**Acute stress alters probabilistic reversal learning in healthy participants**Lara Wieland1, Claudia Ebrahimi1, Teresa Katthagen1, Martin Panitz2, Lennart Luettgau3, Andreas Heinz1, Florian Schlagenhauf1\*, Zsuzsika Sjoerds4\* (\*shared last author ship)

Affiliations:

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

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

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

4Cognitive 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: sz.sjoerds@fsw.leidenuniv.nl

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

Text no. of words:

no. of references: 38

**Abstract**

**Introduction** Stressful situations can improve or impair flexibility of reward learning. Reversal learning requires flexible adaptation to sudden changes in reward contingencies. Although acute stress effects on reversal learning are highly relevant in the context of psychiatric disorders,it has rarely been investigated. Here, we employed functional MRI informed by computational modeling parameters in a within-subject design with healthy human volunteers to investigate the effect of acute psychosocial stress on flexible behavioral adaptation.

**Methods** Participants (n=28) underwent fMRI during a reversal learning task, once after the Trier Social Stress Test (TSST) (ST), a validated psychosocial stress induction method, and once after a control condition (CT) in two separate sessions. During the task, participants chose between two stimuli with anti-correlated and probabilistic reward contingencies in order to obtain rewards in three phases with different levels of volatility on reward contingencies. Effects of stress on choice behavior were investigated using generalized linear mixed-effects models and a set of computational models describing different learning processes that might have generated the data (hybrid Pearce-Hall and Rescorla-Wagner). Computational models were fitted using a hierarchical Bayesian approach (Piray, 2019), and model-derived reward prediction errors (RPE) were used as regressors for fMRI analyses.

**Results** Cortisol responses and measures of subjective stress demonstrated that the TSST induced a state of acute stress. Stress significantly albeit subtly increased correct responses. Model comparison revealed that a Rescorla-Wagner model with individual scaling of the inverse decision temperature best explained the observed behavior under stress. On the neural level, RPEs signals were coded in striatum and vmPFC. No whole-brain correctable effects of stress on RPE representations were found.

**Discussion** Our study shows that acute social stress has an impact on reversal learning with high interindividual variability.

**1. Background** (1028 words)

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 volatile, less predictable factors. Both, stable and volatile 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. This might especially be relevant in situations that afford high behavioral flexibility, for instance in constantly changing environments.

**1.1. Psychiatric relevance**

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). Stress is an important factor in causing and maintaining psychiatric conditions (McEwen, 2004) and health-related behavior in general (Cohen et al., 2016). Understanding the impact of stress on decision-making is essential to develop a mechanistic understanding of psychiatric disorders and necessitates a model of how stress affects choice behavior in healthy individuals. So far small sample sizes, heterogenous subdomains in the operationalization of decision-making and methodological considerations with regard to the type of stressor have complicated the picture (Porcelli and Delgado, 2017).

**1.2 Stress and learning**

In operationalizing decision-making under stress, it is crucial to evaluate which behavior constitutes optimal decision-making in different paradigms. Flexible decision-making requires several subdomains such as goal implementation, risk-taking, reward and loss sensitivity, task switching or learning from feedback. 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 had 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 investigating the effects of acute stress on cognitive flexibility in a smaller study concluded that stress had an impairing effect (Shields et al., 2016). Both meta-analyses predominantly focused on 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. Further variance in these measurements is introduced by 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.

**1.3 Stress and learning (computational)**

In the last decade computational approaches combined with cognitive neuroscience to study cognitive control and decision-making gained traction in clinical application (Huys et al., 2016; Maia and Frank, 2011). A useful dichotomy in differentiating decision-making is habitual versus goal-directed behavior and its respective neural correlates (Balleine and O’Doherty, 2010). Both types of learning find their computational equivalent in reinforcement learning (RL) models. Generally, learning results from so-called reward prediction errors (RPE), signaling the difference between an observed and expected reward. RPEs are used to update choice values of choice options in order to maximize rewards. When behaving habitually or model-free, a learner is guided by model-free RPE in seeking rewards without taking the environmental structure into account. The respective neural signature is found in the ventral striatum. Goal-directed behavior on the other hand is also called model-based because a model of the environment is necessary. A model-based RPE incorporates knowledge about higher-order task structure, such as interdependent choice options. On the neural level, the medial prefrontal cortex (mPFC) is a key network for model-based learning. In general, learners presumably use a mixture of both types for learning depending on their cognitive resources (Dolan and Dayan, 2013). As stress is a key factor impacting cognitive resources, both processes should be prone to interruption by stress (Schwabe and Wolf, 2011, 2009). Experimental attempts to disentangle model-free and model-based learning show that stress affects both types of learning in different ways (Cremer et al., 2021; Otto et al., 2013).

**1.4 Probabilistic reversal learning**

Probabilistic reversal learning is a central concept in the realm of reward-related decision-making because it operationalizes cognitive flexibility in a volatile environment. Successful performance in this task is subsumed under model-based learning. Paradigms tapping at this type of learning require participants to choose between two stimuli with anti-correlated reward contingencies. Contingencies are reversed several times throughout the task unannounced and therefore require flexible adaptation to a volatile environment. Optimal learning necessitates that the anti-correlated task structure is understood and taken into consideration. In this case probabilistic losses should still lead to stay-behavior and probabilistic wins should lead to switch-behavior. Meta-analytically, neural correlates of reversal learning could be expected to be represented in two major neural networks: whereas the RPE is typically encoded in human reward circuitry, such as the ventral striatum and the ventromedial prefrontal cortex (vmPFC), surprise is usually represented in the anterior cingulate cortex, anterior insula and dorsal striatum (Fouragnan et al., 2018).

**1.5 Limitations of previous studies**

Since 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 – it is not clear how stress affects the neural correlates of probabilistic reversal learning. The few studies using within-subject designs (Radenbach et al., 2015) to investigate learning, are either purely behavioral or employ electroencephalography lacking the possibility of precise spatial signal localization (Cavanagh et al., 2011) 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. Applying a state-of-the-art hierarchical Bayesian modeling approach (Piray et al., 2019), allowed us to model the impact of stress. To our best of knowledge probabilistic reversal learning has not been investigated after a psychosocial stress intervention in a within-subject design before.

**2. Methods** (1642 words)

2.1. Study Design:

Employing a within-subject design, 36 healthy adult participants (n=28 in the final analyzed sample) performed a probabilistic reversal learning task during fMRI in two separate sessions seven days apart. 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 without presence of the committee. Session type (stress vs. control) was counter-balanced across participants. In order to prevent confounding effects of circadian rhythm on cortisol levels both experimental sessions were scheduled at exactly the same time of the day (Kudielka et al., 2004). Acute stress response was assessed at physiological (cortisol and α-amylase) and subjective (self-report ) level at six timepoints throughout the session (see Figure 3).

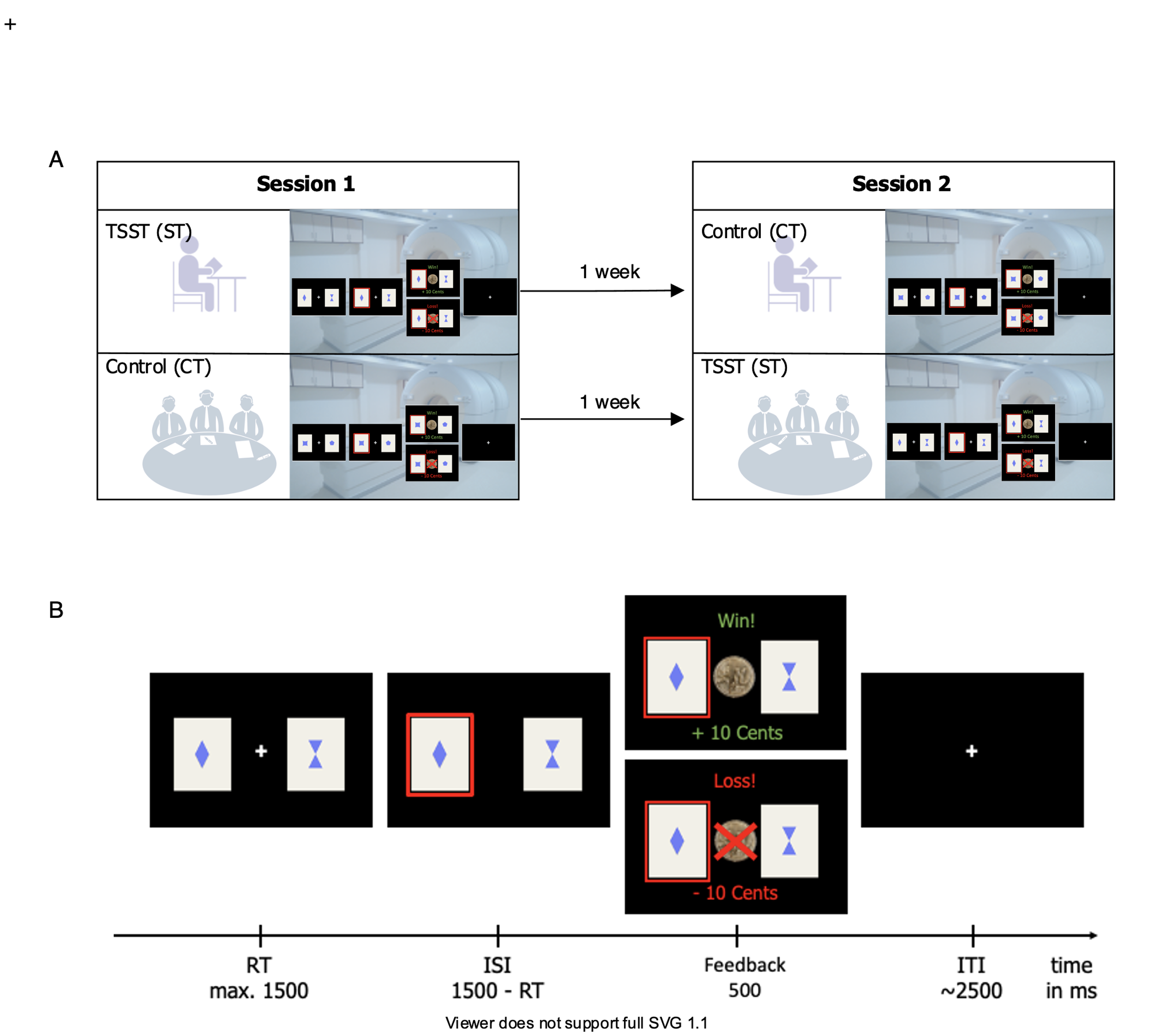


Figure 1 Study design (A) and task design (B)

2.2. Physiological stress response:

We assessed physiological stress response via salivary cortisol, which were assessed 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. 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 with respect to ground (AUCg-stress and AUCg-control, according to Pruessner et al., 2003) separately for both conditions and subtracting AUCg-control from AUCg-stress. The AUC was calculated based on individualized subject wise time points, taking into account slight dispersion in the testing protocol. We further applied a criterion of 1.5nmol per liter (nmol/l, Miller et al., 2013) from post-stress (T3 through T6) above the pre-stress level (T1/2) during the stress condition (Goodman et al., 2017) to define stress-responders vs. non-responders. Using this method three participants were classified as non-responders. However, excluding these participants did not change our behavioral findings and therefore they remained in the final sample.

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). Arousal was rated on the scale "Do you feel active or sleepy?" from 0 (sleepy) to 100 (active), valence on the scale "Do you feel happy or unhappy?" from 0 (unhappy) to 100 (happy) and stress on the scale "Do you feel stressed or not stressed?" from 0 (not stressed) to 100 (stressed).

2.4. Task Design

Participants performed a probabilistic reversal learning task, which included 160 trials and comprised around 15 minutes. The task was programmed in Matlab (The MathWorks, Natick, MA) with Psychtoolbox (Boehme, Reiter). On every trial, participants had to decide between two cards, depicting a different geometric stimulus. The underlying reward structure was not explicitly instructed but could be inferred: reward probabilities associated with the two choice options were anticorrelated (i.e. whenever card A was rewarded, card B was punished and vice versa). Furthermore, participants were informed on the probabilistic nature of the task: the respective winning card was only rewarded in 80% of all trials. Right-side versus left-side location of the stimulus was randomized on each trial. After a fixed number of trials, contingencies reversed and these reversals repeated over the middle experimental phase (see Figure 2). Participants were instructed to win as much money as possible and received a monetary bonus at the end of the experiment. Participants in both groups were matched on differences in proportion of the number of informative and misleading events between the two sessions because feedback was drawn probabilistically. Due to a technical error in the task script, 8 participants had to be excluded from the final sample, leaving a total of 28 participants for final analyses.

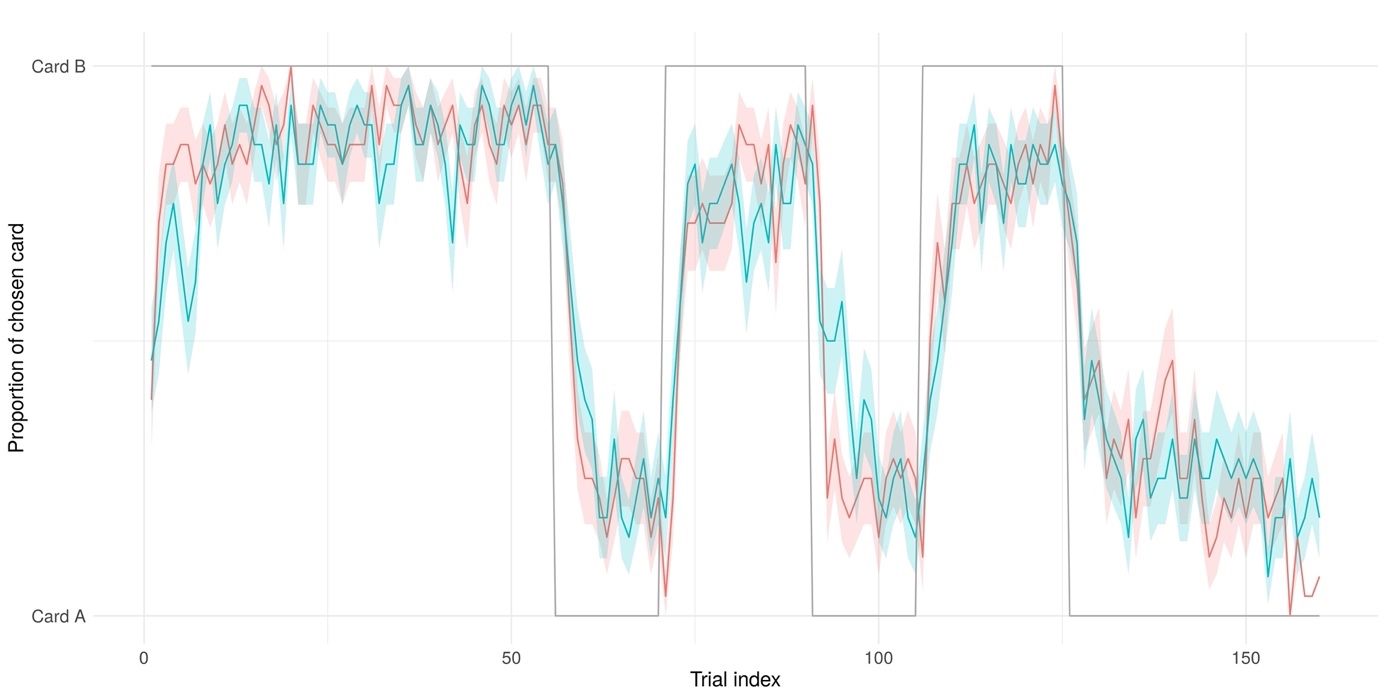


Figure 2 Empirical choice behavior in both conditions with underlying task structure in grey

2.5. Analyses

*2.5.1. Stress response analyses*

Cortisol responses (AUC-g) and the three subjective VAS scales were compared across conditions (stress vs control) using paired-samples *t*-tests.

*2.5.2. Behavioral data*

Single-trial multilevel linear model (logistic regressions) were conducted using the lme4 package (Bates et al., 2015) in R (Version 3.1.X). Parameter estimates were considered significant at p≤.05. We analyzed trial-by-trial correct responses (chose better option), win-stay (select same stimulus after win) and lose-switch (switch stimulus after loss) behavior with factors *stress condition* (CT vs. ST, effect coding as -0.5 and 0.5) and *experimental phase* (pre, reversal, post) as fixed effects, 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 the Akaike information criterion (AIC) and Bayesian information criterion (BIC) to compare the models to a null model, which predicted outcome variables with the individually varying intercept per subject only. For the best-fitting model, the parameter estimates odd's ratio were 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%). This was the case for one participant. Across all trials, participants missed a relatively low number of trials (0.71%).

*2.5.3 Computational models*

In order to describe different learning processes that might have generated the data under stress and control condition 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. The model space comprised Rescorla-Wagner (RW), Pearce-Hall (PH; Diederen et al., 2016) models and a null model (no-learning). In the RW and PH models, the expected value of an action at trial is updated via the RPE (eq. 1), which is defined as the difference between received reward and previously expected reward value for the chosen stimulus (eq. 2):

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,for single update (SU), for full double update (DU) and freely fitted for individually weighted double update (iDU)). We further varied whether learning rates 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 Diederen et al., 2016). 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 following wins and losses:

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

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 and (RW-DU-2al-StressLerning) one model with stress only affecting the temperature parameters and , (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 (see Supplement). As in step 1, model fits were then compared between models.

*2.5.4. 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 follows a mixed effects 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.

*2.5.5. 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 mm2 in plane voxel resolution, 0.5 mm gap between slices, voxel size, 3 × 3 × 5 mm) . The scanning procedure further comprised a T1-weighted MPRAGE recorded within seven days before the first test session, and a field map to account for individual homogeneity differences of the magnetic field. Functional imaging data analyses were performed using SPM12 in Matlab.

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 further nuisance regressors. The stress and control condition were modeled separately.

Rewardwere 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 across conditions (stress/control) on the group level. To control for multiple comparisons, family-wise error correction (*pFWE*) was applied at the whole-brain level.

The association of RPE BOLD signal with behavioral learning performance was tested as follows: A flexible factorial design was used with condition (stress/control) as within-subject factor and learning (median split on percentage of correct responses resulting in two groups: improved vs. impaired learners under stress) and subject as random effect (see Supplement: Exploratory fMRI analyses).

**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), a mean of 12.2 (*SD* = 1.2) educational years, and a mean verbal intelligence of 103.8 (*SD* = 10.1).

3.2. Physiological and subjective results

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

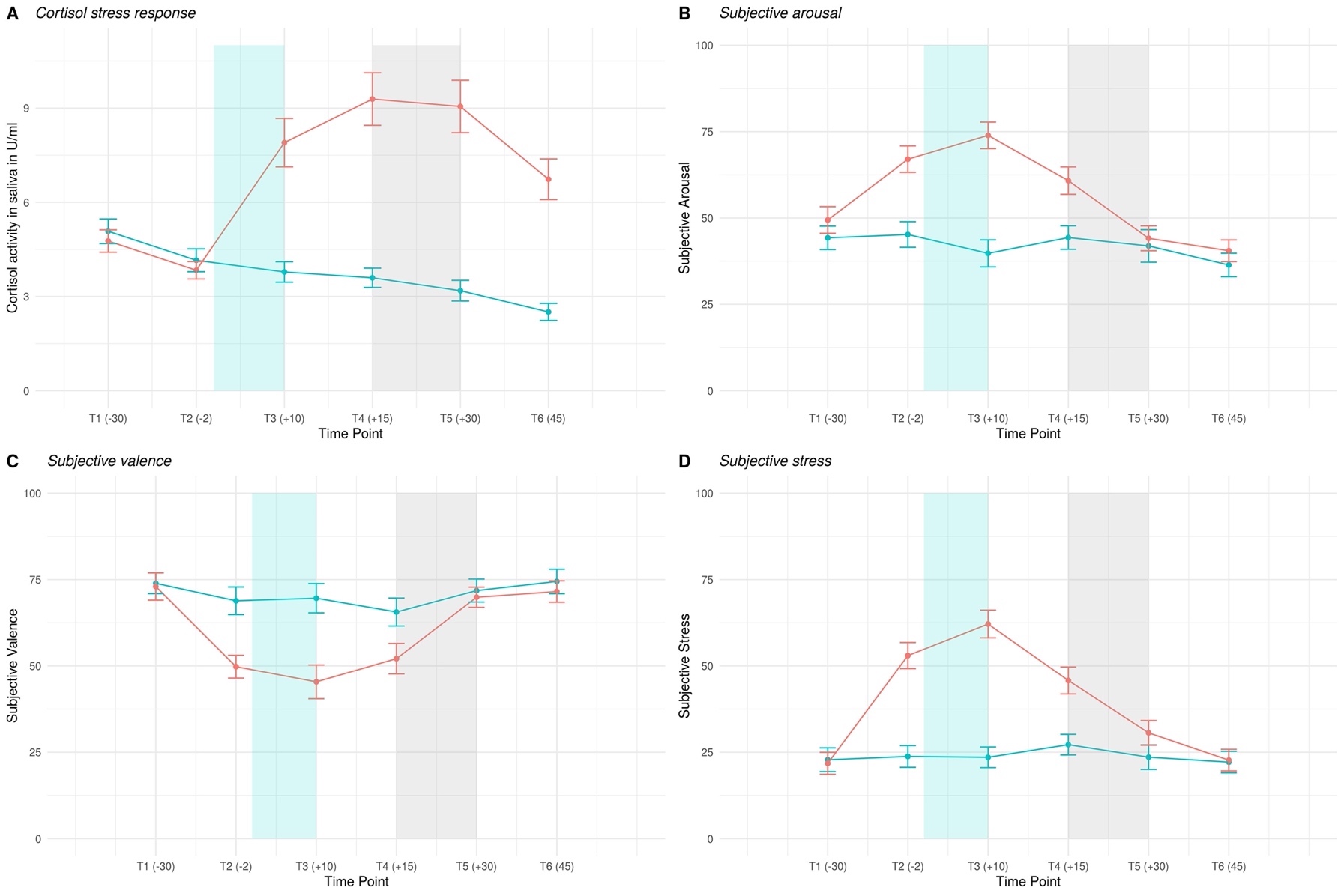


Figure 3 Physiological (cortisol) (A) and subjective stress response (B-D) over the course of the session. Blue shaded area: period of intervention (either stress induction (TSST) or control intervention), grey shaded area: reversal learning task was administered in the MR scanner.

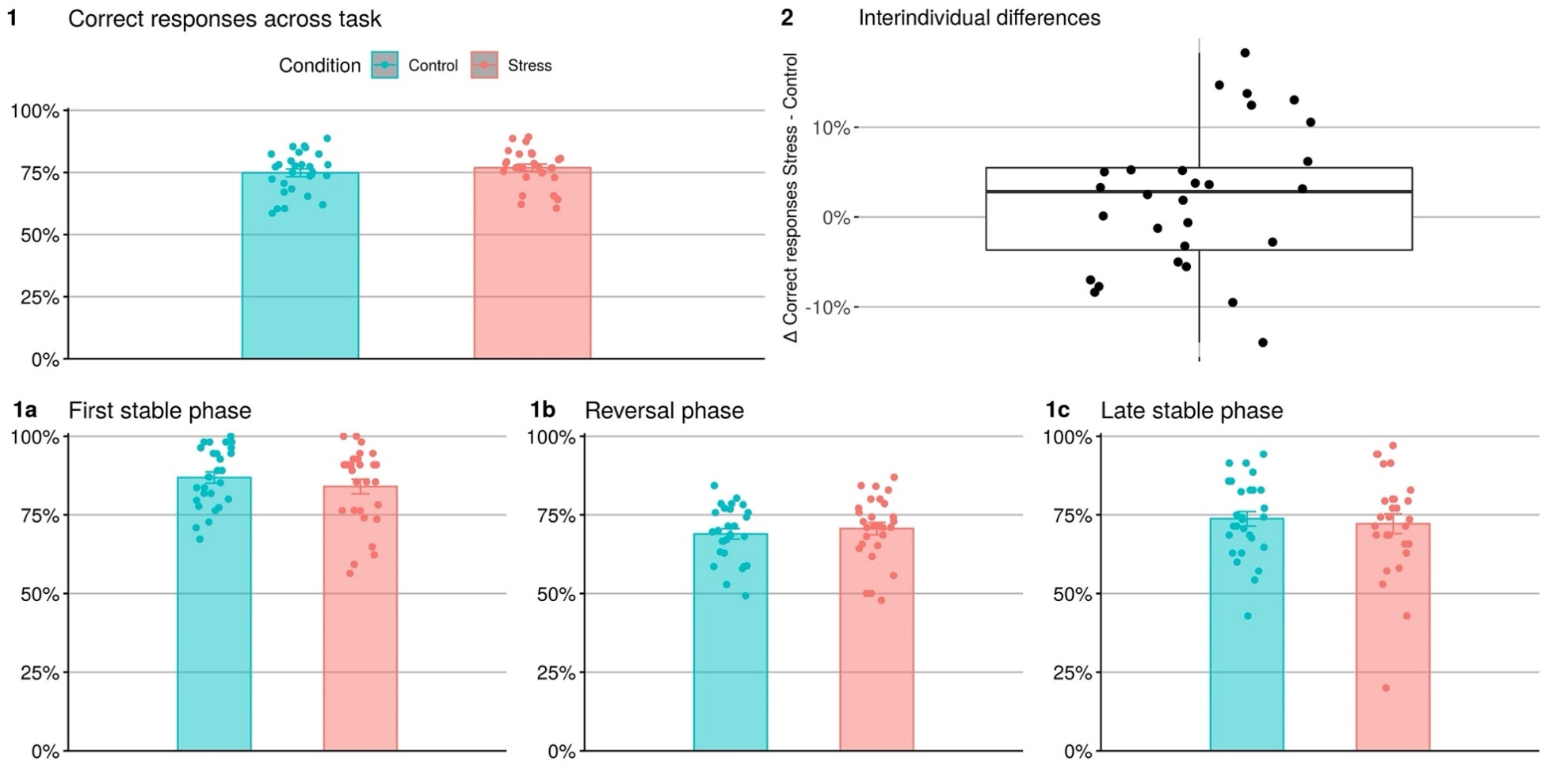


Figure 4 Correct responses across task (1), as well as phases (1a-c) and interindividual differences between conditions (2)

3.3. Behavioral results

Predicting correct responses on a single-trial basis with generalized mixed effects modeling indicated the expected task effect under reversal (*p* < 0.001) and in the second 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 for correct responses under stress (see Figure 4.1, Figure 5). These results were supported by a significant main effect (*p* = 0.030) of stress when physiological stress level (AUC) was used as a continuous predictor instead of experimental condition (see Figure S1). In this model, task effects were again significant for the reversal phase (*p* < 0.001) as well as the stable phase (*p* < 0.001).

Regarding win-stay behavior, task effects of reversal phase ﻿(*p* < 0.001) and stable phase ﻿(*p* < 0.001) were significant, but not experimental condition (*p* = 0.22). Similarly, lose-switch behavior resulted in significant task effects of reversal phase (*p* < 0.001) and stable phase ﻿(*p* < 0.001), but not experimental condition (*p* = 0.73).

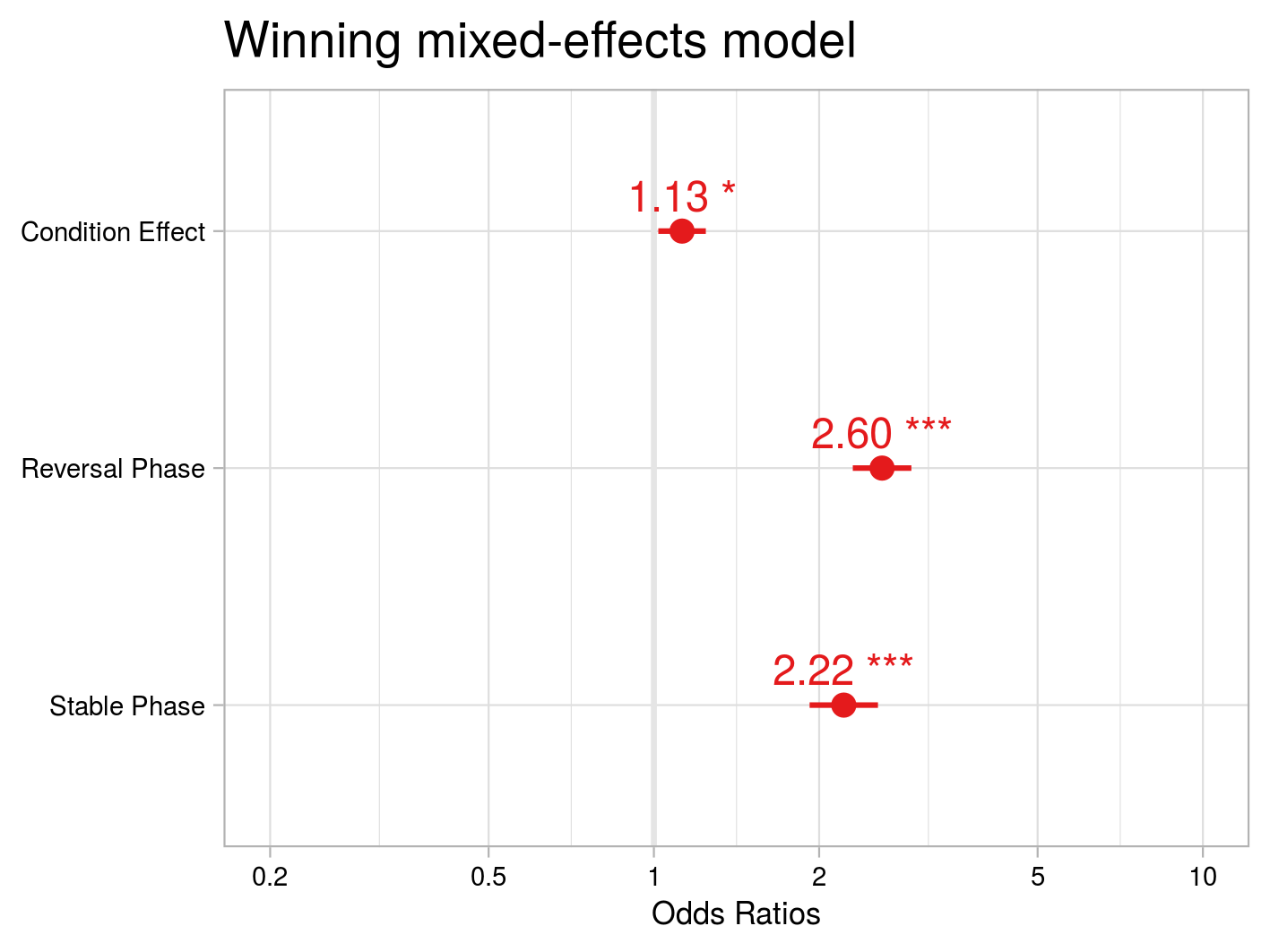


Figure 5 Odd's Ratio of condition (CI: 1.02-1.24), reversal phase (CI: 2.30-2.94) and stable phase (CI:1.92-2.56) contrasts from fixed-effects model.

Table 1 Mixed-effects modeling results

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Correct Responses** | | | | |  |
| **Predictors** | | *Estimate (SE)* | *CI* | *Z* | *p* |  |
| 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 |  |
| Reversal Phase | | 0.96 (0.06) | 0.83-1.07 | 15.27 | < 0.001 |  |
| Stable Phase | | 0.8 (0.07) | 0.65-0.94 | 10.89 | < 0.001 |  |
| ICC | | 0.04 |  |  |  |  |
| N subject | | 28 |  |  |  |  |
| Observations | | 8893 |  |  |  |  |
| Marginal R2 / Conditional R2 | | 0.053/0.088 |  |  |  |

3.4. Computational modeling results

Behavior in the control condition (‘step 1 model space’)was best explained by the RW-DU-2al model across all participants with a PXP = 0.62 (see Figure 6). Subsequently, adding free parameters for potential stress effects (the ‘step 2 model space’) resulted in a best fit for RW-DU-2al-StressBetas (PXP = 0.92), only affecting the temperature parameter and (see Table 2 for parameter estimates and Figure S2/3 for violin plots of distributions). Model selection resulted in protected exceedance probabilities (PXP) < 0.1 for all other models (see Figure 6).

Table 2 Parameter mean estimates of the winning model allowing for stress effect of decision noise (beta) with similar learning rates for control and stress condition.

|  |  |  |  |
| --- | --- | --- | --- |
| **Variable** | *M* | *SD* |  |
|  | 0.19 | 0.11 |  |
|  | 0.36 | 0.17 |  |
|  | 6.01 | 3.99 |  |
|  | 3.21 | 2.52 |  |
|  | 5.61 | 4.68 |  |
|  | 3.08 | 3.33 |  |



Figure 6 Protected exceedance probability - 'step 1' and 'step 2' model spaces

fMRI results

We found a main effect of RPE combined over both conditions in the vmPFC, bilateral striatum, posterior cingulate cortex (PCC) and bilateral insula (FWE-corrected p < 0.05 for the whole brain, see Table 1).

There was no significant effect of condition on RPE-related activation in the whie brain analyses.

Table 3 Main effects of task on RPE representation across conditions

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Main Effect of Task across Conditions (CT and ST)** | | | | | | | | |
| Region | Side | Cluster size | MNI - peak coordinate | | | *T value* | *puncorrected* | *pFWE* |
| corrected |
|  |  |  | x | y | z |  |  |  |
| Middle frontal gyrus  Precuneus | L/R  L | 401  222 |  |  |  |  |  |  |
| -10 | 42 | -12 | 9.48 | < .001 | < .001 |
| -8 | -52 | 32 | 9.05 | < .001 | < .001 |
| Ventral striatum  Insula | R | 16 | 10 | 2 | 12 | 8.62 | < .001 | < .001 |
| L | 56 | -36 | 2 | 12 | 8.08 | < .001 | .001 |
| Inferior frontal gyrus | L | 39 | -22 | -32 | -12 | 8.05 | < .001 | .001 |
| Precentral gyrus | R | 85 | 32 | -20 | 58 | 7.69 | < .001 | .002 |
| Superior frontal gyrus | L | 19 | -18 | 38 | 44 | 7.59 | < .001 | .003 |
| Postcentral gyrus | R | 75 | 38 | -26 | 46 | 7.56 | < .001 | .003 |
| Rolandic operculum | R | 11 | 46 | 2 | 10 | 7.35 | < .001 | .005 |
| Median cingulate gyri | R | 11 | 16 | -14 | 46 | 7.17 | < .001 | .007 |
| Paracentral lobule | L | 12 | -4 | -26 | 50 | 7.00 | < .001 | 0.010 |
| Postcentral gyrus | L | 21 | -34 | -30 | 48 | 6.81 | < .001 | 0.015 |
|  |  |  |  |  |  |  |  |  |

*Note*. Uncorrected as well as whole-brain corrected fMRI results (T-values, cluster sizes, side) with all clusters > 10 voxels from the main task effect across conditions in n = 28 participants are illustrated above. Abbreviations: fMRI = functional magnetic resonance imaging, FWE = family-wise error correction, L = left, MNI = Montreal Neurological Institute, R = right.

**Discussion** (552 words)

The present study investigated the behavioral and neural effects of acute psychosocial stress on reversal learning. In healthy male human participants, we found that probabilistic reversal learning improves slightly under acute psychosocial stress. Using intraindividual cortisol levels instead of condition (stress vs. control) as predictor confirmed these results. Win-stay, as well as lose-switch behavior was not affected. Choice behavior under stress was best explained by a model with altered choice stochasticity but unchanged learning rate under stress. Neurally, a network of vmPFC, bilateral ventral striatum, posterior cingulate cortex and insula represented reward-based learning across conditions with no whole-brain correctable effect of stress.

The existing literature on decision-making is unclear regarding the directional impact of stress on learning. While meta-analyses on older studies found that stress slightly impairs learning (Shields et al., 2016; Starcke and Brand, 2016), more recent studies using cognitive computational modeling found a shift in the balance of model-free and model-based learning (Cremer et al., 2021; Otto et al., 2013; Raio et al., 2020). High interindividual variability and a variety of paradigms and stressors used possibly obscured effects in previous studies and call for further investigation. Our neural findings are in line with previous studies using the same paradigm (Katthagen et al., 2020; Andrea M.F. Reiter et al., 2016; Reiter et al., 2017). Future studies should always consider interindividual differences in the reactivity to stress. The neural finding of no whole-brain correctable effect of stress may be related to the low effect size of the behavioral stress effects uncovered by multilevel linear modeling as well as computational modeling.

**Moderator variables**

A further potential explanation for the inconsistent impact of acute stress on learning are moderating variables, such as chronic stress exposure, cognitive capacities or personality traits. For instance, 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).

**Limitations**

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.

**Reversal learning and psychiatric relevance**

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, in patients with alcohol use disorder (Andrea M F Reiter et al., 2016), anorexia nervosa (Bernardoni et al., 2017), binge-eating disorder (Reiter et al., 2017), ADHD (Hauser et al., 2021) or schizophrenia (Katthagen et al., 2020) cognitive flexibility and its neural correlates are impaired.

**Conclusion**

Our study combines the advantages of a within-subject design and fine-grained computational measures to investigate the effect of acute psychosocial stress on healthy male adults. Several lines of analysis showed altered choice stochasticity and slightly improved performance under stress.

**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 1969/1-2/3-1/5-1) (FS),

**Acknowledgements**

We thank all the authors contacted who provided additional information.

**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; Roles/Writing - original draft.  
CE, TK, ZS, FS, LL, MP, AH; Writing - review & editing.

**References**

Balleine, B.W., O’Doherty, J.P., 2010. Human and rodent homologies in action control: Corticostriatal determinants of goal-directed and habitual action. Neuropsychopharmacology 35, 48–69. https://doi.org/10.1038/npp.2009.131

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

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

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

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

Goodman, W.K., Janson, J., Wolf, J.M., 2017. Meta-analytical assessment of the effects of protocol variations on cortisol responses to the Trier Social Stress Test. Psychoneuroendocrinology 80, 26–35. https://doi.org/10.1016/j.psyneuen.2017.02.030

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

Huys, Q.J.M., Maia, T. V., Frank, M.J., 2016. Computational psychiatry as a bridge from neuroscience to clinical applications. Nat. Neurosci. 19, 404–413. https://doi.org/10.1038/nn.4238

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

Maia, T. V., Frank, M.J., 2011. From reinforcement learning models to psychiatric and neurological disorders. Nat. Neurosci. 14, 154–162. https://doi.org/10.1038/nn.2723

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

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., 2015. 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, Andrea M.F., Deserno, L., Kallert, T., Heinze, H.J., Heinz, A., Schlagenhauf, F., 2016. 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, Andrea M F, Deserno, L., Kallert, T., Heinze, H.J., Heinz, A., Schlagenhauf, F., 2016. 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

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

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