**Working Title: Effects of acute stress on probabilistic reversal learning in healthy participants**

* Lara Wieland1, Claudia Ebrahimi1, Teresa Katthagen1, Martin Panitz2, Lennart Luettgau, Andreas Heinz1, Zsuzsika Sjoerds3\*, Florian Schlagenhauf1\* (\*shared last author ship)

Abstract: Stressful situations can alter reward-based learning. In some studies stress leads to increased learning from rewards, in others it does not show an effect (Starcke & Brand, 2016). A task addressing reward-based learning is the reversal learning task, which uses probabilistic rewards as feedback and incorporates sudden changes in reward contingencies. Deficits in reversal learning have been described in various psychiatric patient populations such as psychosis or addiction, known to be susceptible to stress. The effects of acute stress on reversal learning have been rarely investigated. Here, we investigated the effect of acute social stress in a within-subject design in healthy control participants.

A sample of *n* = 28 male non-clinical participants performed the task in a control condition versus the Trier Social Stress Test (TSST), a validated method to induce psychosocial stress. In our version of the reversal learning task (Reiter, 2016) participants choose between two anti-correlated stimuli in order to obtain rewards in three blocks. Reward contingencies remain stable for the first 55 trials and the last 35 trials. During the second block, in between the stable blocks, reward contingencies change four times, which requires participants to flexibly adapt their behavior. Performance was measured in correct responses, frequency of switches after losses and wins.

Cortisol responses and subjective stress responses showed that the stress induction was successful. Preliminary analyses showed no significant effect of stress induction on the number of correct responses or the frequency of switches after losses and wins. These results demonstrate that reversal learning, at least regarding the overall performance measures in our task, is robust to stress-related changes. Behavioral modeling of the task could yield further insights into more subtle behavioral changes after stress induction. Further research should test more diverse samples and patient populations as well.

Study Design:

The study entailed a within-subjects design, in which participants performed the reversal learning task in two separate test sessions seven days apart.

Physiological stress response:

Saliva was collected using Salivette saliva sampling tubes (SalivetteCortisol®, Sarstedt, Nuembrecht, Germany) to extract and measure salivary cortisol (see Supplement). Physiologically relevant cortisol increases were defined at a threshold of 1.5nmol per liter (nmol/l, Miller et al., 2013) increase post-stress (t3 through t6) above the lowest pre-stress level (t1 or t2) during the stress condition as de- scribed previously (Radenbach et al., 2015; Schommer et al., 2003). 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 subjectwise time points, taking into account slight dispersion in the testing protocol,

Past subjective stress assessment:

Participants filled in trait questionnaires at home in an online survey program (Limesurvey, www.limesurvey.org). This survey included the German version of the Perceived Stress Scale (PSS-10, Cohen et al., 1983), a self-report questionnaire assessing past subjective stress during the last 30 days. Here, past subjective stress is defined in terms of the subjective evaluation of situations in one’s life as stressful. The PSS-10 is described in more detail in the Supplement.

Task Design

Participants performed a probabilistic reversal learning task, which included 160 trials and comprised around 12 minutes. The task was programmed in Matlab (The MathWorks, Natick, MA) with Psychtoolbox. 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 probabilites associated with the two choice options were anticorrelated (i.e. whenever card A was rewarded, card B was a loss 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

Analysis of choice behavior:

Behavioral performance was measured in the percentage of correct responses and was analyzed

**Task** The task included four blocks with 70 trials each and comprised around 30 mins in total. The feedback-dependent learning was implemented using a task adapted from Nassar (2012) and programmed in Matlab (The MathWorks, Natick, MA) with Psychtoolbox. Learning was embedded in the task of inferring the position of a moving helicopter in order to catch the bags it dropped. The helicopter was hidden behind clouds and dropped a bag from the sky on each trial. Its position remained stationary but varied noisily on every trial and occasionally underwent a larger change. There were two noise conditions with a low SD = 8 and a high SD = 16. The helicopter dropped money bags on half of all trials (140) and sand bags on the other half (140) according to a randomized distribution. Participants were told they could move a white cursor on a black line with a trackball mouse for 3 sec. They could confirm the final position via mouse click and after confirmation the white cursor turned black. They saw the bag drop (Bag Drop Phase, 0.5 secs) and were presented with their prediction from the last trial, the actual position of the helicopter and the difference (prediction error) visualized as a red line (Feedback Phase, 1 sec). If they caught the bag it flashed and it was shown until the end of the trial. If they did not catch the bag it disappeared after falling to the altitude of the black line. Additionally feedback of catching the bag (independent of bag type) was followed with a beep tone after the bag drop and not catching it was followed with a different less pleasant beep tone after the bag drop. Accordingly, each trial lasted approximately 4.5 secs in total. If participants were faster than 3 secs in placing the bucket their reaction time was subtracted from the Cue Phase and added onto the ITI in order to keep trial length equally long.

**Training Phase** Before the scanning session participants practiced a training version of the task on a laptop

**Payment** The maximum amount resulted of an approximate value of 0.05€ per trial. Since there were 140 reward trials participants were paid an amount that varied between 4-7€. The lower boundary was chosen to ensure participants would receive a financial compensation additional to scanner costs.

**fMRI Acquisition**

Scans were acquired on a

Scans were acquired on a Siemens 3 T Magnetom Triosystem with a 64-channel head coil (Siemens, Erlangen, Germany) with the following parameters: T2\*-weighted gradient-echo echo-planar, (repetition time, 800 ms; echo time, 37 ms; flip angle, 52°; field of view, X×X mm2 ; voxel size, 2 × 2 × 2 mm) . The scanning procedure (A) comprised a T1 MPRAGE at the start, (B) prior to each run, a field map was collected to account for individual homogeneity differences of the magnetic field, (C) an average total of X EPI volumes were recorded per subject. The task was programmed using Matlab 2016 [X] with Psychophysics Toolbox Version 3 [X]. It was presented on a Dell laptop screen (training phase) and on a projector via a mirror system in the scanner environment. Participants wore MR-compatible Siemens headphones. Responses were calibrated using a wheel on the Current Design MR-compatible Trackball Fiber Optic Response Pad and logged using the right index finger.