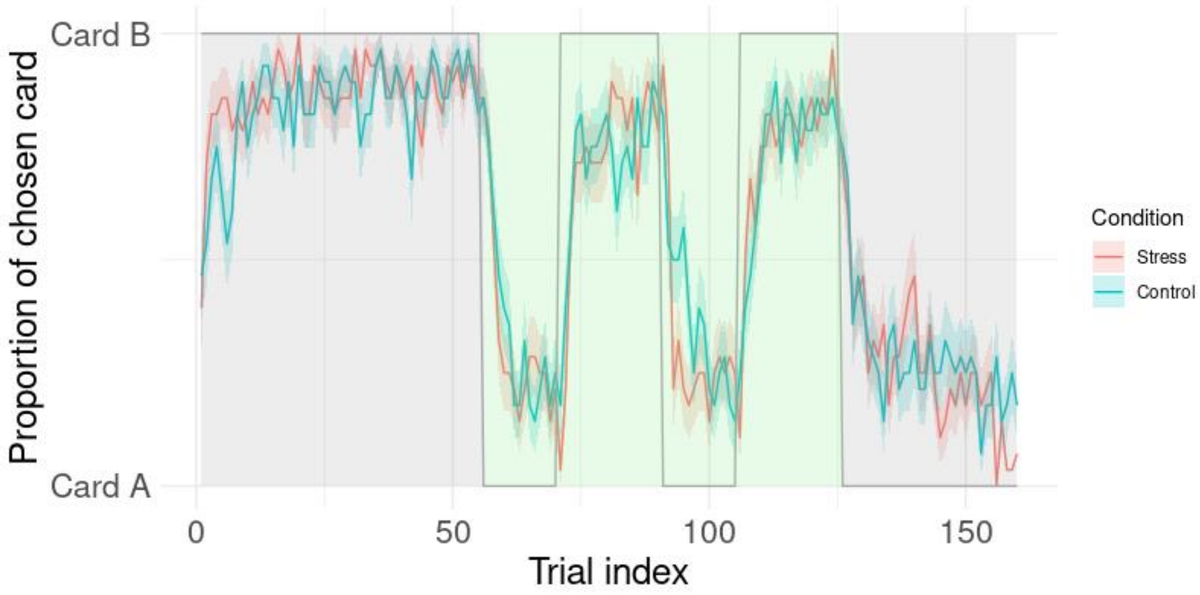


Figure 2. Empirical choice behavior in both conditions (lines showing the mean percentage of chosen card for stress (ST) in red and control (CT) in blue and shaded red and blue areas showing standard errors) with underlying task structure in grey line and shaded areas in grey for stable and light green for volatile phases.

2.6. Analyses

2.6.1. Stress response analyses

Cortisol responses (AUC-g) and the three subjective VAS scales were compared across conditions (stress vs control) using one-tailed paired-sample *t*-tests at a significance level of $p < .05$.

2.6.2. Behavioral data

Single-trial multilevel generalized linear models (logistic regressions) were conducted using the lme4 package (Bates et al., 2015) in R (Version 4.0.3). Parameter estimates were considered significant at $p \leq .05$. We analyzed trial-by-trial correct responses (choose better option), win-stay (select same stimulus after win) and lose-switch (switch stimulus after loss) behavior with the factors *stress condition* (CT vs. ST, effect coding as -0.5 and 0.5) and *experimental phase* (pre, reversal, post) as fixed effects and subject as a random effect, allowing for an individually varying intercept per subject. For the factor *experimental phase* we specified a custom centered contrast, testing the null hypothesis of performance differences between first stable vs. reversal and late stable vs. reversal phase using the hypr package (Rabe et al., 2020). Main effects of condition and phase, as well as an interaction effect were added incrementally in two steps. We used χ^2-tests based on log-likelihood changes to compare a null model, which predicted outcome variables with the individually varying intercept per subject to a model including varying intercepts and all main effects. If this showed a significant better fit, we compared the main effect model to an interaction effect model. For the best-fitting model, the parameter estimates' odd's ratio was computed to assess effect size. Additionally, we performed the same analysis using the cortisol AUC-g values instead of condition labels as predictor. Participants were excluded when their performance was below chance (correct responses < 50%), as described in the methods section 2.5. Across all trials, participants missed a relatively low number of trials (0.71%).

Furthermore, as an exploratory analysis on the potential moderating impact of chronic stress exposure as well as working memory performance (Otto et al., 2013; Radenbach et al., 2015a) we dichotomized the sample into improved vs. impaired learning performance based on total correct responses and between time points/sessions (delta correct response > 0 vs. delta correct response < 0). For these groups, we conducted independent t-tests on chronic stress exposure, as well as working memory performance. Due to missing values for four participants, regarding the PSS-10, the latter analysis was conducted with a reduced sample of 24 participants.

2.6.3 Computational models

In order to describe different learning processes that might have generated the data under the stress versus the control condition, we set up the following model space. It comprised Rescorla-Wagner (RW), Pearce-Hall (PH; (Pearce and Hall, 1980) models and a null model (no-learning). In the RW and PH models, the expected value $Q_{a,t}$ of an action a at trial t is updated via the RPE $\delta_{Q_{a,t}}$ (eq. 1), which is defined as the difference between received reward R_t and previously expected reward value for the chosen stimulus $Q_{a,t}$(eq. 2):

(1) $\delta_{Q_{a,t}} = R_t - Q_{a,t}$

(2) $Q_{a,t+1} = Q_{a,t} + \alpha_{\mathrm{win/loss}}\delta_{Q_{a,t}}$

(3) $Q_{ua,t+1} = Q_{ua,t} + \kappa * \alpha_{\mathrm{win/loss}} * \delta_{Q_{ua,t}}$

In RW models, we accounted for learning about the unchosen option as indicated by the implicit anti-correlated task structure in different sub-models (eq. 3, $\kappa = 0$ for single update (SU), $\kappa = 1$ for full double update (DU) and freely fitted κ for individually weighted double update (iDU)). We further varied whether learning rates α differed for wins and losses. The PH model encompasses eq. 1 and 2 with a dynamic learning rate depending on a decay over time as and the absolute prediction error (see Supplement or Pearce and Hall, 1980). In the no-learning model, a stable bias towards one of the stimuli was implemented (Supplement). For all learning models, trial-wise Q-action values are transformed into choice probabilities by a softmax response model with different inverse decision noise temperatures β following wins and losses:

(4) $\mathrm{p}(a_i) = \frac{exp\left(\beta_{\mathrm{win/loss}} Q_{a_i}\right)}{\sum_{j=1}^{K} exp\left(\beta_{\mathrm{win/loss}} Q_{a_j}\right)}$

The free softmax temperature parameter β reflects choice stochasticity with higher values equating more deterministic and lower values equating more stochastic choices.

We followed a two-step procedure: First, we fit our model space to the behavioral data of the control condition. Then, the best fitting model from the control condition was used for modelling behavior under stress now with additional 'stress weights' on the free parameters. Taken together, the 'step 1 model space' consisted of 8 models for learning under the control condition: RW-SU-1al, RW-SU-2al, RW-DU-1al, RW-DU-2al, RW-iDU-1al, RW-iDU-1al, PH and no-learning. We applied Bayesian model comparison (Piray et al., 2020) to find out which of these models explained the data best (see protected exceedance probabilities (PXP) in Figure 5).

To model learning under the stress condition, we added stress weights to the free parameters of the best-fitting model from the first step (RW-DU-2al). The 'step 2 model space' included the DU-2al model without stress effects (RW-DU-2al-NoStress), one with stress weights affecting only the learning parameters α_{win} and α_{loss} (RW-DU-2al-StressLearning) one model with stress only affecting the temperature parameters β_{win} and β_{loss}, (RW-DU-2al-StressBetas) and a full model with stress affecting all free parameter (RW-DU-2al-StressAll). This model space was fitted to combined data from both conditions: trials were concatenated across control and stress conditions within subjects, with the free stress parameters quantifying the additive effect on the respective parameters for the trials of the stress condition. As in step 1, model fits were then compared between models.

2.6.4. Model fitting

Models from both steps were fitted under the hierarchical Bayesian inference approach as implemented in the cbm toolbox (Piray et al., 2020) run in in Matlab R2018a. This procedure allowed for concurrent model comparison and parameter estimation. Thereby, the latter also followed a multilevel modelling approach: the group mean parameter affects individual parameter estimation and vice versa, but the relationship is scaled by how (relatively) well the model explains the individual subject's behavior.

2.6.5. fMRI data

Scans were acquired on a Siemens 3 T high-resolution PRISMA MR-System with a 20-channel head coil (Siemens, Erlangen, Germany). Covering the whole brain, 40 slices were acquired in oblique orientation at 20° to the anterior commissure-posterior commissure line and in ascending order with the following parameters: T2*-weighted gradient-echo echo-planar imaging (EPI) (TR: 2.09 s; TE: 22 ms; flip angle: 90°; 3 × 3 mm^2 in-plane voxel resolution, 0.5 mm gap between slices, voxel size, 3 × 3 × 4 mm) and a field map to account for individual homogeneity differences of the magnetic field. The scanning procedure further comprised a T1-weighted MPRAGE recorded within seven days before the first test session. fMRI data were preprocessed and analyzed using SPM8 (http://www.fil.ion.ucl.ac.uk/spm/) in Matlab. The first 5 volumes of each functional time series were discarded. Before preprocessing, the origins of the functional imaging series were reoriented to the anterior–posterior commissure plane in native space. Preprocessing included slicetiming, realignment, coregistration, and warping to Montreal Neurological Institute (MNI) space. The obtained normalization parameters were applied to the realigned images, which were resliced with a voxel size of 3 x 3 x 4 mm. All images were smoothed with a Gaussian kernel of 6 mm full width at half-maximum (FWHM).

On the first, individual subject level the feedback onsets were modeled with the reward prediction errors (RPE) included as parametric modulator. The six realignment parameters, the derivative of the translation parameters and a dummy regressor for scans with excessive motion were added as nuisance regressors. The stress and control condition were modeled separately.

Contrast images were computed for the RPE for the control and stress condition and subsequently submitted to random-effects group statistics (second level). A paired t-test was used to compare activation between conditions (stress/control). To control for multiple comparisons, family-wise error correction (p_{FWE}) was applied at the whole-brain level at $p_{FWE}<0.05$. For testing the condition effect, a mask of the RPE main effect over both conditions were used at $p_{FWE}<0.05$.

3. Results

3.1. Sample characteristics

The final sample consisted of n = 28 healthy male adult human participants with a mean age of 26.9 (*SD* = 5.7) years, a mean of 12.2 (*SD* = 1.2) educational years, and a mean verbal intelligence of 103.8 (*SD* = 10.1).

3.2. Stress response analyses

The stress intervention significantly increased subjective stress responses (arousal, valence and subjective stress), as well as physiological responses (cortisol levels). For detailed statistics refer to the Supplement and Figure 3.

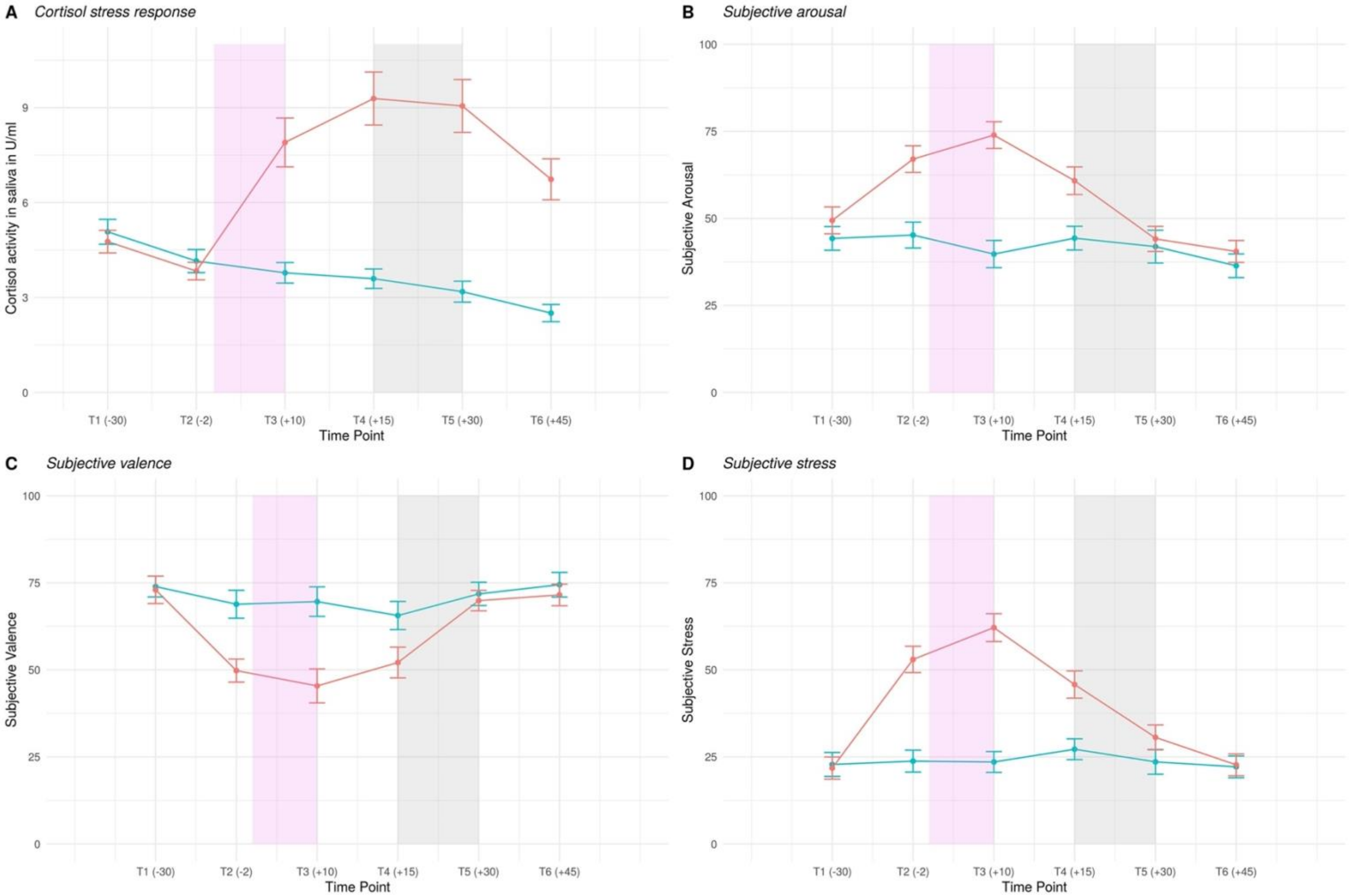


Figure 3 Physiological (cortisol) (A) and subjective stress response (B-D) over the course of the session. Violet shaded area: period of intervention (either stress induction (TSST) or control intervention), grey shaded area: reversal learning task was administered in the MR scanner.

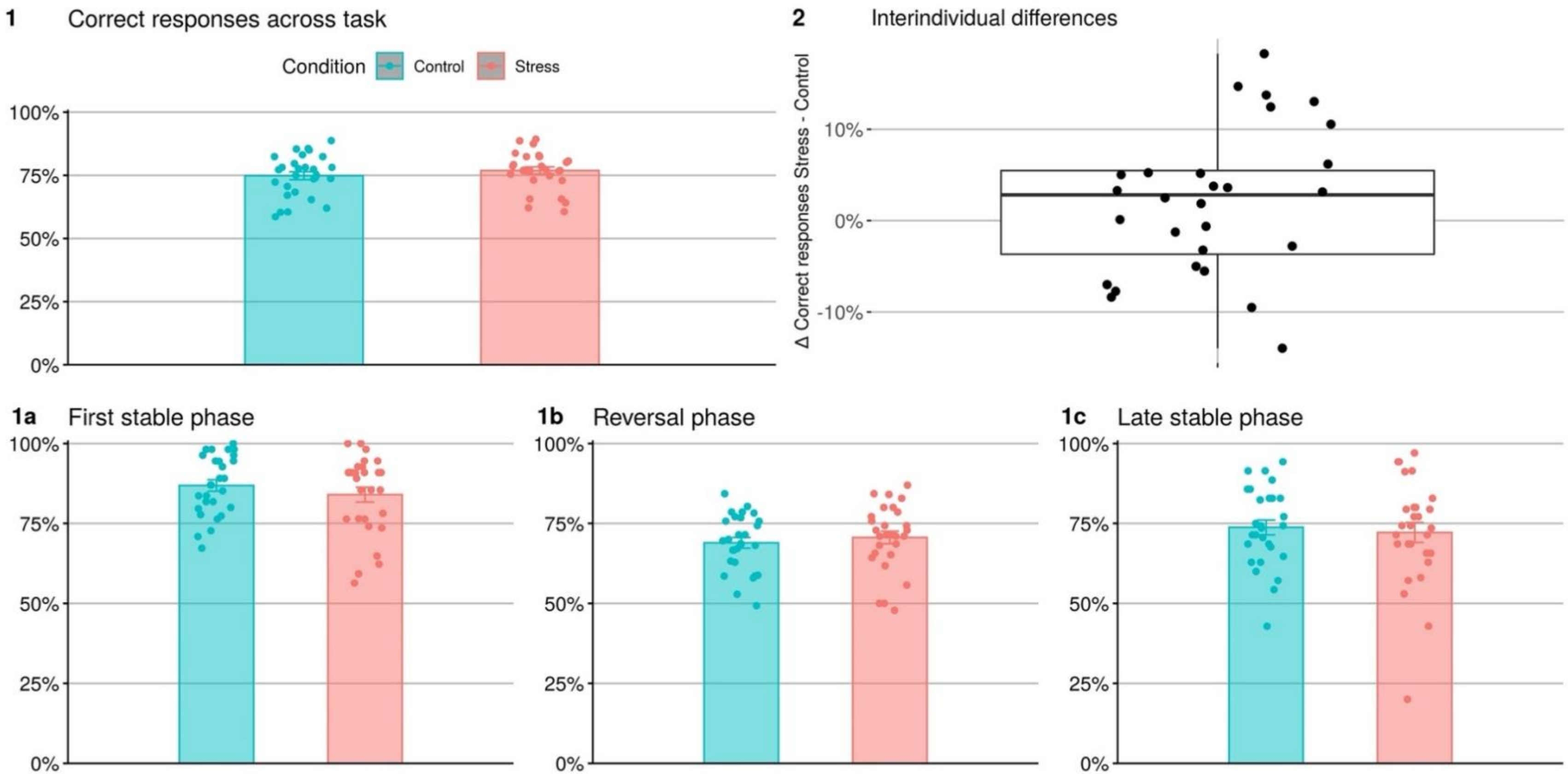


Figure 4 Correct responses during stress (green) and control condition (red) across task (1), as well as phases (1a-c) and interindividual differences between conditions (2).

3.3. Behavioral results

Best-fitting multilevel linear modeling included a subject-specific intercept, as well as main effects of condition and phase. Predicting correct responses on a single-trial basis with multilevel linear modeling indicated the expected task effect in the reversal ($p < 0.001$) and in the last stable phase ($p < 0.001$). For both phases, correct responses decreased with respect to the first reference phase. Furthermore, there was a main effect of condition ($p = 0.020$), suggesting that participants' correct responses subtly increased with a 1.13 higher chance (OR = 1.13) for correct responses under stress (see Table 1 and Supplement: Figure S1). As shown in Figure 4.2, the effects of stress on correct responses were quite heterogenous with high interindividual variability. The findings on correct responses were supported by a significant main effect ($p = 0.030$) of stress when the physiological stress level (AUC) was used as a continuous predictor instead of experimental condition (see Supplementary Figure S2 and Table S-A). In this model, task effects were again significant for the reversal phase ($p < 0.001$) as well as the last stable phase ($p < 0.001$). Regarding win-stay behavior, best-fitting multilevel linear modeling included a subject-specific intercept, as well as a main effect of condition and phase. Task effects of the reversal phase ($p < 0.001$) and the last stable phase ($p < 0.001$) were significant, but not the experimental condition ($p = 0.22$). Similarly, lose-switch behavior resulted in significant task effects of reversal phase ($p < 0.001$) and last stable phase ($p < 0.001$), but not experimental condition ($p = 0.73$) (see Supplementary Tables S-B and S-C).

Exploratory behavioral analysis of moderator variables:

Improved and impaired learners under stress did not differ in their working memory capacity ($t(26) = -0.84$, $p = 0.4$) nor in their chronic stress exposure ($t(22) = 0.79$, $p = 0.4$).

Table 1 Multilevel generalized linear modeling results of the winning model predicting correct responses

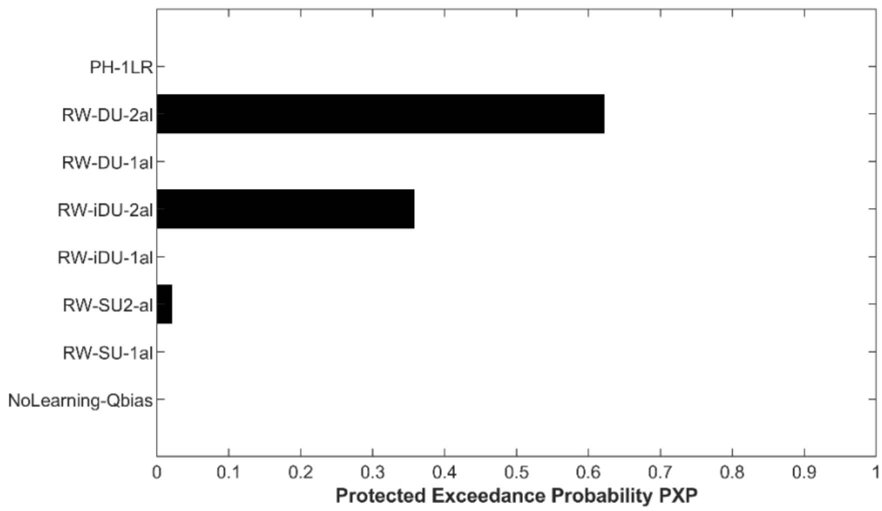
	Correct Responses				
Predictors	*Estimate (SE)*	*CI*	*Z*	*p*	*OR*
Intercept	1.23 (0.07)	1.08-1.38	17.13	< 0.001	
Condition	0.12 (0.05)	0.01-0.22	2.32	0.020	1.13
Reversal Phase	0.96 (0.06)	0.83-1.07	15.27	< 0.001	2.6
Last Stable Phase	0.8 (0.07)	0.65-0.94	10.89	< 0.001	2.22
ICC	0.04				
$N_{subject}$	28				
Observations	8893				
Marginal R^2 / Conditional R^2	0.053/0.088				

3.4. Computational modeling results

Behavior in the control condition ('step 1 model space') was best explained by a Rescorla-Wagner model with full double update and two learning rates (the RW-DU-2al) across all participants with a PXP = 0.62 (see Figure 6). This indicates that participants used the anticorrelated task structure and updated the chosen and the unchosen choice option to a similar extent (full double update model, DU). Furthermore, the learning rate in win trials was lower than in loss trials (paired t-test on alpha win vs alpha loss: $t(27) = -6.7$, $p < 0.001$), resulting in stronger influence of RPEs in loss compared to win trials. In a next step, additional free parameters for potential stress effects were entered for this winning model (the 'step 2 model space'). This resulted in a best fit for RW-DU-2al-StressBetas (PXP = 0.92), indicating that only the temperature parameters β_{win} and β_{loss} were different between control and stress condition but not the learning rates (see Table 2 for parameter estimates). Model comparison resulted in lower protected exceedance probabilities (PXP < 0.1) for all other models (see Figure 6). Choice temperature parameters were significantly higher after win trials compared to loss trials ($F(1,27)= 22.77$, $p < .001$) and numerically higher during the control compared to the stress condition, although the latter effect was not significant ($F(1, 27) = 0.25$, $p = .623$). We observed a large interindividual variance for the temperature parameters (see Supplementary Figures S3 and S4 for violin plots of parameter distributions).

Table 2 Parameter mean estimates of the winning model of ‘step 2 model space’.

Variable	*M*	*SD*
α_{win}	0.19	0.11
α_{loss}	0.36	0.17
$\beta_{\mathrm{control\ win}}$	6.01	3.99
$\beta_{\mathrm{control\ loss}}$	3.21	2.52
$\beta_{\mathrm{stress-win}}$	5.61	4.68
$\beta_{\mathrm{stress-loss}}$	3.08	3.33



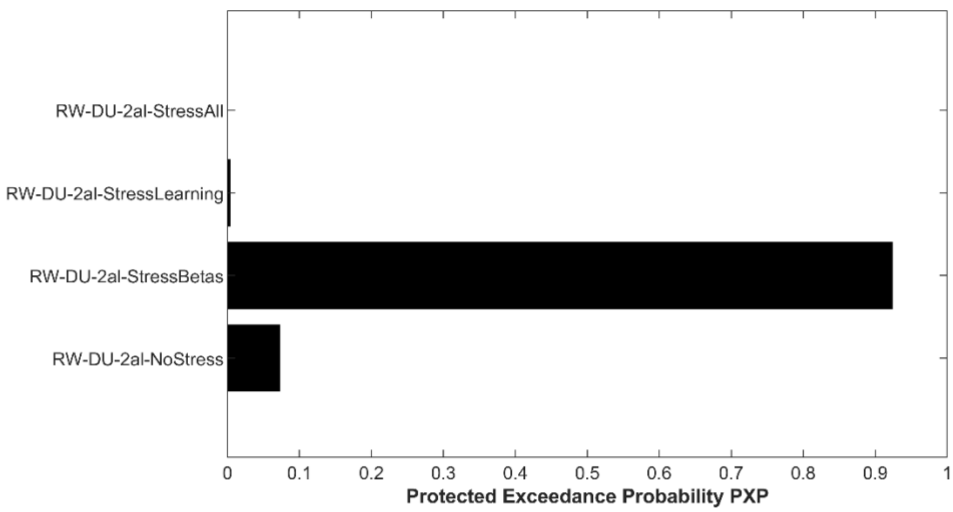


Figure 5 Protected exceedance probability: (a) 'step 1' model space explaining behavior in the control condition (top), (b) 'step 2' model space with added free stress parameters to the best fitting model of the control condition in order to detect stress-related parameter differences between control and stress condition (bottom).

fMRI results

We found a main effect of RPE combined over both conditions in the vmPFC, bilateral ventral striatum, posterior cingulate cortex (PCC) and bilateral insula ($p_{FWE} < .05$ for the whole brain, see Figure 6 and Supplementary Table S-D). We did not observe significant RPE-related activation differences between control and stress condition. On a trend level, there was higher activation in the right insula during stress compared to the control condition ([46 4 10], $t = 4.02$, $p_{FWE\ SVC\ main\ effect} = .068$, $p_{uncorrected} < 0.001$; see Supp. Figure S7).

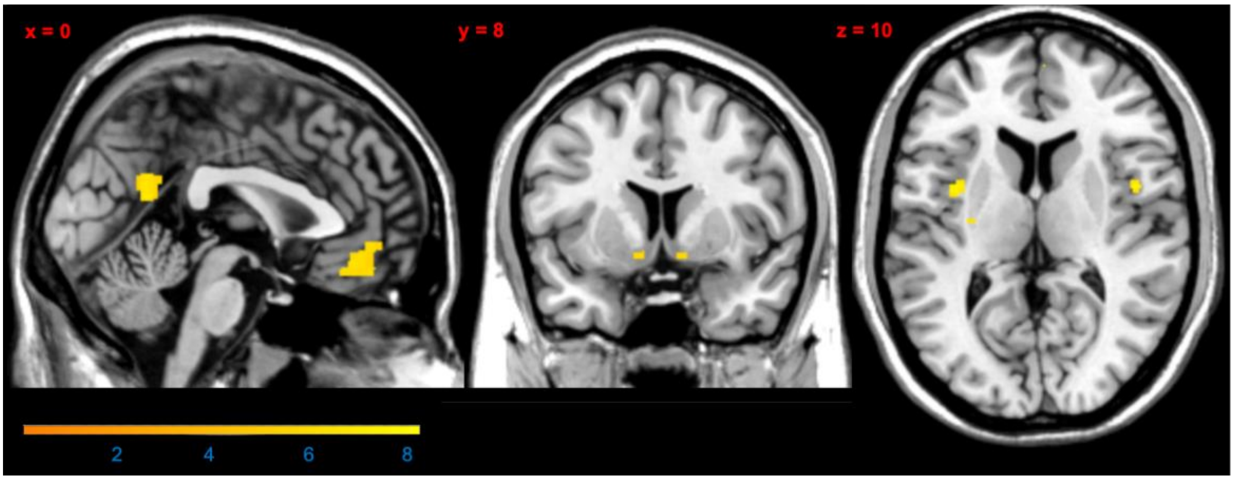


Figure 6: Neural activation related to reward prediction error across both conditions. Displayed are clusters showing significant RPE coding in vmPFC, ventral striatum and insula at $p_{\text{FWE whole brain corrected}} < 0.05$ combining stress and control conditions (main effect of task).

Discussion

The present study investigated the behavioral and neural effects of acute psychosocial stress on reversal learning in healthy male human participants. While participants made slightly more correct responses under acute stress, the neural representation of RPE signals were not significantly altered by acute stress in our sample. Computational modelling of choice behavior showed no stress effect on learning rates, but high stress-related interindividual variability in the use of learned values.

On the behavioral level, participants learned to choose the correct (i.e. more often rewarded) stimulus and adapted their choices after changes in reward contingencies (reversals) during both the control as well as the stress condition. Unlike previous studies (Shields et al., 2016), we observed more correct responses during the stress compared to the control condition in our analysis, but the effect size was small (OR = 1.13) and other behavioral measures such as win-stay or lose-switch behavior were not affected. Furthermore, participants displayed substantial interindividual variability including better, worse, or non-different performance under acute stress in our within-subjects design.

Computational modelling of choice behavior showed that participant's behavior was best explained by a RL model where prediction errors update the expected values of both the chosen and the unchosen choice option, indicating that participants considered the anticorrelated task structure. Acute stress did not affect the learning rate, which scales the influence of the RPE in updating of the expected values. Therefore, within our model space there was no evidence that stress affected the updating speed of learned expected values itself. However, our modeling analysis suggests that the degree to which participants used the learned values (temperature parameter) differed between the stress annd control condition. Interestingly, there was no overall condition effect, but model comparison showed that introducing dissimilar temperature parameters for the control and the stress condition individually explained the observed behavior best. The absence of a significant condition effect on the temperature parameter together with the model selection result indicates meaningful interindividual variability of choice behavior in response to acute stress. Other studies using cognitive computational modeling during learning tasks also observed effects of acute stress on choice temperature, mostly with higher stochasticity (Radenbach 2015, Cremer 2020), while other studies observed attenuation of model-based behavior (Otto et al., 2013) or an increased tendency for win-stay behavior (Raio et al., 2020). However, comparability is limited due to the different tasks used, mainly focusing on the balance between model-free and model-based learning (Cremer et al., 2021; Otto et al., 2013; Raio et al., 2020), which was not the focus of the present study.

On the neural level, RPE signals were coded in a network comprising vmPFC, bilateral ventral striatum, posterior cingulate cortex and insula across both conditions, in line with previous studies using the same paradigm (Boehme et al., 2015; Katthagen et al., 2020; Reiter et al., 2017, 2016a) and with meta-analytic findings of RPE fMRI studies (Fouragnan et al., 2018). The fact that we did not find whole-brain correctable effects of stress on RPE representation is in line with our behavioral findings that learning parameters were not affected by acute stress and that the behavioral stress effects uncovered by multilevel linear modeling (increased correct responses) were very subtle. The trendwise increase of RPE-related activation in the insula during the stress compared to the control condition, might contribute to the behavioral effect as the insula has been implicated in error processing, mainly interpreted to code salience signals (Fouragnan et al., 2018). However, this finding did not survive correction for multiple testing and therefore needs to be interpreted with caution. On the other hand, our neural findings might suggest that (model-free) reward prediction error processing is not affected by acute stress and that behavioral stress effects may be more related to the value representation and utilizing of those values during the decision process as indicated by our modeling findings. Although speculative at this point, our finding of altered choice stochasticity parameters may hint towards this and aligns with recent findings on the importance of computational noise directly affecting value representation (Findling et al., 2019). In rodents acute stress improved reversal learning whereas chronic stress impaired reversal learning (Bryce and Howland, 2015; Hurtubise and Howland, 2017). Differential long-term stress exposure may have led to the heterogenous effects of stress on reversal learning in our sample. In humans, chronic stress increased the detrimental influence of acute stress on model-based learning (Radenbach et al., 2015b). Apart from chronic stress exposure, cognitive capacities or personality traits are further potential explanations for the inconsistent impact of acute stress on learning. A high working memory capacity seems to hold a protective function against the attenuation of model-based learning (Otto et al., 2013), while trait impulsivity interacts with different aspects of learning differentially, but particularly seems to increase perseveration (Raio et al., 2017). As probabilistic reversal learning does not disentangle model-based and model-free learning these effects of moderators were difficult to replicate here. Exploratory analyses on working memory capacity and chronic stress exposure did not reveal any respective effects on stress in our sample.

Our findings are limited by some of the following factors. Considering the gender differences in decision-making (Shields et al., 2016) which may be amplified by stress (Mather and Lighthall, 2012) and potential impact of cyclical changes in female individuals we decided to investigate an exclusively male sample. Furthermore, our sample was homogenously young and highly educated. Therefore, our findings cannot be generalized to the general population or patient samples. Our task does not allow to temporally disentangle value and RPE representations in the brain. Dissociating these computations might be a promising avenue for future studies to determine the neurocomputational processes underlying reversal learning performance increases under acute stress.

While our relatively young and healthy study sample has shown only slight beneficial effects of acute stress, other more vulnerable populations may show different patterns. Stress, especially when long-term or chronic, is an important factor in causing and maintaining psychiatric illness (McEwen, 2004). While healthy individuals can adapt to a certain level of stress and even find it beneficial (Lighthall et al., 2013), decision-making frequently goes awry in psychiatric disorders (Cáceda et al., 2014; Voon et al., 2017). Our results suggest that it might be worthwhile assessing decision-making under acute stress in populations at risk of developing psychiatric conditions to reveal how stress is involved in maladaptive decision-making. Identification of altered choice behavior and relevant neural networks in healthy individuals make it possible to disentangle how stress affects healthy decision-making and what might be a maladaptive psychiatric alteration. As an operationalization of cognitive flexibility, reversal learning is a construct with high relevance for several psychiatric disorders. For instance, cognitive flexibility and its neural correlates are impaired in patients with alcohol use disorder (Reiter et al., 2016b), anorexia nervosa (Bernardoni et al., 2017), binge-eating disorder (Reiter et al., 2017), ADHD (Hauser et al., 2021) or schizophrenia (Schlagenhauf et al., 2014).

Conclusion

Our study combines the advantages of a within-subjects design and fine-grained computational measures to investigate the effect of acute psychosocial stress on healthy male adults. Several lines of analysis showed slightly improved performance, reflected in altered choice stochasticity, but without whole-brain-correctable neural effects of stress in the coding or RPEs.

Funding

This work was supported by the Rubicon award granted to ZS by the Netherlands Organization for Scientific Research (NWO, Rubicon 2014/05563/ALW), by the Einstein Center for Neurosciences (LW), and by the Heisenberg program from the German Research Foundation (grant number SCHL 1969/5-1), as well as the Max Planck Society and grants from the German Research Foundation awarded to FS (DFG SCHL 1969/2-2 as part of FOR1617 and SCHL1969/1-2 and 402170461 as part of TRR265).

Acknowledgements

We thank all the authors contacted who provided additional information. The authors would like to thank all volunteers who participated in this study. The authors would also like to thank C. Edwards, K. Hudl, L. Golz, M. Jochemko, A. Kummer, N. Pampus, J. Lepsien, A. Pampel, T. Schlumm, S. Thiers, C. Backhauß, M. Falk, and L. Fisher for their valuable assistance in recruitment and data acquisition throughout the study.

Disclosure Statement

The authors have declared that there are no conflicts of interest in relation to the subject of this study.

Data availability

Data and analysis scripts are available via https://github.com/agschlagenhauf/SALAD

Author Contributions

ZS, FS; Conceptualization.
LL, ZS; Data curation.
LW, CE, TK, FS; Formal analysis.
LW; Writing - original draft.
CE, TK, ZS, FS, LL, MP, AH; Writing - review & editing.

References

Bates, D., Mächler, M., Bolker, B.M., Walker, S.C., 2015. Fitting linear mixed-effects models using lme4. J. Stat. Softw. 67. https://doi.org/10.18637/jss.v067.i01

Bernardoni, F., Geisler, D., King, J.A., Ritschel, F., Murr, J., Reiter, A.M.F., Smolka, M.N., Kiebel, S., Ehrlich, S., 2017. Altered medial frontal feedback learning signals in anorexia nervosa. Biol. Psychiatry. https://doi.org/10.1016/j.biopsych.2017.07.024

Boehme, R., Deserno, L., Gleich, T., Katthagen, T., Pankow, A., Behr, J., Buchert, R., Roiser, J.P., Heinz, A., Schlagenhauf, F., 2015. Aberrant salience is related to reduced reinforcement learning signals and elevated dopamine synthesis capacity in healthy adults. J. Neurosci. 35, 10103–10111. https://doi.org/10.1523/JNEUROSCI.0805-15.2015

Brainard, D.H., 1997. The Psychophysics Toolbox. Spat. Vis. 10, 433–436.

Bryce, C.A., Howland, J.G., 2015. Stress facilitates late reversal learning using a touchscreen-based visual discrimination procedure in male Long Evans rats. Behav. Brain Res. 278, 21–28. https://doi.org/10.1016/j.bbr.2014.09.027

Cáceda, R., Nemeroff, C.B., Harvey, P.D., 2014. Toward an understanding of decision making in severe mental illness. J. Neuropsychiatry Clin. Neurosci. 26, 196–213. https://doi.org/10.1176/appi.neuropsych.12110268

Cavanagh, J.F., Frank, M.J., Allen, J.J.B., 2011. Social stress reactivity alters reward and punishment learning. Soc. Cogn. Affect. Neurosci. 6, 311–320. https://doi.org/10.1093/scan/nsq041

Cohen, S., Gianaros, P.J., Manuck, S.B., 2016. A Stage Model of Stress and Disease. Perspect. Psychol. Sci. 11, 456–463. https://doi.org/10.1177/1745691616646305

Cremer, A., Kalbe, F., Gläscher, J., Schwabe, L., 2021. NeuroImage Stress reduces both model-based and model-free neural computations during flexible learning. Neuroimage

229, 117747. https://doi.org/10.1016/j.neuroimage.2021.117747

Doherty, J.P.O., Dayan, P., Friston, K., Critchley, H., Dolan, R.J., 2003. Temporal Difference Models and Reward-Related Learning in the Human Brain 28, 329–337.

Dolan, R.J., Dayan, P., 2013. Review Goals and Habits in the Brain. Neuron 80, 312–325. https://doi.org/10.1016/j.neuron.2013.09.007

Findling, C., Skvortsova, V., Dromnelle, R., Palminteri, S., Wyart, V., 2019. Computational noise in reward-guided learning drives behavioral variability in volatile environments. Nat. Neurosci. 22, 2066–2077. https://doi.org/10.1038/s41593-019-0518-9

Fouragnan, E., Retzler, C., Philiastides, M.G., 2018. Separate neural representations of prediction error valence and surprise: Evidence from an fMRI meta-analysis. Hum. Brain Mapp. 39, 2887–2906. https://doi.org/10.1002/hbm.24047

Goldfarb, E. V, Froböse, M.I., Cools, R., Phelps, E.A., 2015. Stress and Cognitive Flexibility : Cortisol Increases Are Associated with Enhanced Updating but Impaired Switching 14–24. https://doi.org/10.1162/jocn

Hauser, T.U., Iannaccone, R., Ball, J., Mathys, C., Brandeis, D., Walitza, S., Brem, S., 2021. Role of the Medial Prefrontal Cortex in Impaired Decision Making in Juvenile Attention-Deficit/Hyperactivity Disorder 71, 1165–1173. https://doi.org/10.1001/jamapsychiatry.2014.1093

Hurtubise, J.L., Howland, J.G., 2017. Effects of stress on behavioral flexibility in rodents. Neuroscience 345, 176–192. https://doi.org/10.1016/j.neuroscience.2016.04.007

Katthagen, T., Kaminski, J., Heinz, A., Buchert, R., Schlagenhauf, F., 2020. Striatal Dopamine and Reward Prediction Error Signaling in Unmedicated Schizophrenia Patients. Schizophr. Bull. 46, 1535–1546. https://doi.org/10.1093/schbul/sbaa055

Kirschbaum, C., Pirke, K.-M., Hellhammer, D.H., 1993. The ‘Trier Social Stress Test’ – A Tool for Investigating Psychobiological Stress Responses in a Laboratory Setting.

Neuropsychobiology 28, 76–81. https://doi.org/10.1159/000119004

Kudielka, B.M., Schommer, N.C., Hellhammer, D.H., Kirschbaum, C., 2004. Acute HPA axis responses, heart rate, and mood changes to psychosocial stress (TSST) in humans at different times of day. Psychoneuroendocrinology 29, 983–992. https://doi.org/https://doi.org/10.1016/j.psyneuen.2003.08.009

Lighthall, N.R., Gorlick, M.A., Schoeke, A., Frank, M.J., Mather, M., 2013. Stress Modulates Reinforcement Learning in Younger and Older Adults 28, 35–46. https://doi.org/10.1037/a0029823

Luettgau, L., Schlagenhauf, F., Sjoerds, Z., 2018. Psychoneuroendocrinology Acute and past subjective stress in fl uence working memory and related neural substrates. Psychoneuroendocrinology 96, 25–34. https://doi.org/10.1016/j.psyneuen.2018.05.036

Lupien, S.J., Maheu, F., Tu, M., Fiocco, A., Schramek, T.E., 2007. The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition. Brain Cogn. 65, 209–237. https://doi.org/10.1016/j.bandc.2007.02.007

Mather, M., Lighthall, N.R., 2012. Risk and reward are processed differently in decisions made under stress. Curr. Dir. Psychol. Sci. 21, 36–41. https://doi.org/10.1177/0963721411429452

McEwen, B.S., 2004. Protection and damage from acute and chronic stress: Allostasis and allostatic overload and relevance to the pathophysiology of psychiatric disorders. Ann. N. Y. Acad. Sci. 1032, 1–7. https://doi.org/10.1196/annals.1314.001

Otto, A.R., Raio, C.M., Chiang, A., Phelps, E.A., Daw, N.D., 2013. Working-memory capacity protects model-based learning from stress. Proc. Natl. Acad. Sci. U. S. A. 110, 20941–20946. https://doi.org/10.1073/pnas.1312011110

Pearce, J.M., Hall, G., 1980. A model for Pavlovian learning: Variations in the effectiveness of conditioned but not of unconditioned stimuli. Psychol. Rev.

https://doi.org/10.1037/0033-295X.87.6.532

Piray, P., Dezfouli, A., Heskes, T., Frank, M.J., Daw, N.D., 2019. Hierarchical Bayesian inference for concurrent model fitting and comparison for group studies. PLoS Comput. Biol. 15. https://doi.org/10.1371/journal.pcbi.1007043

Plessow, F., Fischer, R., Kirschbaum, C., Goschke, T., 2011. Inflexibly Focused under Stress: Acute Psychosocial Stress Increases Shielding of Action Goals at the Expense of Reduced Cognitive Flexibility with Increasing Time Lag to the Stressor. J. Cogn. Neurosci. 23, 3218–3227. https://doi.org/10.1162/jocn_a_00024

Plessow, F., Kiesel, A., Kirschbaum, C., 2012. The stressed prefrontal cortex and goal-directed behaviour: Acute psychosocial stress impairs the flexible implementation of task goals. Exp. Brain Res. 216, 397–408. https://doi.org/10.1007/s00221-011-2943-1

Porcelli, A.J., Delgado, M.R., 2017. Stress and Decision Making: Effects on Valuation, Learning, and Risk-taking. Curr. Opin. Behav. Sci. 14, 33–39. https://doi.org/10.1016/j.cobeha.2016.11.015

Rabe, M.M., Vasishth, S., Hohenstein, S., Kliegl, R., Schad, D.J., 2020. hypr: An R package for hypothesis-driven contrast coding. J. Open Source Softw. 5, 2134. https://doi.org/10.21105/joss.02134

Radenbach, C., Reiter, A.M.F., Engert, V., Sjoerds, Z., Villringer, A., Heinze, H.J., Deserno, L., Schlagenhauf, F., 2015a. The interaction of acute and chronic stress impairs model-based behavioral control. Psychoneuroendocrinology 53, 268–280. https://doi.org/10.1016/j.psyneuen.2014.12.017

Radenbach, C., Reiter, A.M.F., Engert, V., Sjoerds, Z., Villringer, A., Heinze, H.J., Deserno, L., Schlagenhauf, F., 2015b. The interaction of acute and chronic stress impairs model-based behavioral control. Psychoneuroendocrinology 53, 268–280. https://doi.org/10.1016/j.psyneuen.2014.12.017

Raio, C.M., Hartley, C.A., Orederu, T.A., Li, J., Phelps, E.A., 2017. Stress attenuates the flexible updating of aversive value. Proc. Natl. Acad. Sci. U. S. A. 114, 11241–11246. https://doi.org/10.1073/pnas.1702565114

Raio, C.M., Konova, A.B., Otto, A.R., 2020. Trait impulsivity and acute stress interact to influence choice and decision speed during multi-stage decision-making. Sci. Rep. 10. https://doi.org/10.1038/s41598-020-64540-0

Reiter, A.M.F., Deserno, L., Kallert, T., Heinze, H.-J., Heinz, A., Schlagenhauf, F., 2016a. Behavioral and Neural Signatures of Reduced Updating of Alternative Options in Alcohol-Dependent Patients during Flexible Decision-Making. J. Neurosci. 36, 10935 LP – 10948. https://doi.org/10.1523/JNEUROSCI.4322-15.2016

Reiter, A.M.F., Deserno, L., Kallert, T., Heinze, H.J., Heinz, A., Schlagenhauf, F., 2016b. Behavioral and neural signatures of reduced updating of alternative options in alcohol-dependent patients during flexible decision-making. J. Neurosci. 36, 10935–10948. https://doi.org/10.1523/JNEUROSCI.4322-15.2016

Reiter, A.M.F., Heinze, H.J., Schlagenhauf, F., Deserno, L., 2017. Impaired Flexible Reward-Based Decision-Making in Binge Eating Disorder: Evidence from Computational Modeling and Functional Neuroimaging. Neuropsychopharmacology 42, 628–637. https://doi.org/10.1038/npp.2016.95

Schlagenhauf, F., Huys, Q.J., Deserno, L., Rapp, M.A., Beck, A., Heinze, H.J., Dolan, R., Heinz, A., 2014. Striatal dysfunction during reversal learning in unmedicated schizophrenia patients. Neuroimage 89, 171–180. https://doi.org/10.1016/j.neuroimage.2013.11.034

Schwabe, L., Dalm, S., Schächinger, H., Oitzl, M.S., 2008. Neurobiology of Learning and Memory Chronic stress modulates the use of spatial and stimulus-response learning strategies in mice and man 90, 495–503. https://doi.org/10.1016/j.nlm.2008.07.015

Schwabe, L., Tegenthoff, M., Höffken, O., Wolf, O.T., 2013. Mineralocorticoid Receptor

Blockade Prevents Stress-Induced Modulation of Multiple Memory Systems in the Human Brain. Biol. Psychiatry 74, 801–808. https://doi.org/10.1016/j.biopsych.2013.06.001

Schwabe, L., Wolf, O.T., 2011. Stress-induced modulation of instrumental behavior: From goal-directed to habitual control of action. Behav. Brain Res. 219, 321–328. https://doi.org/10.1016/j.bbr.2010.12.038

Schwabe, L., Wolf, O.T., 2009. Stress prompts habit behavior in humans. J. Neurosci. 29, 7191–7198. https://doi.org/10.1523/JNEUROSCI.0979-09.2009

Shields, Grant S., Sazma, M.A., Yonelinas, A.P., 2016. The effects of acute stress on core executive functions: A meta-analysis and comparison with cortisol. Neurosci. Biobehav. Rev. 68, 651–668. https://doi.org/10.1016/j.neubiorev.2016.06.038

Shields, Grant S, Trainor, B.C., Lam, J.C.W., Yonelinas, A.P., 2016. Acute stress impairs cognitive flexibility in men, not women. Stress 19, 542–546. https://doi.org/10.1080/10253890.2016.1192603

Starcke, K., Brand, M., 2016. Effects of stress on decisions under uncertainty: A meta-analysis. Psychol. Bull. 142, 909–933. https://doi.org/10.1037/bul0000060

Tewes, U., Wechsler, D., 1991. Hamburg-Wechsler-Intelligenztest f�r Erwachsene : HAWIE-R. Huber, Bern; Stuttgart.

Voon, V., Reiter, A., Sebold, M., Groman, S., 2017. Model-Based Control in Dimensional Psychiatry. Biol. Psychiatry 82, 391–400. https://doi.org/10.1016/j.biopsych.2017.04.006