Behavioral/Systems/Cognitive

Preventing the Stress-Induced Shift from Goal-Directed to Habit Action with a β-Adrenergic Antagonist

Lars Schwabe,[1] Oliver Höffken,[2] Martin Tegenthoff,[2] and Oliver T. Wolf[1]

[1]Department of Cognitive Psychology, Ruhr-University Bochum, 44780 Bochum, Germany, and [2]Department of Neurology, Ruhr-University Bochum, BG-Kliniken Bergmannsheil, 44789 Bochum, Germany

Stress modulates instrumental action in favor of habit processes that encode the association between a response and preceding stimuli and at the expense of goal-directed processes that learn the association between an action and the motivational value of the outcome. Here, we asked whether this stress-induced shift from goal-directed to habit action is dependent on noradrenergic activation and may therefore be blocked by a β-adrenoceptor antagonist. To this end, healthy men and women were administered a placebo or the β-adrenoceptor antagonist propranolol before they underwent a stress or a control procedure. Shortly after the stress or control procedure, participants were trained in two instrumental actions that led to two distinct food outcomes. After training, one of the food outcomes was selectively devalued by feeding participants to satiety with that food. A subsequent extinction test indicated whether instrumental behavior was goal-directed or habitual. As expected, stress after placebo rendered participants' behavior insensitive to the change in the value of the outcome and thus habitual. After propranolol intake, however, stressed participants behaved, same as controls, goal-directed, suggesting that propranolol blocked the stress-induced bias toward habit behavior. Our findings show that the shift from goal-directed to habitual control of instrumental action under stress necessitates noradrenergic activation and could have important clinical implications, particularly for addictive disorders.

Introduction

Learning how to achieve a pleasant state or how to avoid an unpleasant state (i.e., instrumental learning) can be controlled by two anatomically and functionally distinct systems that operate in tandem (Adams, 1982; Balleine and Dickinson, 1991): (1) a goal-directed system that learns causal relationships between an action and the incentive value of the outcome and is supported by the medial prefrontal cortex, the dorsomedial striatum, and the dorsomedial thalamus (Balleine and Dickinson, 1998a; Corbit et al., 2003; Killcross and Coutureau, 2003; Yin et al., 2005; Valentin et al., 2007); and (2) a dorsolateral striatum-dependent habit system that encodes associations between a response and preceding stimuli, without any link to the outcome that is engendered by the response (Yin et al., 2004, 2006; Tricomi et al., 2009; Balleine and O'Doherty, 2010).

Recent research demonstrates that stressful experiences, whether acute or chronic, modulate the systems controlling instrumental learning in a manner that favors habit behavior over goal-directed behavior (Dias-Ferreira et al., 2009; Schwabe and Wolf, 2009, 2010). This stress effect can be mimicked by the simultaneous administration of glucocorticoid stress hormones (cortisol in humans) and yohimbine, a α2-adrenoceptor antagonist that increases noradrenergic stimulation (Schwabe et al., 2010). Glucocorticoids or yohimbine alone did not induce habit action, thus suggesting that glucocorticoids and noradrenaline interact to shift instrumental action from goal-directed to habitual control.

Similar interactions between noradrenaline and glucocorticoids are well documented for stress effects on hippocampus-dependent spatial or episodic memory processes, which require concurrent glucocorticoid and noradrenergic activity (Roozendaal et al., 2009). Most interestingly, it has been repeatedly shown that the effects of stress on hippocampus-dependent memory can be abolished when individuals are tested in a non-arousing environment or after administration of a β-adrenergic antagonist (Kuhlmann and Wolf, 2006; Roozendaal et al., 2006a; de Quervain et al., 2007; Schwabe et al., 2009). If stress effects on hippocampus-dependent memory and stress effects on the interaction of multiple (i.e., goal-directed and habitual) memory systems in instrumental learning share a common neuroendocrine mechanism, it is tempting to hypothesize that the shift from goal-directed to habit action under stress may be prevented by a β-adrenergic antagonist too.

The present experiment tested this hypothesis. In a placebo-controlled between-subject design, participants received either a placebo or the β-adrenergic antagonist propranolol (double-blind administration) before they were exposed to stress or a control condition. Shortly after the stress (or control) procedure, participants were trained in two instrumental actions leading to two distinct food rewards. To separate goal-directed from habit

Received June 29, 2011; revised Aug. 8, 2011; accepted Oct. 12, 2011.

Author contributions: L.S. and O.T.W. designed research; L.S. performed research; L.S. analyzed data; L.S., O.H., M.T., and O.T.W. wrote the paper.

This work was supported by DFG Grant SCHW1357/2-1. We gratefully acknowledge the assistance of Florian Watzlawik, Valerie Kinner, and Carsten Siebert during data collection. We thank Tobias Otto for his technical assistance.

Correspondence should be addressed to Dr. Lars Schwabe, Department of Cognitive Psychology, Ruhr-University Bochum, Universitaetsstrasse 150, 44780 Bochum, Germany. E-mail: Lars.Schwabe@rub.de.

DOI:10.1523/JNEUROSCI.3304-11.2011

 0270-6474/11/3117317-09$15.00/0

processes, we used a devaluation procedure (Balleine and Dickinson, 1998b; Valentin et al., 2007): participants were allowed to eat one of the rewards to satiety. A subsequent extinction test assessed whether participants' instrumental behavior was sensitive or insensitive to the change in the value of the goal, i.e., whether instrumental action was under goal-directed or habitual control. We predicted that stress would bias behavior toward habits and that this stress-induced shift from goal-directed to habitual action would be blocked by prior administration of propranolol.

Materials and Methods

Eighty healthy, normal-weight adults [40 women, 40 men; age, mean (M) ± SEM: 24.3 ± 0.3 years; body mass index: 22.8 ± 0.3 kg/m^2) participated in this experiment. Exclusion criteria were checked in a standardized interview and comprised any current or chronic medical condition, use of medication, current or lifetime history of any psychiatric disorder, smoking, use of hormonal contraceptives, current or planned diet, as well as any food intolerances. Furthermore, we prescreened participants to exclude those who do not like the food rewards that were used in this study (i.e., orange juice and chocolate milk). Nevertheless, 10 participants (5 men, 5 women) revealed during the experiment that they did not like at least one of the used food rewards (pleasantness ratings and percentage of food-associated high-probability actions during training >2 SDs below the mean) and were therefore excluded from the analyses (Valentin et al., 2007; Schwabe and Wolf, 2009).

We used a fully crossed between-subject design with the factors drug (placebo vs 40 mg of propranolol) and treatment (control vs stress), thus resulting in four experimental groups: placebo/control ($n = 18$), placebo/stress ($n = 18$), propranolol/control ($n = 17$), and propranolol/stress ($n = 17$). The study procedure was approved by the ethics committee of the Medical Faculty of the Ruhr-University Bochum.

Stress protocol. Participants in the stress condition were exposed to the Socially Evaluated Cold Pressor Test (SECPT), as described in detail previously (Schwabe et al., 2008). In brief, participants immersed their right hand up to and including the wrist for 3 min (or until they could not tolerate it any more) into ice water (0−2°C). During hand immersion, they were videotaped and monitored by a rather cold and unsociable experimenter. Participants in the control condition submerged their right hand up to and including the wrist for 3 min into warm water (35−37°C); they were neither videotaped nor monitored.

To assess the efficacy of the stress manipulation and the action of the drug, we took subjective stress ratings, salivary cortisol measurements, and blood pressure measurements at several time points across the experiment.

Subjective stress ratings. Immediately after the SECPT or control condition participants indicated on a scale from 0 ("not at all") to 100 ("very much") how stressful, painful, and unpleasant they had experienced the previous situation.

Salivary cortisol. Participants collected saliva samples before drug intake, before the SECPT/control condition as well immediately, 25 min, and 65 min after the SECPT/control condition with Salivette collection devices (Sarstedt). Saliva samples were kept at −20°C until analysis. Free cortisol concentrations were measured using an immunoassay (IBL International). Interassay and intra-assay coefficients of variance were below 10%.

Blood pressure measurement. Blood pressure measurements were taken with a Dinamap system (Critikon) before drug intake, before the beginning of the SECPT/control condition (~40 min after drug intake), during the SECPT/control condition as well as immediately, 25 min, and 65 min after the SECPT/control condition.

Instrumental learning task. The instrumental learning task that was used in the present experiment is described in detail previously (Valentin et al., 2007; Schwabe and Wolf, 2009). Briefly, participants were presented three trial types: chocolate, orange, and neutral (Fig. 1*A*). On each trial, participants were asked to choose between two actions represented by two distinct symbols on a computer screen. After subjects had selected one of the symbols by moving the mouse cursor to the symbol and pressing the left mouse button, the referring symbol was highlighted for 3 s and 1 ml of a liquid food or else no liquid was delivered, according to the reward schedule associated with the chosen action. The liquids were delivered with separate electronic pumps (one pump for each liquid) and transferred via 3-m-long tubes (diameter: 3 mm) to the participants who kept the ends of the tubes between the lips. Importantly, the two actions per trial type differed in the probability with which a food outcome was delivered. While one action was followed with a probability of $p = 0.70$ by a food outcome (high-probability action), the probability of a food outcome was $p = 0.20$ for the other action (low probability action). On the chocolate and orange trials, the high-probability action led to chocolate milk and orange juice, respectively, with a probability of $p = 0.50$ and to a common outcome (peppermint tea) with a probability of $p = 0.20$ (the reward and the common outcome were never presented in the same trial). On both trial types, the low probability action was never associated with the rewards but led only to the common outcome with a probability of $p = 0.20$. In neutral trials, water was delivered, either with a probability of $p = 0.70$ (high-probability action) or $p = 0.20$ (low probability action). By comparing performance in these trials to the performance in chocolate and orange trials, this neutral condition served as a control to assess the effect of the rewards (chocolate milk, orange juice) on participants' choice behavior.

Participants completed 75 trials for each trial type, resulting in 225 trials in total (intertrial interval: 8 s). The occurrence of the trial types was fully randomized. The specific assignment of the symbols and the positions on the computer screen to each action was held constant for each subject but counterbalanced across participants.

Selective outcome devaluation. After subjects had completed the learning task, they were invited to eat either oranges or chocolate pudding until they did not want to eat any more. This served to decrease the value of one food outcome (e.g., eating chocolate pudding to satiety should decrease the value of chocolate milk) whereas the value of the other outcome (orange juice in the example) should remain intact. The specific food used for devaluation was counterbalanced across participants.

Extinction test. The effect of the selective outcome devaluation on instrumental behavior was assessed in an extinction test given shortly after the devaluation procedure. Participants completed another 75 trials of each of the three trial types in which they were asked to choose between the two possible actions. The basic procedure was the same as during the learning session. This time, however, the rewards (the devalued and nondevalued food outcomes) were no longer presented, i.e., participants were tested in extinction for these outcomes. Both in the chocolate and in the orange trials, the two alternative actions delivered the common outcome (peppermint tea) with a probability of $p = 0.20$. In the neutral trials, water was now available with the equal probability of $p = 0.20$ for both actions.

Performance in this extinction test revealed whether instrumental behavior was goal-directed or habitual. Decreased responding to the action associated with the devalued food outcome relative to the action associated with the valued food outcome indicated goal-directed behavior. The ongoing choice of the devalued instrumental action was indicative for habit behavior (Balleine and Dickinson, 1998b; Valentin et al., 2007).

Procedure. All testing took place between 1:30 and 6:30 P.M. to control for the diurnal rhythm of the stress hormone cortisol. Participants were asked to refrain from caffeine and physical exercise within the 6 h before testing and not to eat anything for at least 3 h before the experiment started. After their arrival at the lab, participants gave a first saliva sample and baseline measurements of blood pressure were taken. Next, participants took a placebo or a propranolol (40 mg, Propra-ratiopharm) pill. Neither the participants nor the experimenter knew which drug was administered (double-blind drug administration). The dose of propranolol was chosen in line with earlier studies that examined the influence of a β-adrenergic antagonist on stress-induced changes in learning and memory (de Quervain et al., 2007; Schwabe et al., 2009). After a 40 min break during which participants were allowed to read, another saliva sample and blood pressure measurements were taken before participants underwent the SECPT or the control condition; blood pressure was also measured during the treatment. Afterwards, another saliva sample was

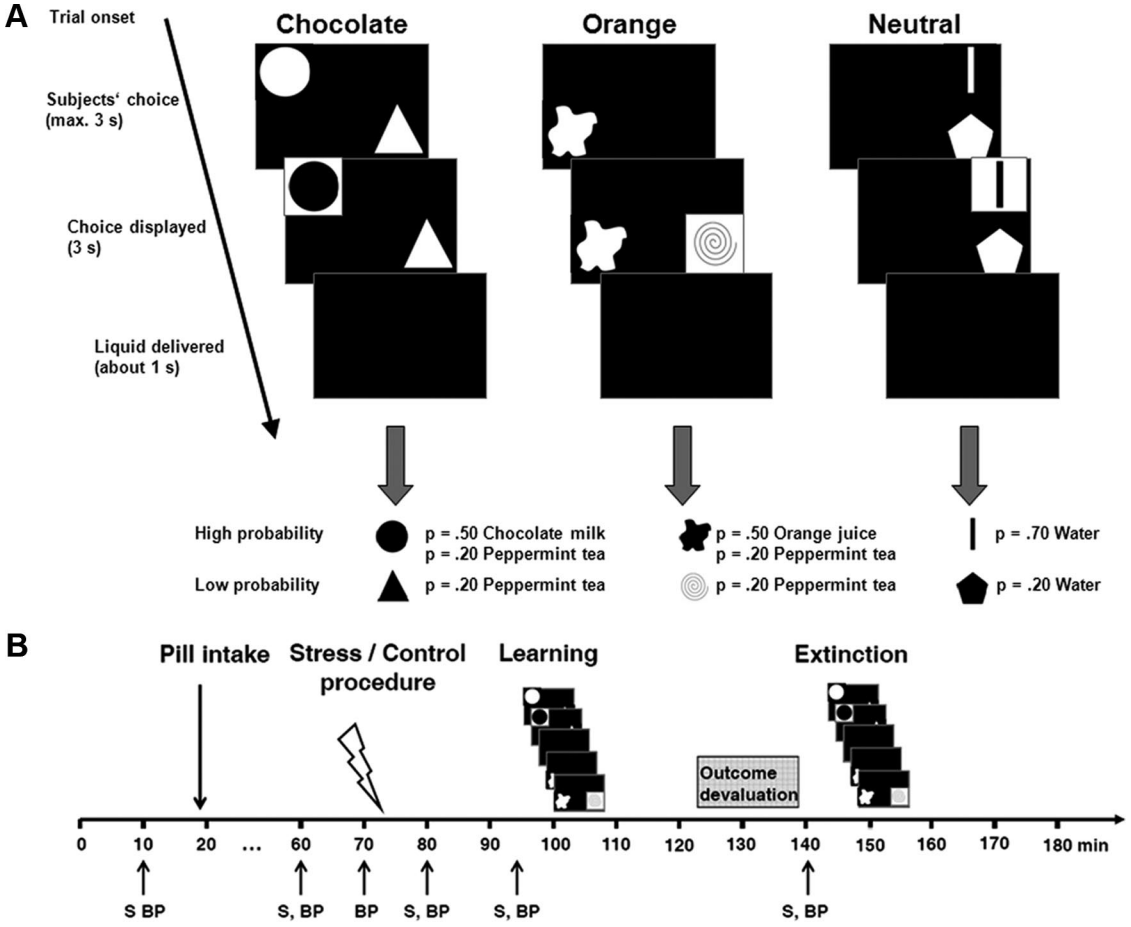

Figure 1. Instrumental task and time line of the experiment. ***A***, Participants completed three trial types (chocolate, orange, and neutral). In each trial type, there was one action that led with a high probability to a food outcome and one action that led with a low probability to a food outcome. Depending on the trial type, the high-probability action yielded chocolate milk or orange juice with a probability of $p = 0.5$, a common outcome (peppermint tea) with a probability of $p = 0.2$, or nothing. The low probability action led to the common liquid with a probability of $p = 0.2$. After an action was chosen, the referring symbol was highlighted for 3 s before 1 ml of the food was delivered. ***B***, Approximately 45 min after the intake of a placebo or a propranolol pill, participants were exposed to a stressor (Socially Evaluated Cold Pressor Test) or a control condition. Twenty-five minutes later, they were trained in the instrumental task. After training, one of the food outcomes was devalued by feeding participants to satiety with that food. Finally, participants completed an extinction test during which the food rewards (chocolate milk and orange juice) were no longer presented. ***S***, Saliva sample; BP, blood pressure measurement. Parts of this figure are reproduced from Schwabe and Wolf (2009), with permission of the Society for Neuroscience.

taken and blood pressure was measured. In addition, participants rated how stressful, painful, and unpleasant they had experienced the previous stress and control condition, respectively. Twenty-five minutes after the treatment and ~75 min after the drug intake, another saliva sample was taken, blood pressure was measured again, and participants rated on a scale from 0 ("not at all") to 100 ("very much") how pleasant they found the liquids that were used in the study. Next, the learning session of the instrumental learning task started. After learning and another pleasantness rating, participants could eat as many oranges or as much chocolate pudding as they wanted (selective outcome devaluation) and rated the pleasantness of the liquids again afterwards. Immediately after the outcome devaluation, a last saliva sample was taken and blood pressure was measured before the participants completed the critical extinction test (Fig. 1*B*).

Statistical analysis. Blood pressure, salivary cortisol, and subjective data were analyzed by treatment (stress vs control) × drug (placebo vs propranolol) × time point of measurement ANOVAs. Participants' behavior in the instrumental learning task was subjected to treatment × drug × value × block (5 blocks with 15 trials per block) ANOVAs. Significant main or interaction effects were pursued by appropriate follow-up tests that were Bonferroni corrected if indicated. All reported p-values are two-tailed; the partial η^2 was used as a measure of effect size.

Results

Subjective, endocrine, and autonomic changes following stress and propranolol intake

Participants' subjective stress ratings in combination with elevations in salivary cortisol and blood pressure verified the successful stress induction by the SECPT.

Subjective feeling

As expected, participants that were exposed to the SECPT rated the hand immersion as significantly more painful (mean ± SEM: 67.4 ± 4.0 vs 1.4 ± 1.2), stressful (50.3 ± 3.3 vs 4.6 ± 1.6), and unpleasant (65.7 ± 4.0 vs 4.0 ± 1.5; all $p < 0.001$) than participants in the control condition. Propranolol had no effect on the subjective stress ratings (all main or interaction effects $p > 0.10$).

Salivary cortisol

Salivary cortisol increased in response to the SECPT but not in response to the control condition (time × treatment interaction: $F_{(4,252)} = 15.18$, $p < 0.001$, $\eta^2 = 0.19$). As shown in Figure 2*A*, participants in the stress and control groups did not differ in their cortisol concentrations before drug intake, before or immediately after SECPT/control condition (all p values >0.25). However,

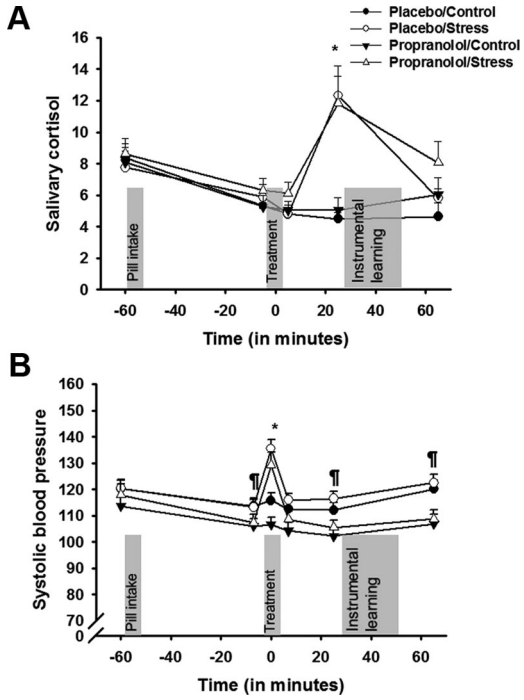

Figure 2. Physiological responses to stress and propranolol intake. ***A***, Salivary cortisol (in nmol/L) increased significantly in response to the SECPT but not in response to the control condition. Cortisol levels were not affected by propranolol. ***B***, Systolic blood pressure (in mmHg) was elevated during the SECPT but not during the control condition. Propranolol reduced systolic blood pressure, while it did not abolish the stress-induced blood pressure increase. Data represent mean ± SEM. Significant difference between the stress and the control groups, $^{*}p < 0.05$. Significant difference between the placebo and the propranolol groups, $^{¶}p < 0.05$.

stressed participants had significantly higher cortisol levels at the beginning of the instrumental learning task ($p < 0.001$). Propranolol did not affect baseline cortisol levels or the cortisol response to the SECPT (all p values >0.38).

Blood pressure

The exposure to the SECPT led to a significant increase in systolic blood pressure, which was not seen in the control condition (time × treatment interaction: $F_{(5,325)} = 28.99$, $p < 0.001$, $\eta^2 = 0.31$). Figure 2*B* shows that participants in the stress group had a significantly higher systolic blood pressure during the SECPT/control condition ($p < 0.001$), whereas there were no group differences at any time point before or after the SECPT/control condition (all p values >0.16). The same pattern of results was obtained for diastolic blood pressure (time × treatment interaction: $F_{(5,325)} = 32.11$, $p < 0.001$, $\eta^2 = 0.33$; group difference during treatment: $p < 0.001$, all other $p > 0.45$; data not shown).

Moreover, the action of the drug (placebo vs propranolol) was reflected in systematic changes in systolic and diastolic blood pressure: Although there were no group differences before drug intake, blood pressure decreased significantly over time in the propranolol groups but not in the placebo groups (time × drug interactions for systolic and diastolic blood pressure: both $F_{(5,325)} > 2.65$, both $p < 0.05$, both $\eta^2 > 0.04$). It is important to note, that the stress-induced increase in blood pressure was also present under propranolol (time × treatment × drug interactions for systolic and diastolic blood pressure: both $F_{(5,325)} < 1.20$, both $p > 0.30$, both $\eta^2 < 0.02$; Fig. 2*B*).

Instrumental learning remained unaffected by stress and propranolol

Figure 3 shows that all participants increasingly preferred the high-probability actions that were associated with chocolate milk and orange juice, respectively, over their low-probability counterparts across training (block effects for chocolate and orange trials: both $F_{(4,264)} > 17$, both $p < 0.001$, both $\eta^2 > 0.20$), thus indicating successful instrumental learning. In neutral trials, however, participants showed no such preference ($p = 0.19$; block × trial type interaction: $F_{(8,524)} = 3.93$, $p < 0.01$, $\eta^2 = 0.06$) which suggests that participants were indifferent as to whether they received the effectively neutral outcome or not. Importantly, there were no effects of treatment or drug on the learning curves in the instrumental task (all main or interaction effects: all p values >0.25).

Selective outcome devaluation was not influenced by stress or propranolol

During the outcome devaluation procedure, participants ate on average 1.74 cups of chocolate pudding (SEM: 0.11; 150 g per cup) or 2.12 oranges (SEM: 0.13). Not surprisingly, eating chocolate pudding or oranges to satiety led to a significant drop in subjective hunger ratings, from 61.6 (SEM: 2.9) before the devaluation procedure to 36.1 (SEM: 3.0) after the devaluation procedure (main effect time: $F_{(1,66)} = 95.89$, $p < 0.001$, $\eta^2 = 0.59$). Participants' subjective pleasantness ratings, however, revealed that the devaluation procedure affected primarily the value of the food that was eaten to satiety (Fig. 4). Although pleasantness ratings decreased significantly for chocolate milk and orange juice after the devaluation procedure (both p values < 0.05), which is most likely due to general satiety effects, this decrease was significantly stronger for the food that was eaten to satiety (food × time interaction: $F_{(1,66)} = 16.93$, $p < 0.001$, $\eta^2 = 0.20$).

Neither the amount of food that was consumed, nor the subjective hunger or pleasantness ratings were affected by stress or propranolol (all main and interaction effects: all p values >0.35).

Propranolol blocked the stress-induced shift from goal-directed to habit action

The performance in the extinction test shortly after the outcome devaluation revealed whether behavior was goal-directed or habitual. Goal-directed behavior is indicated by decreased responding to the action that was previously associated with the now devalued action, habit action by the absence of such a decrease in responding to the devalued action.

As shown in Figure 5, participants in the placebo/control group behaved in a goal-directed manner. Consistent with their pleasantness ratings, they chose the high-probability action that was associated with the valued outcome significantly more often than the high-probability action that was associated with the devalued outcome ($F_{(1,17)} = 11.22$, $p < 0.001$, $\eta^2 = 0.40$). At the beginning of the extinction session, before they could know that chocolate milk and orange juice would not be presented any longer, they still preferred the valued high-probability action over its low-probability counterpart (binomial test; $p < 0.01$). In the devalued trials, however, they did not favor the high-probability action but even tended to avoid the action that was previously associated with the now devalued outcome ($p = 0.08$).

In contrast to participants in the placebo/control group, participants in the placebo/stress group chose the devalued high-probability action as often as the valued high-probability action ($F_{(1,17)} < 1$, $p = 0.75$, $\eta^2 < 0.01$); they preferred both high-probability actions over the referring low-probability actions

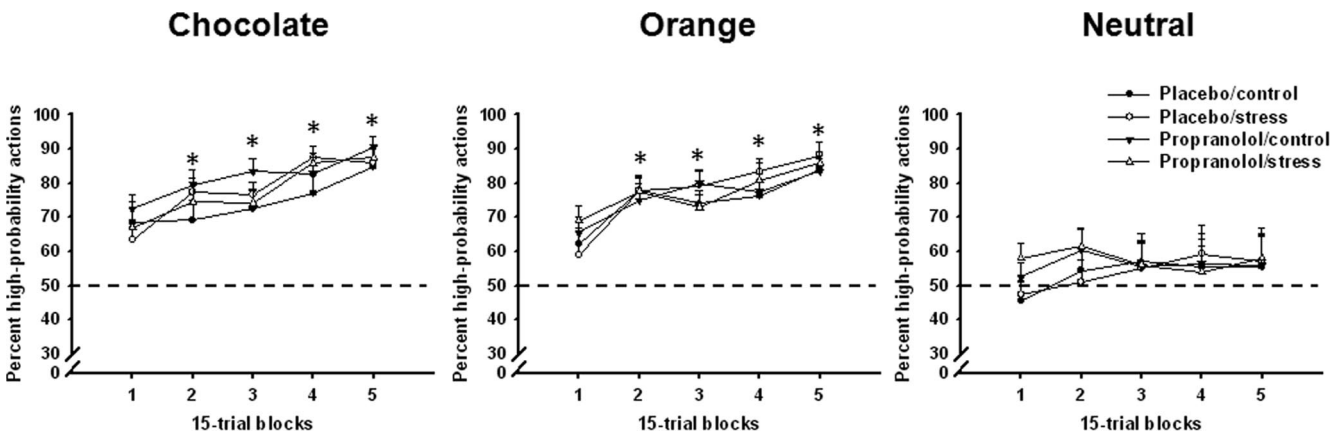

Figure 3. Learning curves in the instrumental task. Regardless of the experimental group, participants increasingly favored the high-probability actions associated with chocolate milk and orange juice over the referring low-probability actions across the learning session (*$p < 0.05$). No such preference was found in neutral trials. The dashed line indicates the percentage of high-probability actions of 50%, where participants were completely indifferent between the high- and low-probability actions. Data represent mean ± SEM.

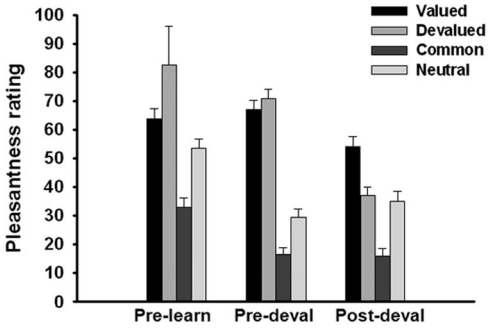

Figure 4. Subjective pleasantness ratings before instrumental learning as well as before and after the outcome devaluation procedure. Before the outcome devaluation, participants indicated that they found both rewards (i.e., valued and later devalued outcomes) pleasant. Eating either chocolate pudding or oranges to satiety led to a significant drop in the subjective pleasantness of the food that was eaten relative to the food that was not eaten. Pleasantness ratings were given on a scale from 0 ("not pleasant") to 100 ("very pleasant"). Data represent mean ± SEM.

(both p values <0.01). Thus, although participants in the placebo/stress condition indicated (same as those in the placebo/control group) that they did not want the devalued outcome anymore, they still performed the action associated with that outcome which suggests that the behavior of the stressed participants was habitual.

This stress-induced shift from goal-directed to habit action disappeared after propranolol intake. Under propranolol, all participants, regardless of whether they underwent the stress or the control condition, acted goal-directed. They chose the valued high-probability action significantly more often than the devalued high-probability action (both $F_{(1,16)} > 7.30$, both $p < 0.02$, both $\eta^2 > 0.30$) and preferred the valued high-probability action (both p values <0.01) but avoided the devalued high-probability action (both p values <0.05) in the first 15 extinction test trials.

In line with these interpretations, a value (valued vs devalued trials) × block (5 blocks with 15 trials per block) × treatment (stress vs control) × drug (placebo vs propranolol) ANOVA yielded a significant three-way interaction between value, treatment, and drug ($F_{(1,66)} = 3.84$, $p = 0.05$, $\eta^2 = 0.06$). Follow-up tests showed that stress altered participants' behavior in the devalued trials under placebo (value × treatment: $F_{(1,34)} = 5.32$, $p < 0.05$, $\eta^2 = 0.14$) but not under propranolol ($F_{(1,32)} < 1$, $p = 0.87$, $\eta^2 < 0.01$).

Because the interactive influence of drug and treatment appeared to be strongest at the beginning of the extinction session and because the influence of the outcome devaluation is clearest in the first extinction test trials, we analyzed the changes in instrumental behavior from the last training trials to the first extinction test trials by a value × block (last 15 training trials vs first 15 extinction trials) × treatment × drug ANOVA. This analysis yielded a significant four-way interaction suggesting that the impact of the treatment on the change in behavior from training to testing was modulated by the drug ($F_{(1,66)} = 7.23$, $p < 0.01$, $\eta^2 = 0.10$). Follow-up ANOVAs revealed that control participants in the placebo group decreased responding to the devalued action after the outcome devaluation (value × block interaction: $F_{(1,17)} = 19.16$, $p < 0.001$, $\eta^2 = 0.53$), whereas there was no such decrease in the stressed participants that were administered a placebo ($F_{(1,17)} < 1$, $p = 0.96$, $\eta^2 < 0.01$; value × block × treatment interaction: $F_{(1,34)} = 12.66$, $p = 0.001$, $\eta^2 = 0.27$). Thus, as shown in Figure 6, under placebo the behavior of controls but not the behavior of stressed participants was sensitive to changes in the value of the outcome. After propranolol administration, stress had no effect on the changes in instrumental action in response to the outcome devaluation ($F_{(1,32)} < 1$, $p = 0.99$, $\eta^2 < 0.01$) and both the stress and control group decreased responding to the devalued action from training to extinction testing (both $F_{(1,16)} > 20$, both $p < 0.001$, both $\eta^2 > 0.55$).

Propranolol abolished the association between cortisol and habit performance

To assess the role of glucocorticoids in the stress-induced shift from goal-directed to habit action, we correlated peak cortisol levels (i.e., cortisol levels 25 min after the treatment) with the percentage of devalued high-probability actions in the first 15 trials of the extinction test. These analyses showed that stress-induced cortisol levels correlated significantly with the sensitivity of participants' behavior to the outcome devaluation: higher cortisol levels were associated with habitual responding ($r = 0.42$, $p < 0.02$; Fig. 7*A*). Corroborating the idea that glucocorticoid effects on instrumental behavior require concurrent noradrenergic activity (Schwabe et al., 2010), the correlation between peak cortisol levels and habit action disappeared in participants that were administered propranolol ($r = 0.17$, $p = 0.34$; Fig. 7*B*).

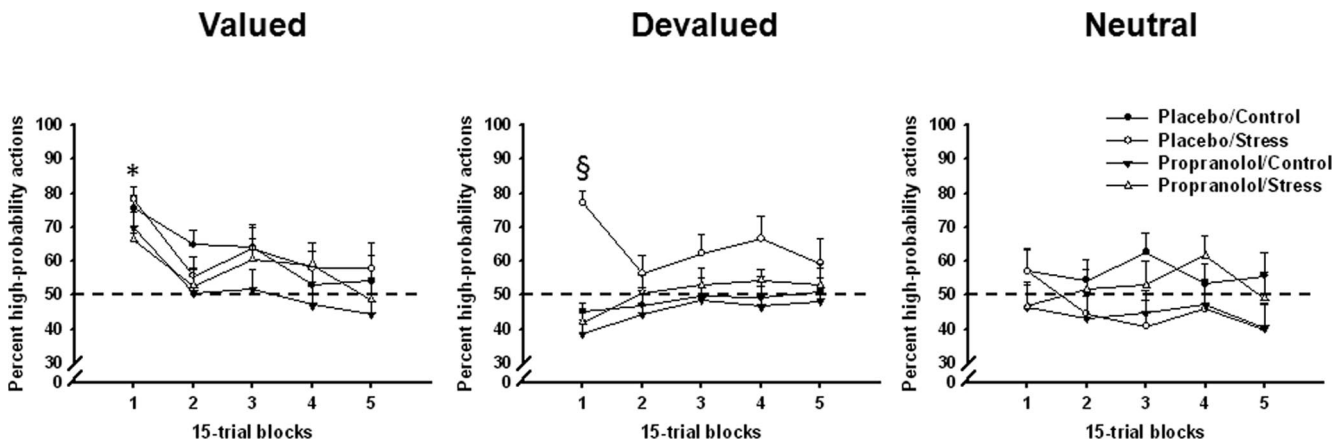

Figure 5. Valued, devalued, and neutral high-probability actions across the extinction test. All participants, regardless of the experimental group, favored the valued high-probability action over its low-probability counterpart in the first extinction test block (*$p < 0.05$), before they could know that the rewards were not presented any more. However, only stressed participants that had received a placebo favored the devalued high-probability action over the referring low-probability action ($^{§}p < 0.05$), suggesting that they were insensitive to the change in the value of the outcome (i.e., that they performed habitual). Participants in the other three groups tended to avoid the devalued high-probability action at the beginning of the extinction test. The dashed line indicates the percentage of high-probability actions of 50%, where participants were completely indifferent between the high-and low-probability actions. Data represent mean ± SEM.

Discussion

Recent findings indicate that stress promotes a shift from goal-directed to habitual instrumental action (Schwabe and Wolf, 2011). Here, we demonstrate that this stress-induced shift in instrumental action necessitates noradrenergic activity and thus provide first evidence how this shift can be prevented. We show that the administration of the β-adrenergic antagonist propranolol before the stress exposure abolished the stress-induced bias toward habit action as well as the association between stress-induced cortisol elevations and habitual behavior. Propranolol itself, however, did not affect instrumental behavior.

It is by now well established that stress effects on spatial memory or working memory processes depend on simultaneous glucocorticoid and noradrenergic activation in the basolateral part of the amygdala which then modulates the activity in other brain areas such as the hippocampus or the prefrontal cortex (McGaugh et al., 1996; Roozendaal et al., 2006b, 2009). Together with our previous findings (Schwabe et al., 2010), the present data suggest that concurrent glucocorticoid and noradrenergic activity is not only required for stress effects on a single (hippocampal or prefrontal cortical) memory system (Kuhlmann and Wolf, 2006; Roozendaal et al., 2006b; de Quervain et al., 2007) but also for the interaction of multiple memory systems in instrumental learning. Although there is some evidence for a role of the amygdala in the modulation of goal-directed and habit processes in instrumental action (Balleine et al., 2003; Balleine and Killcross, 2006), whether the (basolateral) amygdala is also for stress effects on instrumental learning the central mediator where glucocorticoid and noradrenaline effects converge remains to be shown by future lesion studies in rodents or human neuroimaging studies.

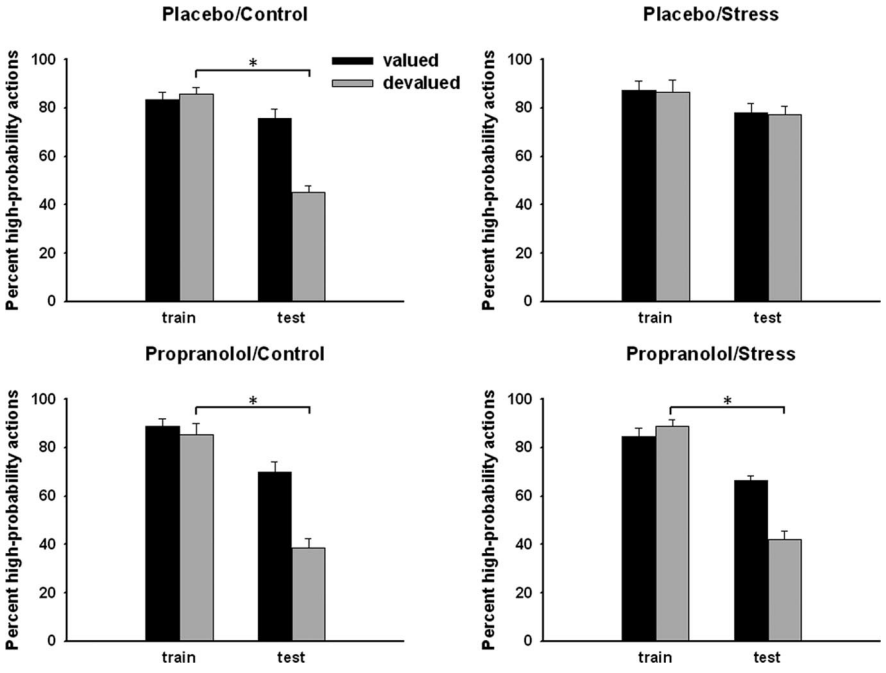

Figure 6. Changes in valued and devalued high-probability actions from the last 15 training trials to the first 15 extinction test trials. Under placebo, control participants but not stressed participants were sensitive to the change in the value of the outcome and decreased responding to the devalued high-probability action from training to extinction testing (*$p < 0.05$). After administration of propranolol, both control and stressed participants changed their instrumental behavior in response to the outcome devaluation, indicating that propranolol blocked the stress-induced shift from goal-directed to habitual control of instrumental action. Data represent mean ± SEM.

Our findings confirm the important role of noradrenergic arousal (in combination with glucocorticoids) for the shift from goal-directed to habit action that was suggested by our previous study (Schwabe et al., 2010). However, the present findings go far beyond the previous ones. Whereas our previous study showed that the combined (pharmacological) elevation of glucocorticoid and noradrenergic activity may bias instrumental behavior toward habits, we show here that the stress-induced shift from goal-directed to habit action can be prevented by a blockade of noradrenergic activity. It is important to note that stress is much more than an increase in glucocorticoid and noradrenaline levels.

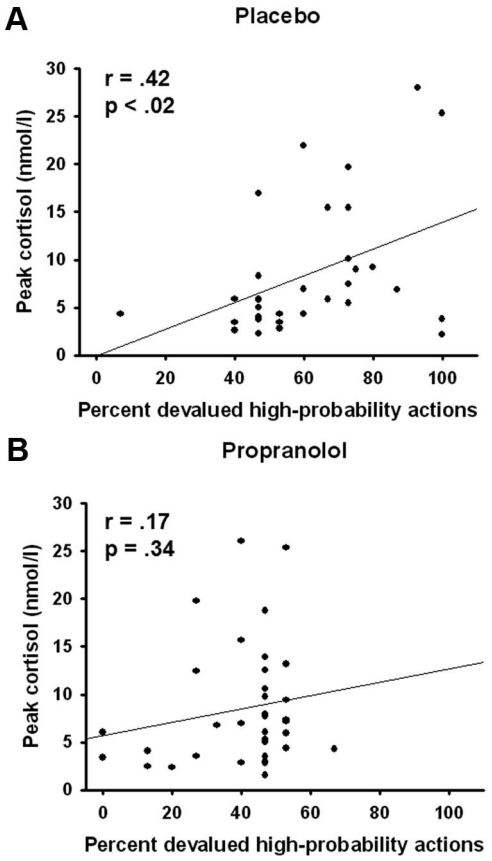

Figure 7. Correlations between peak salivary cortisol levels (25 min post-treatment) and the percentage of devalued high-probability actions in the first 15 extinction test trials. ***A***, High cortisol levels correlated significantly with habitual performance in the extinction test if participants received a placebo before learning. ***B***, Under propranolol, however, the correlation between cortisol and the percentage of devalued high-probability actions disappeared.

Numerous neurotransmitters, peptides, and hormones are released in response to stress (de Kloet et al., 2005; Joëls and Baram, 2009), many of which, including dopamine or corticotrophin releasing hormone, are known to affect learning and memory processes (Rossato et al., 2009; Chen et al., 2010). Thus, our previous findings could not rule out the possibility that other stress mediators might alter instrumental action as well, independently of glucocorticoids and noradrenaline. In other words, whereas our previous study (Schwabe et al., 2010) showed that concurrent glucocorticoid and noradrenergic activity is sufficient to render behavior habitual, the present findings show that noradrenergic activity is necessary for stress-induced changes in instrumental behavior. Knowing the necessary conditions for the shift from goal-directed to habit action is crucial for any attempts to prevent this shift. Indeed, the present findings are, to the best of our knowledge, the first to show how the stress-induced shift toward habit action can be prevented.

Interestingly, while the stress-induced shift from goal-directed to habit action became apparent in the extinction test after outcome devaluation, there was no visible effect of stress (or propranolol) on learning curves. This finding raises the question when stress exerted its effect on instrumental action. Did stress affect the acquisition or the expression of goal-directed versus habit behavior? If stress affected already the acquisition of the instrumental task, this would implicate that both goal-directed and habitual processes could control instrumental action equally well from early training on. This interpretation, however, is in conflict with the predominant view that instrumental action is initially under goal-directed control, whereas habits develop gradually over time, especially after overtraining (Balleine and Dickinson, 1991, 1998a; Dickinson et al., 1995; Yin and Knowlton, 2006). Alternatively, stress may not have affected the acquisition of instrumental action but the expression of the learned actions in the extinction test, implicating that both goal-directed and habit processes were developed at the end of the 75-trial learning session. In line with this latter view, there is evidence that stress may promote habit actions even when it is induced shortly before extinction testing, i.e., without having any effects on learning (Schwabe and Wolf, 2010). However, the conclusion that stress has no effect on the acquisition of instrumental actions might be premature because the influence of stress on goal-directed versus habit action was more pronounced when it was presented before learning than when it was presented before extinction testing (Schwabe and Wolf, 2009, 2010).

Depending on whether stress affected primarily the acquisition or the expression of instrumental action, genomic or non-genomic modes of glucocorticoid actions may have been involved. During the past years, it has become increasingly clear that, in addition to the classic, non-genomic pathway of glucocorticoid action, there are also rapid, non-genomic glucocorticoid effects that are mediated by membrane-bound receptors that activate a G-protein-coupled signaling cascade (Karst et al., 2005, 2010; Roozendaal et al., 2010; Groeneweg et al., 2011). These rapid, non-genomic glucocorticoid effects occur within minutes and are relatively short-lasting, whereas genomic glucocorticoid effects develop with a delay of more than an hour. In the present experiment, participants were exposed to stress ~30 min before learning and ~70 min before extinction testing, i.e., genomic glucocorticoid effects could not be present during learning but possibly during extinction testing. Interestingly, it has been suggested that genomic glucocorticoid effects reduce the functioning of the prefrontal cortex (Joëls et al., 2006), one of the key substrates of goal-directed action (Balleine and Dickinson, 1998a; Valentin et al., 2007). Thus, it might be tempting to speculate that genomic glucocorticoid effects set in during extinction testing and modulate brain systems in favor of those that are involved in habit action. This view is challenged, however, by the finding that stress favors habits also when it is induced shortly before extinction testing (i.e., when genomic glucocorticoid effects cannot have developed) (Schwabe and Wolf, 2010) and by evidence suggesting that noradrenaline interacts primarily with non-genomic glucocorticoid actions in modulating memory processes (Barsegyan et al., 2010; Roozendaal et al., 2010). Unraveling the exact roles of non-genomic and genomic glucocorticoid actions in stress effects on instrumental learning and other forms of learning and memory is still an ongoing endeavor and an important route for future research.

Given the well known role of glucocorticoids in energy regulation and reports of stress- or glucocorticoid-induced increases in appetite (Stone and Brownell, 1994; Bell et al., 2000; Epel et al., 2001), one might argue that stress does not influence the systems controlling instrumental action but that it just increases participants urge for food. Our data, however, speak clearly against such an interpretation. First, we did not find any effect of stress or glucocorticoids on the amount of food consumed, on subjective hunger or pleasantness ratings, in any of our studies, including the present one. Second, pharmacological manipulations of glucocorticoid levels did not influence participants' instrumental

behavior (Schwabe et al., 2010). Third, there is in our view no good reason why the metabolic effects of stress-induced elevations in glucocorticoid levels should disappear after β-adrenergic blockade or why they should occur specifically for the devalued outcome. Consequently, we do not think that metabolic glucocorticoid effects can account for the impact of stress on participants' sensitivity to the outcome devaluation and rather suggest that stress and stress hormones change the contribution of goal-directed and habit systems to instrumental action.

Together, our results demonstrate that the stress-induced shift from goal-directed to habit action is dependent on noradrenergic activity. Although this shift toward habits may be generally adaptive in that it promotes efficient behavior during stressful times, it may also contribute to psychiatric disorders such as addiction. Addictive behavior can be seen as the endpoint of a number of transitions from initially goal-directed, voluntary drug consumption to habitual and ultimately compulsive drug intake (Robbins and Everitt, 1999; Everitt and Robbins, 2005). Stress is one of the major risk factors for addiction and particularly for relapse (Sinha, 2001; Koob and Kreek, 2007). The aberrant engagement of habitual processes during stressful experiences may reinstate previously developed drug-taking routines and could hence facilitate relapse to drug use (Schwabe et al., 2011). Thus, the present findings provide not only novel insights into the modulation of goal-directed and habit action under stress but may also point to a potential use of β-adrenergic antagonists in the prevention of relapse to addictive behaviors.

References

Adams C (1982) Variations in the sensitivity of instrumental responding to reinforcer devaluation. Q J Exp Psychol 34:77–98.

Balleine BW, Dickinson A (1991) Instrumental performance following reinforcer devaluation depends upon incentive learning. Q J Exp Psychol 43:279–296.

Balleine BW, Dickinson A (1998a) Goal-directed instrumental action: contingency and incentive learning and their cortical substrates. Neuropharmacology 37:407–419.

Balleine BW, Dickinson A (1998b) The role of incentive learning in instrumental outcome evaluation by sensory-specific satiety. Anim Learn Behav 26:46–59.

Balleine BW, Killcross S (2006) Parallel incentive processing: an integrated view of amygdala function. Trends Neurosci 29:272–279.

Balleine BW, O'Doherty JP (2010) Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. Neuropsychopharmacology 35:48–69.

Balleine BW, Killcross AS, Dickinson A (2003) The effect of lesions of the basolateral amygdala on instrumental conditioning. J Neurosci 23:666–675.

Barsegyan A, Mackenzie SM, Kurose BD, McGaugh JL, Roozendaal B (2010) Glucocorticoids in the prefrontal cortex enhance memory consolidation and impair working memory by a common neural mechanism. Proc Natl Acad Sci U S A 107:16655–16660.

Bell ME, Bhatnagar S, Liang J, Soriano L, Nagy TR, Dallman MF (2000) Voluntary sucrose ingestion, like corticosterone replacement, prevents the metabolic deficits of adrenalectomy. J Neuroendocrinol 12:461–470.

Chen Y, Rex CS, Rice CJ, Dubé CM, Gall CM, Lynch G, Baram TZ (2010) Correlated memory defects and hippocampal dendritic spine loss after acute stress involve corticotrophin-releasing hormone signalling. Proc Natl Acad Sci U S A 107:13123–13128.

Corbit LH, Muir JL, Balleine BW (2003) Lesions of mediodorsal thalamus and anterior thalamic nuclei produce dissociable effects on instrumental conditioning in rats. Eur J Neurosci 18:1286–1294.

de Kloet ER, Joëls M, Holsboer F (2005) Stress and the brain: from adaptation to disease. Nat Rev Neurosci 6:463–475.

de Quervain DJ, Aerni A, Roozendaal B (2007) Preventive effect of β-adrenoceptor blockade on glucocorticoid-induced memory retrieval deficits. Am J Psychiatry 164:967–969.

Dias-Ferreira E, Sousa JC, Melo I, Morgado P, Mesquita AR, Cerqueira JJ, Costa RM, Sousa N (2009) Chronic stress causes frontostriatal reorganization and affects decision-making. Science 325:621–625.

Dickinson A, Balleine BW, Watt A, Gonzalez F, Boakes RA (1995) Motivational control after extended instrumental training. Anim Learn Behav 23:197–206.

Epel E, Lapidus R, McEwen B, Brownell K (2001) Stress may add a bite to appetite in women: a laboratory study of stress-induced cortisol and eating behavior. Psychoneuroendocrinology 26:37–49.

Everitt BJ, Robbins TW (2005) Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. Nat Neurosci 8:1481–1489.

Groeneweg FL, Karst H, de Kloet ER, Joëls M (2011) Rapid non-genomic effects of corticosteroids and their role in the central stress response. J Endocrinol 209:153–167.

Joëls M, Baram TZ (2009) The neuro-symphony of stress. Nat Rev Neurosci 10:459–466.

Joëls M, Pu Z, Wiegert O, Oitzl MS, Krugers HJ (2006) Learning under stress: how does it work? Trends Cogn Sci 10:152–158.

Karst H, Berger S, Turiault M, Tronche F, Schütz G, Joëls M (2005) Mineralocorticoid receptors are indispensable for nongenomic modulation of hippocampal glutamate transmission by corticosterone. Proc Natl Acad Sci U S A 102:19204–19207.

Karst H, Berger S, Erdmann G, Schütz G, Joëls M (2010) Metaplasticity of amygdalar responses to the stress hormone corticosterone. Proc Natl Acad Sci U S A 107:14449–14454.

Killcross S, Coutureau E (2003) Coordination of actions and habits in the medial prefrontal cortex of rats. Cereb Cortex 13:400–408.

Koob GF, Kreek MJ (2007) Stress, dysregulation of drug reward pathways, and the transition to drug dependence. Am J Psychiatry 164:1149–1159.

Kuhlmann S, Wolf OT (2006) A non-arousing test situation abolishes the impairing effects of cortisol on delayed memory retrieval in healthy women. Neurosci Lett 399:268–272.

McGaugh JL, Cahill L, Roozendaal B (1996) Involvement of the amygdala in memory storage: interaction with other brain systems. Proc Natl Acad Sci U S A 93:13508–13514.

Robbins TW, Everitt BJ (1999) Drug addiction: bad habits add up. Nature 398:567–570.

Roozendaal B, Okuda S, Van der Zee EA, McGaugh JL (2006a) Glucocorticoid enhancement of memory requires arousal-induced noradrenergic activation in the basolateral amygdala. Proc Natl Acad Sci U S A 103:6741–6746.

Roozendaal B, Okuda S, de Quervain DJ, McGaugh JL (2006b) Glucocorticoids interact with emotion-induced noradrenergic activation in influencing different memory functions. Neuroscience 138:901–910.

Roozendaal B, McEwen BS, Chattarji S (2009) Stress, memory and the amygdala. Nat Rev Neurosci 10:423–433.

Roozendaal B, Hernandez A, Cabrera SM, Hagewoud R, Malvaez M, Stefanko DP, Haettig J, Wood MA (2010) Membrane-associated glucocorticoid activity is necessary for modulation of long-term memory via chromatin modification. J Neurosci 30:5037–5046.

Rossato JI, Bevilaqua LR, Izquierdo I, Medina JH, Cammarota M (2009) Dopamine controls persistence of long-term memory storage. Science 325:1017–1020.

Schwabe L, Wolf OT (2009) Stress prompts habit behavior in humans. J Neurosci 29:7191–7198.

Schwabe L, Wolf OT (2010) Socially evaluated cold pressor stress after instrumental learning favors habits over goal-directed action. Psychoneuroendocrinology 35:977–986.

Schwabe L, Wolf OT (2011) Stress-induced modulation of instrumental behavior: from goal-directed to habitual control of action. Behav Brain Res 219:321–328.

Schwabe L, Haddad L, Schachinger H (2008) HPA axis activation by a socially evaluated cold pressor test. Psychoneuroendocrinology 33:890–895.

Schwabe L, Römer S, Richter S, Dockendorf S, Bilak B, Schächinger H (2009) Stress effects on declarative memory retrieval are blocked by a β-adrenoceptor antagonist in humans. Psychoneuroendocrinology 34:446–454.

Schwabe L, Tegenthoff M, Höffken O, Wolf OT (2010) Concurrent glucocorticoid and noradrenergic activity shifts instrumental behavior from goal-directed to habitual control. J Neurosci 30:8190–8196.

Schwabe L, Dickinson A, Wolf OT (2011) Stress, habits and drug addiction:

a psychoneuroendocrinological perspective. Exp Clin Psychopharmacol 19:53–63.

Sinha R (2001) How does stress increase risk of drug abuse and relapse? Psychopharmacology 158:343–359.

Stone A, Brownell K (1994) The stress-eating paradox: multiple daily measurements in adult males and females. Psychol Health 9:425–436.

Tricomi E, Balleine B, O'Doherty JP (2009) A specific role for posterior dorsolateral striatum in human habit learning. Eur J Neurosci 29: 225–2232.

Valentin VV, Dickinson A, O'Doherty JP (2007) Determining the neural substrates of goal-directed learning in the human brain. J Neurosci 27:4019–4026.

Yin HH, Knowlton BJ (2006) The role of the basal ganglia in habit formation. Nat Rev Neurosci 7:464–476.

Yin HH, Knowlton BJ, Balleine BW (2004) Lesions of dorsolateral striatum preserve outcome expectancy but disrupt habit formation in instrumental learning. Eur J Neurosci 19:181–189.

Yin HH, Ostlund SB, Knowlton BJ, Balleine BW (2005) The role of the dorsomedial striatum in instrumental conditioning. Eur J Neurosci 22:513–523.

Yin HH, Knowlton BJ, Balleine BW (2006) Inactivation of dorsolateral striatum enhances sensitivity to changes in the action-outcome contingency in instrumental conditioning. Behav Brain Res 166:189–196.