Nature Neuroscience
Brief Communication

SUPPLEMENTAL MATERIAL

Paradoxical reversal learning enhancement by stress or prefrontal cortical damage: rescue with BDNF

Carolyn Graybeal[1], Michael Feyder[1], Emily Schulman[1], Lisa M. Saksida[2], Timothy J. Bussey[2], Jonathan L. Brigman[1], and Andrew Holmes*[1]

[1] Laboratory of Behavioral and Genomic Neuroscience National Institute on Alcohol Abuse and Alcoholism, National Institutes of Health, Rockville, MD, USA
[2] Department of Experimental Psychology, University of Cambridge, Cambridge, Medical Research Council and Wellcome Trust Behavioral and Clinical Neuroscience Institute, UK

SUPPLEMENTAL METHODS

Experimental subjects

Male C57BL/6J mice were obtained from The Jackson Laboratory (Bar Harbor, ME, USA) at 7-8 weeks of age, and housed 2/cage on the same rack in a temperature (72±5°C) and humidity (45±15%) controlled vivarium under a 12-h light/dark cycle. There was at least 1 week acclimation week prior to experimental manipulation. All experimental procedures were approved by the National Institute on Alcohol Abuse and Alcoholism Animal Care and Use Committee and followed the National Institute of Health guidelines outlined in 'Using Animals in Intramural Research' and the local Animal Care and Use Committees.

Nature Neuroscience: doi:10.1038/nn.2954

Reversal task

Touchscreen-based operant apparatus

The apparatus and pre-training procedure was designed for mice, as described previously[1-3]. An operant chamber measuring 21.6 x 17.8 x 12.7 cm (model # ENV-307W; Med Associates) was housed within a sound and light attenuating box (Med Associates). The grid floor of the chamber was covered with solid Plexiglas to facilitate ambulation. A pellet dispenser delivering a 14-mg reward pellet (#F05684; BioServ) into a food magazine was located at one end of the chamber. A house light and tone generator was located on the same side. At the opposite end was a touch-sensitive screen (Light Industrial Metal Cased TFT LCD Monitor; Craft Data Limited). The touchscreen was covered by a black Plexiglas panel with 5 x 5 cm windows separated by 0.5 cm and located at a height of 6.5 cm from the chamber floor. Stimulus presentation was controlled by custom software ('MouseCat', L.M. Saksida). Nose-pokes were automatically detected by the touchscreen and recorded by the 'MouseCat' Software.

Pre-training

Body weights were slowly reduced and maintained to 85% free-feeding weight throughout testing. Prior to operant testing, mice were fed ~10 reward pellets/mouse in the home cage to prevent a neophagic response during testing. The following day, mice were acclimated to the operant chambers via placement in the chamber for 30 min with access to 10 pellets in the food magazine. Mice that consumed all pellets in a 30-min session proceeded to instrumental pre-training.

Pre-training involved 3 phases. In phase 1, randomly-shaped stimuli were presented in 1 of 2 touchscreen windows for 10 sec. When the stimulus disappeared, a 1-pellet reward was delivered; concurrent with presentation of a 2-sec 65-dB tone and illumination of the magazine light (secondary reinforcers) to support learning. Mice that consumed 30 pellets within 30 min

Nature Neuroscience: doi:10.1038/nn.2954

moved on to phase 2. In phase 2, a stimulus was presented in 1 of the 2 stimulus windows and the mouse was required to touch the window with the stimulus to receive a reward. Mice that earned 30 pellets within 30 min moved on to phase 3. In phase 3, the mouse was now required to initiate each trial (i.e., presentation of stimulus) with a head entry into the food magazine. In addition, to discourage indiscriminate responding at the windows, responses to the blank window resulted in a 15-sec timeout period in which the house lights were extinguished and trial initiation not possible. After each incorrect response, a correction trial was given in which the stimulus was presented in the same spatial configuration until a correct response was made. Mice that responded correctly ≥75% over a 30-trial session (i.e., 30 first presentation trials, excluding correction trials) proceeded to the discrimination and reversal task.

Pairwise discrimination and reversal

Mice were trained to discriminate between 2 novel stimuli presented in a spatially pseudorandomized manner over 30-trial sessions (15-sec ITI), as previously described[1, 3]. Designation of the correct and incorrect stimulus was counterbalanced across strain and stress treatment. Correct (correct trials) and incorrect (error trials) responses were rewarded and punished, respectively, as during pre-training. Error trials were followed by correction trials as during pre-training. The performance criterion was an average of ≥85% correct responses over 2 consecutive sessions. Once criterion was met, mice began reversal training. For reversal, designation of correct and incorrect stimulus for each mouse was opposite to that on discrimination. Testing otherwise was the same as for discrimination. Final reversal criterion was ≥85% correct responses over 2 consecutive sessions. Reversal learning was divided into two stages: early (perseverative) and late (learning)[4, 5]. To do this, performance on each session was determined by calculating the percent of correct responses made (correct trials/total possible rewarded trials) and binning sessions in which performance was <50% correct (=early stage) or ≥50% correct (=late stage). The number of errors and correction errors committed at each stage

were calculated, as was the average stimulus-response latency and reward-retrieval (post-stimulus-response) latency.

Excitotoxic lesions of vmPFC, OFC and DLS

Mice were trained to discrimination criterion and assigned to lesion or sham groups by matching trials to criterion. For surgery, mice were anesthetized with isoflurane and fixed in a stereotaxic apparatus (Kopf Stereotaxic Alignment System 1900). A 33-gauge infusion cannula (Plastics One, Roanoke, VA) attached with polyeurathane tubing to a Hamilton syringe (Hamilton, Reno NV) was directed at the vmPFC (+1.98 mm anteroposterior (AP), ±0.30mm mediolateral (ML), -3.50mm ventral (V) to Bregma), OFC (+2.70mm AP, ±1.35mm ML, -2.80mm V to Bregma) or DLS (+1.18 and +0.22 mm AP, ±2.35 mm ML, -0.25 and -0.30mm V to Bregma). 0.2 µL (vmPFC, OFC) or 0.25 (DLS) µL *N*-methyl-D-aspartate (Sigma-Aldrich, St. Louis, MO) or saline vehicle was infused over 2-2.5 min using a pump (Harvard Apparatus PHD 22/2000, Holliston, MA), with the cannula left in place for an additional 5 min to allow for full diffusion. On removal of the cannula, mice were sutured and returned to their home cages. Mice were given 1 week of recovery before being returned to food restriction for testing.

Sensitivity to outcome devaluation

Pre-devaluation training and stress. Sensitivity of reversal performance to changes in outcome value was assayed via malaise-induced reinforcer devaluation[6, 7]. Mice were trained to discrimination criterion and assigned to stress or control groups by matching trials to criterion. Stress was exactly as described above. Reversal training began the day after the final stress and mice were trained to chance levels of performance (control=50.4 ±1.4 % correct, stressed=49.4 ±1.5 % correct). Mice were then assigned to devalued or non-devalued groups matching for sessions to chance performance (which did not differ between controls (2.9 ±1.3) and stressed (3.4 ±2.1) by this stage of reversal). Devaluation occurred over the next 2 days.

Nature Neuroscience: doi:10.1038/nn.2954

Malaise-induced reward devaluation. The devalued group was given free access to 1.5 g of food reward in the home cage for 15 min and then immediately received an intraperitoneal (i.p.) injection (20 mL/kg body weight) of 0.15 M lithium chloride (LiCl) (Sigma-Aldrich, St. Louis, MO) to induce malaise. The non-devalued group received a LiCl injection and was given free access to 1.5g of the food reward in the home cage for 15 min ~4 hr later. Reward consumption was measured by weighing the available food reward at the start and end of the 15 min free feeding period. This procedure was repeated for both groups the following day. A reversal probe was conducted the following day.

Reversal probe. Sensitivity to outcome devaluation was probed in a single reversal session under non-rewarded extinction conditions. The procedure was the same as a normal reversal session with the exception that responses were not rewarded. Performance was quantified by errors, correction errors, % correct responding, total trials performed, and average response times to stimuli and magazine. The generalization and retention tests were both conducted the following day.

Generalization and retention tests. A context generalization test was performed to ensure that the association between food reward and malaise had successfully generalized from the home cage environment to the operant chamber. 1.5 g of food reward were placed in the magazine of the operant chamber and the amount eaten over 15 min was measured by weighing before and after this period. Finally, to ensure the long-term retention of the reward-malaise association, 1.5 g of reward was made freely available in the home cage and the amount eaten over 15 min was measured by weighing before and after this period.

BDNF vmPFC infusions

Mice were trained to discrimination criterion and assigned to infusion or vehicle groups by matching trials to criterion. Via stereotaxic surgery as above, with 26-gauge bilateral indwelling cannula (Plastics One, Roanoke, VA) were targeted at the vmPFC (+1.80 mm anteroposterior

Nature Neuroscience: doi:10.1038/nn.2954

(AP), ±0.40mm mediolateral (ML), -1.50mm ventral (V) to Bregma) and cemented in place. Mice were given 1 week of recovery before being returned to food restriction for testing, which began with reminder sessions to ensure discrimination retention. Mice were then exposed to 3 days of forced swim stress, or remained in their home cage, exactly as above. Immediately after the final stress exposure, a 33-gauge infusion cannula with a 1 mm projection was inserted into the guide cannula. Mice were infused with 0.4 mg/mL BDNF (Sigma-Aldrich, St. Louis, MO) in a total volume of 0.2 µL (in 0.9% saline) over 2 min with the cannula left in place for an additional 5 min to allow diffusion.

Forced swim stress

Stress was via forced swim exposure, as previously described[8]. In a room different from operant testing, mice were placed in a 20-cm diameter Plexiglas cylinder filled halfway with 24 ±1°C water for 10 min and then removed, dried and returned to the home cage. Mice were exposed to 1 swim stress trial per day, between 1400-1600 h. Non-stressed controls remained in the home cage.

Stress effects on reversal

3x swim stress prior to reversal

Mice were trained to discrimination criterion and assigned to stress or control groups by matching trials to criterion. The day after reaching discrimination criterion, mice were exposed to 3 consecutive days of swim stress or remained in the home cage. Twenty-four hours after the last day of stress mice commenced reversal training and the number of trials per session was increased from 30 to 60 (with a 2 hr maximum session length).

1x swim stress prior to reversal

The procedure was the same as for the *'3x swim stress prior to reversal'* experiment, with the exception that mice were only exposed to 1 swim stress 30 min prior to the first reversal session. Reversal sessions comprised 30 trials.

3x swim stress prior to reversal and daily throughout reversal

The procedure was the same as for the *'3x swim stress prior to reversal'* experiment, with the exception that, in addition to 3x daily swim stress exposure prior to reversal testing, mice received daily exposure to swim stress (at least 2 hr after testing) until reaching reversal criterion. Reversal sessions comprised 30 trials.

Stress effects on 'control' tasks

Stress effects on acquisition of discrimination

Mice were trained through phase 3 of pre-training as above and then exposed to 3 days of swim stress (or remained in the home cage). Twenty-four hours after the last day of stress mice commenced discrimination.

Acquisition of a simple instrumental response

Mice were trained to acquire a simple instrumental response, as previously described[9]. Mice went through pre-training phases 1 and 2 as above (but to a 90% criterion) and then exposed to 3 days of swim stress (or remained in the home cage). Acquisition training began the following day. For this task, mice were required to respond to 1 of 2 stimuli (1 x 2.8 cm^2 white square per window) to obtain reward. Stimuli remained on the screen until a response was made, with 30 trials (5-sec ITI) per session. Acquisition criterion was 30 trials within 12.5 min, on each of 5 consecutive sessions.

Acoustic startle and prepulse inhibition of startle

Mice were tested for acoustic startle and prepulse inhibition of startle as previously described[10]. Mice were exposed to 3 days of swim stress (or remained in the home cage). The following day, mice were placed in a clear Plexiglas cylinder in 1 of 4 SR-Lab System startle chambers (San Diego Instruments, San Diego, CA) for a 5 min acclimation period. A 65-dB broadband background noise was presented during acclimation and throughout the test session. During the test session, mice were presented with startle trials (40-msec 120-dB broadband sound pulse) and prepulse+startle trials (20-msec noise prepulse sound followed, 100 msec later, by a 40-msec 120-dB broadband sound pulse). The prepulse+startle trials were preceded and followed by 5 pulse alone trials, which were not included in the analyses. Test trials consisted of 10 trials of 3 different intensities (3, 6, and 12 dB above background). Each trial type was presented 10 times with a variable interval of 12-30 sec between each presentation.

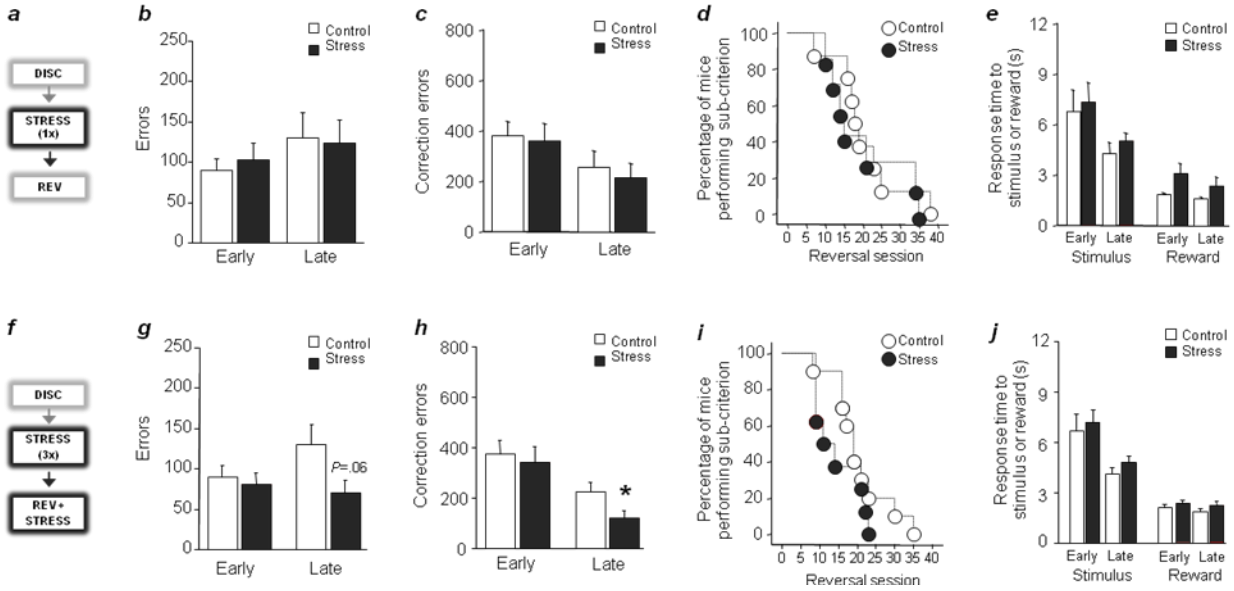
Basal activity in the startle chambers was measured during no-stimulus trials. Startle amplitude was measured every 1 msec, over a 65 msec period beginning at the onset of the startle stimulus. The peak startle amplitude over the sampling period was taken as the dependent variable. Whole-body startle responses were measured via vibrations transduced into analog signals by a piezoelectric unit attached to the platform on which the cylinders rested. Percent prepulse inhibition of startle (PPI) was calculated as 100 – [(startle response for prepulse+startle trials/startle response for startle-alone trials) × 100].

Statistical analysis

The effect of stress and lesion at each stage of learning was analyzed via Student's *t*-tests. Survival analysis of the percentage of mice that remained at sub-criterion across reversal sessions was conducted using the Mantel-Cox test. The effect of stress and prepulse intensity on percent prepulse inhibition was analyzed by ANOVA, with within-subjects analysis for intensity. The threshold for statistical significance was $P<.05$.

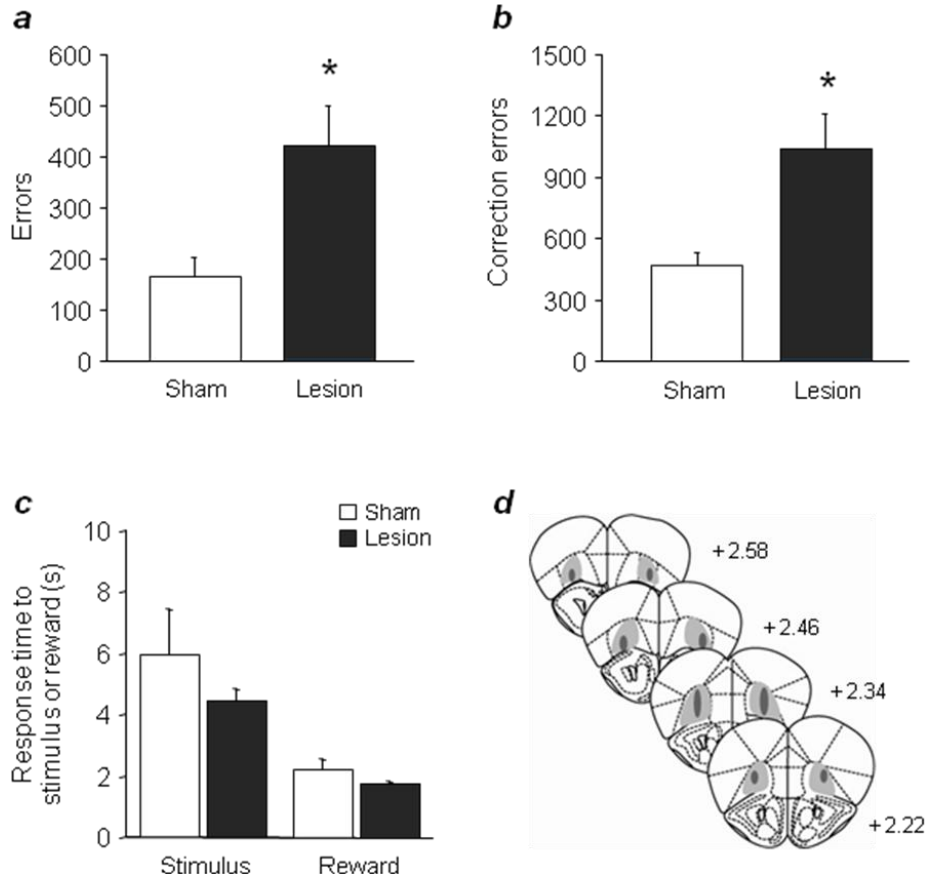
Nature Neuroscience: doi:10.1038/nn.2954

Supplemental Fig 1: Prolonged but not single stress facilitates late reversal. (***a***) A single session of swim stress was applied once 30 minutes prior to reversal. A single session of stress did not alter errors (***b***) or correction errors (***c***) (n=7-8/stress treatment). (***d***) The percentage of mice performing below criterion across sessions was not affected by a single stressor. (***e***) A single session of stress did not alter stimulus but not reward response times during late reversal. (***f***) Swim stress was applied once daily for 3 days prior to reversal and then daily at least 2 hours after each reversal session. This prolonged stress exposure produced a trend for decreased errors ($t(17)=2.05$, $P=.06$) (***g***) and significantly decreased correction errors ($t(17)=2.27$, $P<.05$) (***h***) (n=9-10/stress treatment). (***i***) The percentage of mice performing below criterion across sessions was not affected by prolonged stress. (***j***) Stress did not alter stimulus but not reward response times during late reversal. Data are Means ±SEM. *$P<.05$ versus controls

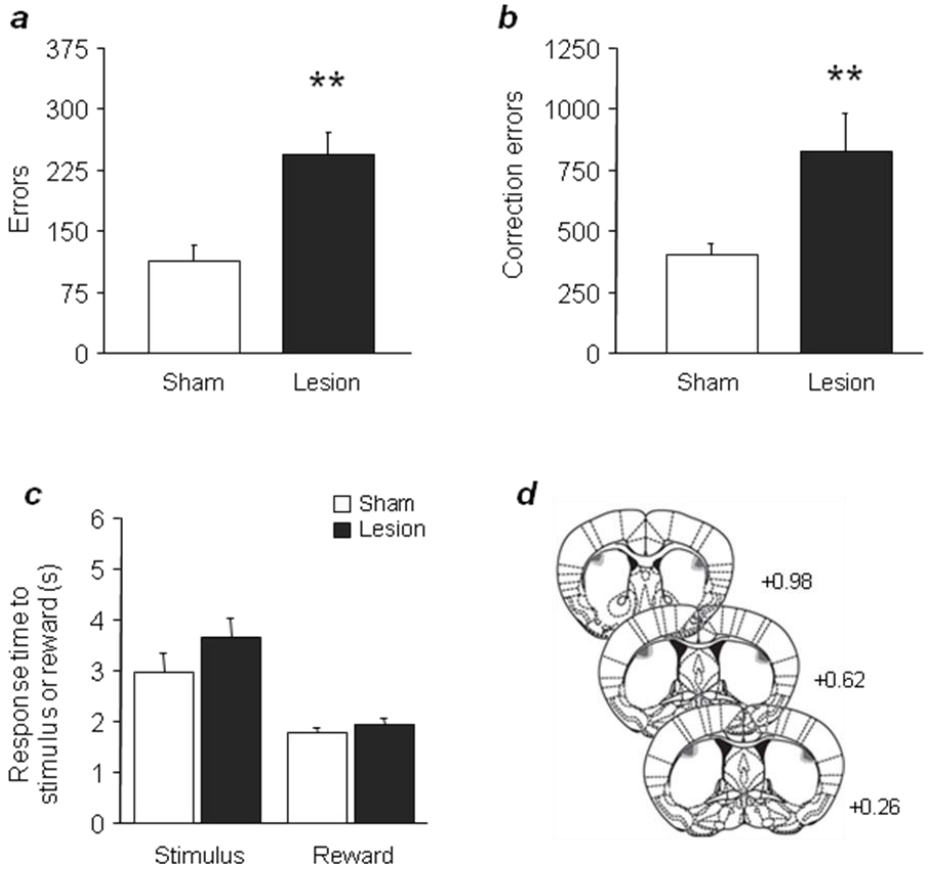


Nature Neuroscience: doi:10.1038/nn.2954

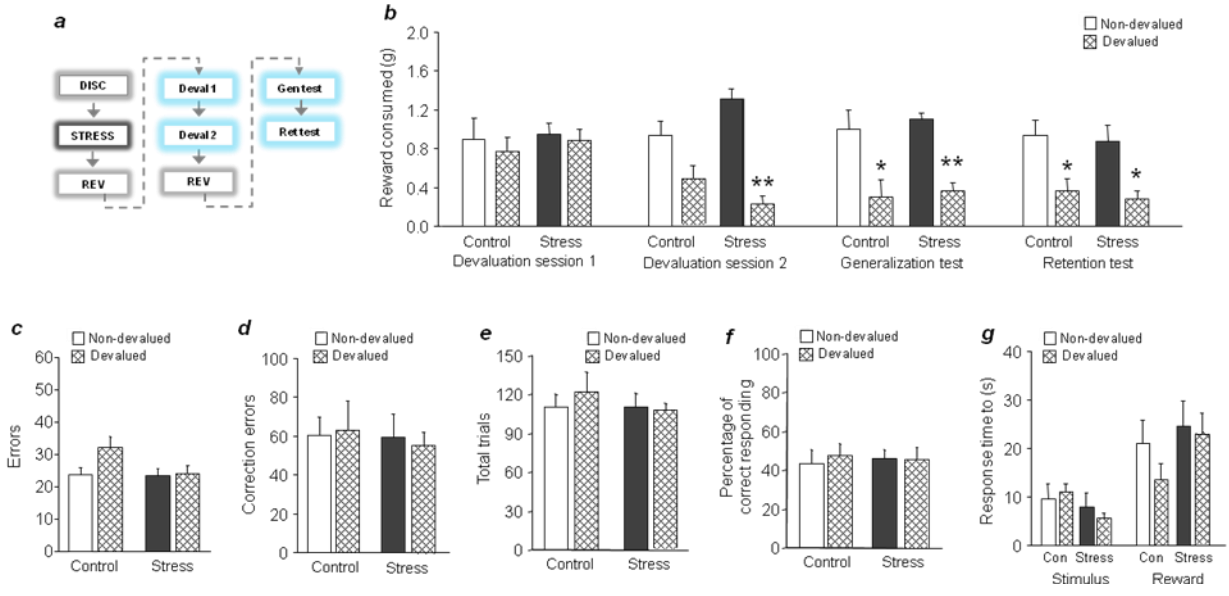
Supplemental Fig 2: Bilateral OFC lesions impair reversal. OFC lesioned mice made significantly more reversal errors (t(18)=2.26, *P*<.05) (***a***) and correction errors (t(18)=2.31, *P*<.05) (***b***) than sham controls. (***c***) Stimulus and reward response latencies were no different between lesioned mice and sham controls. (***d***) Schematic reconstruction of OFC lesions (black=minimum lesion extent, grey=maximum). n=7-13/lesion group. Data are Means ±SEM. **P*<.05 versus sham.



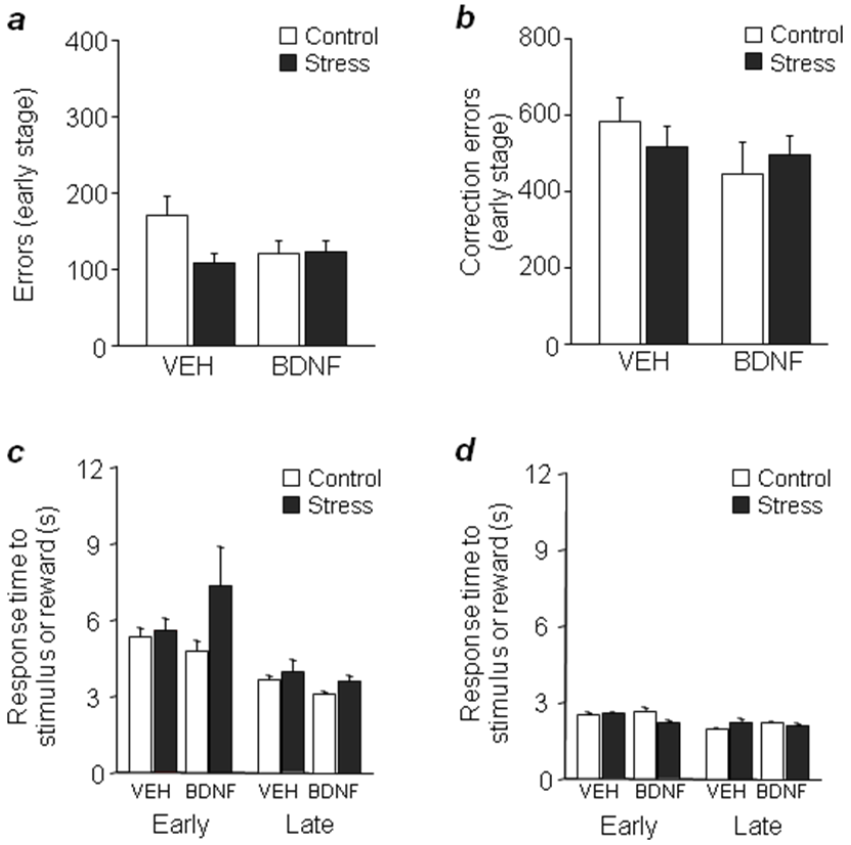
Supplemental Fig. 3: Bilateral DLS lesions impair reversal learning. DLS lesioned mice made significantly more reversal errors (t(14)=3.91, *P*<.01) (***a***) and correction errors (t(14)=2.93, *P*<.01) (***b***) than sham controls. (***c***) Stimulus and reward response latencies were no different between lesioned mice and sham controls. (***d***) Schematic reconstruction of DLS lesions (black=minimum lesion extent, grey=maximum). n=9/lesion group. Data are Means ±SEM. ***P*<.01 versus sham.



Supplemental Fig. 4: Reversal training produces habit-like insensitivity to outcome value. (***a***) Schematic summary of devaluation procedure. (***b***) Stressed and control groups consumed equal amounts of freely available food reward on the devaluation session 1, while devalued stressed mice consumed significantly less reward than non-devalued stressed mice on devaluation session 2 (t(9)=8.04, *P*<.01). Controls showed the same trend. On the (operant chamber) generalization test, the devalued stressed (t(9)=6.85, *P*<.01) and control (t(7)=2.52, *P*<.05) groups showed significant aversion to the food reward, as compared non-devalued counterparts. On the (home cage) retention test, the devalued stressed (t(9)=2.84, *P*<.05) and control (t(7)=3.19, *P*<.05) groups showed significant aversion to the food reward, as compared non-devalued counterparts. During the reversal probe test, devalued stressed and control groups made a similar number of reversal errors (***c***) and correction errors (***d***), and total trials (***e***), and showed similar % correct responding (***f***) and stimulus- and reward-retrieval latencies (***g***) as non-devalued counterparts. n=4-6/devaluation group/stress treatment. Data are Means ±SEM. **P*<.05, ***P*<.01 vs. non-devalued/same stress condition.



Supplemental Fig. 5: BDNF infusions prior to reversal did not affect early reversal learning. Stressed mice did not significantly differ in the number of errors (*a*) or correction errors (*b*) in either the vehicle- or BDNF-infused, though there was a trend for vehicle-infused stressed mice to make fewer errors (t(17)=1.86, *P*=08) than vehicle-infused controls. There was no effect of stress on response times in either the vehicle- or BDNF-infused groups at either the early (c) or late (d) stage of reversal. n=7-12/stress treatment/infusion group. Data are Means ±SEM.



References

1. Karlsson, R.M., *et al.* Assessment of glutamate transporter GLAST (EAAT1)-deficient mice for phenotypes relevant to the negative and executive/cognitive symptoms of schizophrenia. *Neuropsychopharmacology* **34**, 1578-1589 (2009).

2. Izquierdo, A., *et al.* Genetic and dopaminergic modulation of reversal learning in a touchscreen-based operant procedure for mice. *Behav Brain Res* **171**, 181-188 (2006).

3. Brigman, J.L., *et al.* Pharmacological or Genetic Inactivation of the Serotonin Transporter Improves Reversal Learning in Mice. *Cereb Cortex* (2010).

4. Jones, B. & Mishkin, M. Limbic lesions and the problem of stimulus--reinforcement associations. *Exp Neurol* **36**, 362-377 (1972).

5. Chudasama, Y. & Robbins, T.W. Dissociable contributions of the orbitofrontal and infralimbic cortex to pavlovian autoshaping and discrimination reversal learning: further evidence for the functional heterogeneity of the rodent frontal cortex. *J Neurosci* **23**, 8771-8780 (2003).

6. Lederle, L., *et al.* Reward-related behavioral paradigms for addiction research in the mouse: performance of common inbred strains. *PLoS One* **6**, e15536 (2011).

7. Schoenbaum, G. & Setlow, B. Cocaine makes actions insensitive to outcomes but not extinction: implications for altered orbitofrontal-amygdalar function. *Cereb Cortex* **15**, 1162-1169 (2005).

8. Izquierdo, A., Wellman, C.L. & Holmes, A. Brief uncontrollable stress causes dendritic retraction in infralimbic cortex and resistance to fear extinction in mice. *J Neurosci* **26**, 5733-5738 (2006).

9. Hefner, K., *et al.* Impaired fear extinction learning and cortico-amygdala circuit abnormalities in a common genetic mouse strain. *J Neurosci* **28**, 8074-8085 (2008).

10. Millstein, R.A., Ralph, R.J., Yang, R.J. & Holmes, A. Effects of repeated maternal separation on prepulse inhibition of startle across inbred mouse strains. *Genes Brain Behav* **5**, 346-354 (2006).