

Stress Prompts Habit Behavior in Humans

Lars Schwabe and Oliver T. Wolf

Department of Cognitive Psychology, Ruhr University Bochum, 44780 Bochum, Germany

Instrumental behavior can be controlled by goal-directed action–outcome and habitual stimulus–response processes that are supported by anatomically distinct brain systems. Based on previous findings showing that stress modulates the interaction of “cognitive” and “habit” memory systems, we asked in the presented study whether stress may coordinate goal-directed and habit processes in instrumental learning. For this purpose, participants were exposed to stress (socially evaluated cold pressor test) or a control condition before they were trained to perform two instrumental actions that were associated with two distinct food outcomes. After training, one of these food outcomes was selectively devalued as subjects were saturated with that food. Next, subjects were presented the two instrumental actions in extinction. Stress before training in the instrumental task rendered participants’ behavior insensitive to the change in the value of the food outcomes, that is stress led to habit performance. Moreover, stress reduced subjects’ explicit knowledge of the action–outcome contingencies. These results demonstrate for the first time that stress promotes habits at the expense of goal-directed performance in humans.

Introduction

The capacity to predict and control the consequences of one’s own behavior is critical for a successful adaptation to changing environments. The process by which individuals learn which behavior leads to a specific consequence is referred to as instrumental learning. Instrumental behavior is controlled by two systems: a goal-directed system that learns action–outcome associations and a stimulus–response (S–R) or habit system (Dickinson, 1985). During early stages of learning, behavior is mainly goal directed, i.e., it is controlled by the contingency of action and outcome. As training proceeds, however, behavior becomes more and more guided by the triggering stimulus and independent of the outcome, i.e., it becomes habitual (Adams, 1982; Balleine and Dickinson, 1991). In rats, lesions of the medial prefrontal cortex, the dorsomedial striatum, or the mediodorsal thalamus resulted in behavior that was independent of the value of a goal, even after a few training trials (Balleine and Dickinson, 1998; Corbit et al., 2003; Yin et al., 2005). Conversely, lesions of the dorsolateral striatum prevented the formation of habits even after extensive training (Yin et al., 2004, 2005). Corroborating this dissociation, neuropsychological and neuroimaging studies in humans indicated that goal-directed learning is mediated by the prefrontal cortex, whereas habit learning relies on an intact striatum (Knowlton et al., 1996; Valentin et al., 2007).

Converging lines of evidence show that stress and the glucocorticoid stress hormones (mainly cortisol in humans) released from the adrenal cortex can operate as a switch between “cognitive” and “habit” learning systems. Stress before training in

a task that could be solved by hippocampus-dependent spatial (cognitive) and striatum-dependent S–R (habit) systems favored habit over cognitive learning in both rodents and man (Kim et al., 2001; Schwabe et al., 2007). Similar effects occurred after chronic stress or pharmacological manipulation of stress hormone levels (Packard and Wingard, 2004; Schwabe et al., 2008a, 2009a,b). Here, we test the hypothesis that the use of the two systems involved in instrumental learning is also modulated by stress, in a manner that facilitates habit performance, at the expense of goal-directed learning.

To this end, we exposed subjects to stress (or a control condition) before they were trained in two actions leading to two distinct food outcomes. We used a partial reinforcement schedule, in which an action led with a certain probability to the corresponding outcome, because this results in more persistent behavior than continuous reinforcement (Hull, 1943). After training, we devalued selectively one of the two food outcomes by inviting the subjects to eat that food to satiety (Balleine and Dickinson, 1998). Then, participants performed the two actions in extinction. A recent functional magnetic resonance imaging study showed that goal-directed and habit learning in this paradigm rely on the medial prefrontal cortex and the caudate nucleus, respectively (Valentin et al., 2007). Goal-directed behavior is expressed by a decrease in the frequency of the action associated with the devalued outcome, i.e., the food eaten to satiety. If stress favors habit learning, we would expect that the behavior of stressed subjects is insensitive to the change in the value of the outcomes.

Materials and Methods

Eighty healthy, normal weight students of the Ruhr University Bochum participated in this experiment (40 women, 40 men; age, 23.6 ± 0.4 years, mean \pm SEM; body mass index, 22.3 ± 0.3 kg/m², mean \pm SEM). Exclusion criteria were checked in a standardized interview and comprised any current or chronic mental or physical disorders, any food intolerance, as well as a current or planned diet. Smokers as well as

Received Feb. 27, 2009; revised April 24, 2009; accepted April 27, 2009.

This work was supported by Deutsche Forschungsgemeinschaft Grant SCHW1357/2-1. We gratefully acknowledge the assistance of Florian Watzlawik and Karla Luecking 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.0979-09.2009

Copyright © 2009 Society for Neuroscience 0270-6474/09/297191-08\$15.00/0

women taking oral contraceptives were excluded from participation because nicotine and oral contraceptives change the neuroendocrine stress response (Kirschbaum et al., 1999; Mendelson et al., 2005). Furthermore, we prescreened participants to ensure that they find the presented foods (chocolate milk, chocolate pudding, oranges, orange juice, and peppermint tea) pleasant. Nevertheless, 13 subjects had to be excluded from additional analyses because they revealed during the experiment that they disliked at least one of the foods [pleasantness rating below 10 on a scale from 0 (“not pleasant”) to 100 (“very pleasant”) and choosing the high-probability action <20% of the time].

Subjects were asked to refrain from caffeine and physical exercise within the 6 h before testing and to fast for at least 3 h before the experiment started. All participants provided written informed consent for their participation in the protocol as approved by the ethics committee of the German Psychological Society.

Stress protocol. Participants in the stress condition (18 men, 16 women) were exposed to the socially evaluated cold pressor test (SECPT) as described in detail previously (Schwabe et al., 2008b). Briefly, they immersed their right hand up to and including the wrist for 3 min (or until they could no longer tolerate it) into ice water (0–2°C). During hand immersion, they were videotaped and monitored by an unfamiliar person. Participants in the control condition (18 men, 15 women) submerged their right hand up to and including the wrist for 3 min in warm water (35–37°C); they were neither videotaped nor monitored by an unfamiliar person. To assess whether the stress induction by the SECPT was successful, subjective stress ratings, blood pressure, and salivary cortisol were measured.

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

Blood pressure. Blood pressure was measured for 5 min before, for 3 min during, and again for 5 min after the SECPT or control condition using the Dinamap system (Critikon) with the cuff placed on the left upper arm.

Saliva sampling and cortisol analysis. Participants collected saliva samples before as well as 1, 20, and 50 min after the SECPT or control condition with a Salivette collection device (Sarstedt). Saliva samples were kept at –20°C until analysis. Free cortisol concentrations were measured using an immunoassay (IBL). Interassay and intra-assay coefficients of variance were below 10%.

Instrumental learning task. We used a modification of a task introduced by Valentin et al. (2007); the task was created with the help of the Biopsychology toolbox (Rose et al., 2008). In this task, three trial types were presented: chocolate, orange, and neutral. On each trial, participants had to choose between two actions represented by two distinct symbols (Fig. 1). According to the reward schedule associated with the chosen action, 1 ml of a liquid was delivered or else no liquid was delivered. 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 like a straw between the lips. Importantly, the two actions per trial type differed in the probability with which a food outcome was delivered. Although 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

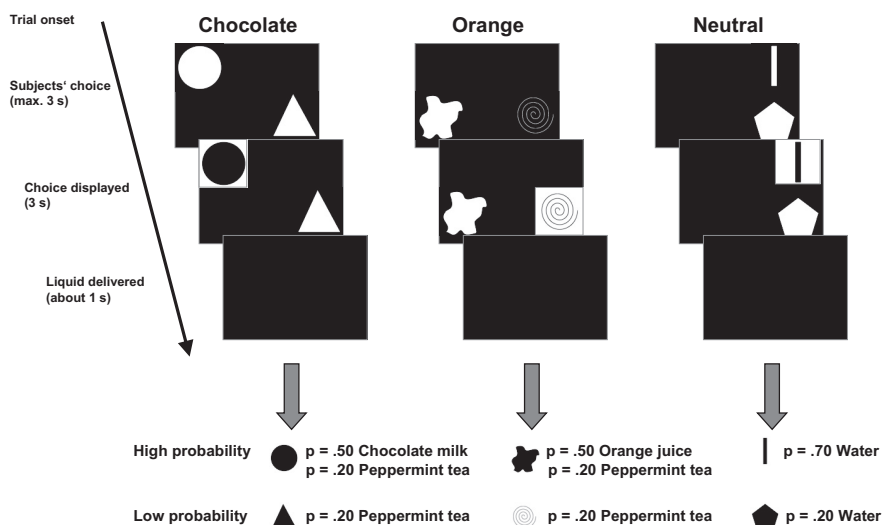


Figure 1. The instrumental learning task (modified from Valentin et al., 2007). Participants completed three trial types (chocolate, orange, and neutral). On each trial, they were asked to choose between two actions represented by unique symbols. In each trial type, there was one action that led with 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 delivered chocolate milk and orange juice, respectively, with a probability of $p = 0.50$, a common liquid (peppermint tea) with a probability of $p = 0.20$, or nothing. The low probability action yielded the common outcome with a probability of $p = 0.20$. When an action was chosen, the related symbol was highlighted for 3 s before the outcome was delivered.

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, with a probability of either $p = 0.70$ (high probability action) or $p = 0.20$ (low probability action). This neutral condition served as a control to assess the effect of the rewards (chocolate milk and orange juice) on participants' choice behavior.

Subjects selected an action by moving the cursor to this symbol and pressing the left mouse button. The referring symbol was highlighted for 3 s and the food outcome delivered (depending on the chosen action and its outcome probability). Then, the screen was cleared and the next trial was started. Participants completed 75 trials in each of the three trial types (chocolate, orange, and neutral) whose occurrence was randomized, resulting in 225 trials in total (intertrial interval, 8 s; total processing time, ~30 min).

Outcome devaluation. After training in the instrumental task, participants were invited to eat either oranges or chocolate pudding until they did not want it anymore (selective satiation). This procedure served to decrease the value of one outcome (e.g., when a subject was satiated with oranges, the value of the orange juice should be decreased), while the value of the other outcome (chocolate milk in the example) should remain high. Which specific food was used for devaluation (oranges or chocolate pudding) was fully counterbalanced across subjects.

Extinction test. After the outcome devaluation, participants were again presented 75 trials of each of the three trial types in random order (intertrial interval, 8 s) and asked to choose between the actions that led to different food outcomes at training. Same as during training, the symbol representing the chosen action was highlighted. This time, however, the rewards (chocolate milk and orange juice) were never delivered, i.e., subjects 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. This extinction procedure ensured that the subjects only use information about the value of the outcome by making use of the previously learned associations between that outcome and a particular action.

A decrease in the choice of the action associated with the devalued food outcome indicated goal-directed performance, whereas the ongoing

choice of the action associated with the devalued food outcome was interpreted as indicative for habit performance.

Procedure. All testing took place between 1:00 P.M. and 5:30 P.M. to control for the diurnal rhythm of cortisol. After subject's arrival at the laboratory, blood pressure measurements were taken and a first saliva sample was collected. Then, subjects were exposed either to the SECPT or a control condition. Immediately thereafter, subjective assessments of the previous situation and another saliva sample were collected and blood pressure was measured again. Twenty minutes after the cessation of the SECPT/control condition, participants collected another saliva sample and started then with the experimental task. This interval between the SECPT/control condition and the instrumental learning task was chosen because cortisol reaches peak levels in response to the SECPT after 20–30 min (Schwabe et al., 2008b). First, subjective ratings of hunger (0, "not hungry" to 100, "very hungry") and pleasantness of the food outcomes (0, "not pleasant" to 100, "very pleasant") were collected. Next, participants completed 225 trials of the instrumental learning task as described above. Afterward, they rated their hunger and pleasantness of the food outcomes again; another saliva sample was collected (~50 min after stress). Then, they were allowed to eat either chocolate pudding or oranges to satiety. This outcome devaluation served to devalue one of the outcomes associated with a particular action but left the value of the other outcome intact. Subjective ratings of hunger and pleasantness of the food outcomes were collected before the start of the extinction test session. During this session, participants were presented the same trials with the same symbols. They were again asked to choose between the two actions, but neither the devalued nor the nondevalued outcome was presented again (i.e., subjects were tested in extinction for these outcomes).

Finally, subjects were asked in a brief, standardized interview to name which symbol (i.e., which action) was associated with which food outcome in the three trial types. They were requested to describe verbally which symbol had to be selected to receive chocolate milk, orange juice, and water, respectively.

Statistical analyses. Data were analyzed by means of mixed-design ANOVAs, χ^2 tests, paired *t* tests, and *t* tests for independent samples. Salivary cortisol data were missing for 10 participants (four controls) because these participants provided not enough saliva for the biochemical analysis. *p* values were Bonferroni's corrected when indicated. All reported *p* values are two tailed.

Results

Subjective and physiological responses to stress

Participants' subjective stress ratings, blood pressure, and salivary cortisol responses verified the success of the stress-induction by the SECPT.

All but six subjects of the stress group (four women, two men; mean duration, 82 s; range, 50–150 s) immersed their hand for the full 3 min in the ice water. These six subjects did not differ in their subjective or physiological stress responses from the rest of the stress group (all *p* > 0.30).

Subjective stress ratings

As expected and shown in Table 1, participants in the stress condition experienced the hand immersion as significantly more stressful, painful, and unpleasant than participants in the control condition (all $F_{(1,63)} > 30$; all *p* < 0.001). Men and women were comparable in their evaluation of the hand immersion (all *p* > 0.23).

Blood pressure

The exposure to the SECPT elicited a significant increase in systolic and diastolic blood pressure (treatment, both $F_{(1,63)} > 9.5$; both *p* < 0.01). As can be seen in Table 1, groups had comparable blood pressure before and after hand immersion, whereas stressed participants had higher blood pressure during hand immersion (time \times treatment, both $F_{(2,126)} > 55$; both *p* < 0.001).

Table 1. Subjective stress ratings and blood pressure values before, during, and after the SECPT or control condition

	Control	Stress
Subjective assessments		
Stressfulness	5.8 (2.4)	37.4 (5.0)
Painfulness	0.9 (0.5)	58.2 (4.0)
Unpleasantness	6.1 (2.1)	52.6 (3.9)
Systolic blood pressure (mmHg)		
Before hand immersion	118.5 (2.6)	119.4 (2.5)
During hand immersion	114.9 (2.5)	133.2 (2.4)
After hand immersion	112.3 (1.9)	115.2 (2.4)
Diastolic blood pressure (mmHg)		
Before hand immersion	65.8 (1.4)	66.9 (1.4)
During hand immersion	66.1 (1.3)	81.0 (1.6)
After hand immersion	63.3 (1.3)	65.5 (1.5)

Stressfulness, painfulness, and unpleasantness were rated on a scale from 0 ("not at all") to 100 ("very much"). Bold indicates significant group difference (*p* < 0.01). Data represent means \pm SEMs are given in parentheses.

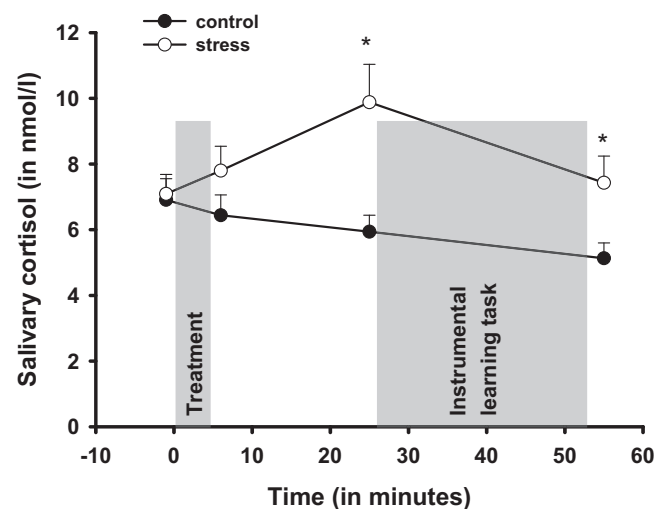


Figure 2. Salivary cortisol response (in nanomoles per liter) to the stress and control condition. The gray bars denote the timing and duration of the treatment (stress vs control condition) and the instrumental learning task, respectively. Note that the learning task was presented during the high cortisol period of the stress group. Data represent means \pm SEM. **p* < 0.05 (corrected), significant group difference.

Overall, men tended to have higher systolic and diastolic blood pressure than women (sex, both $F_{(1,63)} > 2.4$; both *p* < 0.12), but they did not differ in the blood pressure response to the SECPT (treatment \times sex, both $F_{(1,63)} < 1$; both *p* > 0.80).

Cortisol

As shown in Figure 2, the SECPT caused a significant increase in cortisol, whereas the control condition did not (treatment, $F_{(1,53)} = 7.8$, *p* < 0.01; time \times treatment, $F_{(3,159)} = 2.9$, *p* < 0.05). Stressed participants and controls did not differ in their cortisol concentration at baseline and immediately after the treatment but 20 and 50 min after cessation of the SECPT or control condition. Participants learned the instrumental actions when cortisol concentrations were high in the stress group. There was no effect of sex on the cortisol concentration, nor was there an interaction between participants' sex and the treatment (both $F < 1.6$; both *p* > 0.20).

Effects of stress on instrumental learning

Inspection of individual data revealed a subgroup of seven subjects who showed no increase in the choice of the high probability action in the chocolate and orange trials (supplemental Fig. S1,

available at www.jneurosci.org as supplemental material), although they preferred the rewards (chocolate milk and orange juice) over the common outcome ($F_{(2,8)} = 6.1$; $p = 0.02$). None of these seven subjects could name the action–outcome association for any of the three trial types. Thus, these subjects were classified as “nonlearners” and excluded from the following analyses. Interestingly, five of the seven nonlearners were stressed before training ($\chi^2_{(1)} = 1.4$; $p = 0.24$). Although not statistically significant, this might be interpreted as first evidence that stress impedes instrumental learning.

For the remaining 60 participants, Figure 3 shows the percentage of high probability choices associated with the nondevalued, the subsequently devalued, and the neutral outcome over training (whether chocolate milk or orange juice was devalued was counterbalanced across subjects). As training proceeded, all participants, regardless of the stress or control group, favored increasingly the high probability actions associated with the rewards (chocolate milk and orange juice) over their low probability counterparts. This indicates that subjects learned to choose the instrumental action for both the outcome that was devalued later on and the nondevalued outcome. In contrast, participants did not learn to choose the high probability action more often than the low probability action in the neutral trials, suggesting that participants were rather indifferent as to whether they received the effectively neutral control liquid or not. Accordingly, a mixed-design ANOVA with value (neutral, later devalued, and nondevalued outcome trials) and time (five blocks with 15 trials per block) as within-subjects factors and treatment (SECPT vs control condition) and sex (men vs women) as between-subjects factors revealed significant main effects of value ($F_{(2,112)} = 34.6$; $p < 0.001$) and time ($F_{(4,448)} = 20.9$; $p < 0.001$) as well as a significant time \times value interaction ($F_{(8,448)} = 5.0$; $p < 0.001$). Importantly, there was no effect of treatment, indicating that learning curves of stressed and control subjects were comparable, nor did participants' sex have an effect on instrumental learning performance (all $F_{(1,56)} < 1$; all $p > 0.80$).

Effects of selective outcome devaluation on subjective hunger and pleasantness ratings

The selective satiation (devaluation) procedure led to a significant reduction in subjective hunger ratings ($F_{(1,58)} = 160.3$; $p < 0.001$). On average, hunger ratings dropped from 64 ± 2.9 (mean \pm SEM) before the devaluation to 35 ± 2.6 after satiety. The subjective pleasantness ratings as displayed in Figure 4 show that the devaluation was indeed specific to the food eaten to

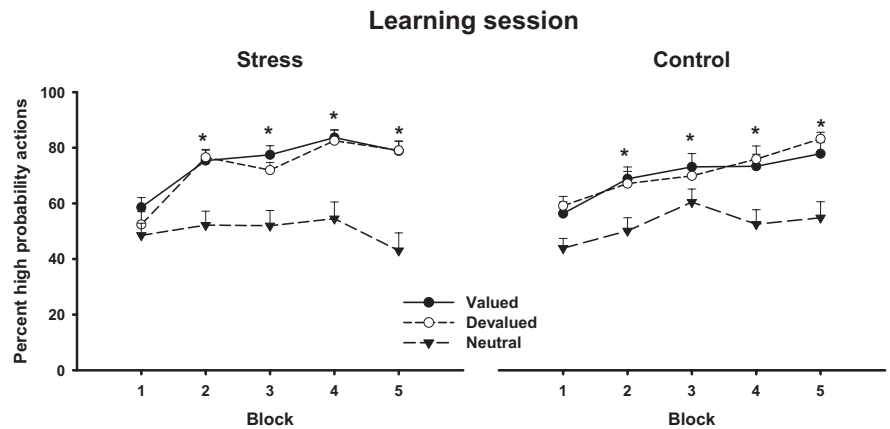


Figure 3. Percentage of high probability choices across the learning session (1 block = 15 trials). Both stressed and control subjects favored increasingly the high probability actions associated with the subsequently valued and devalued outcomes (i.e., chocolate pudding and orange juice) over the corresponding low probability actions ($*p < 0.05$, corrected). No such preference was observed in the neutral outcome trials, in which participants were indifferent between the high and low probability actions. Data represent means \pm SEM.

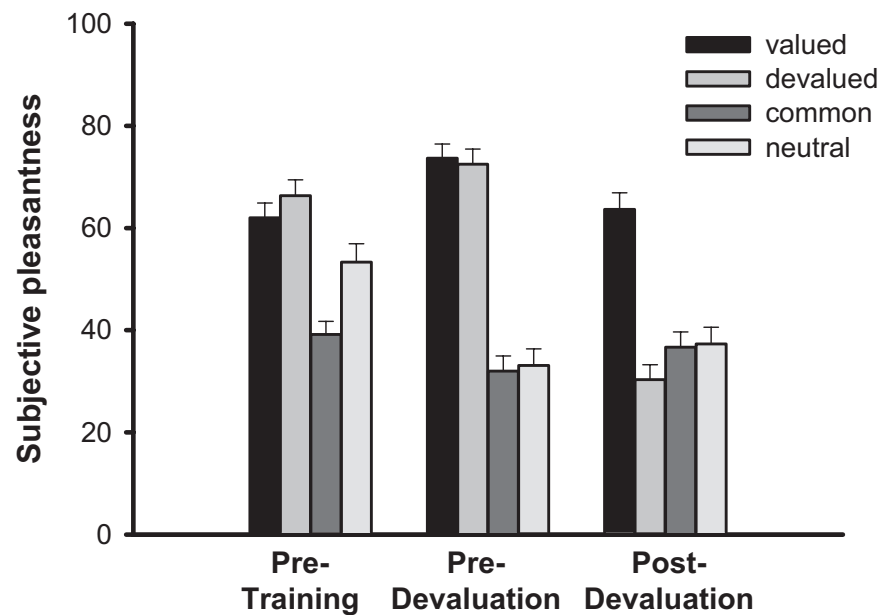


Figure 4. Subjective ratings of the pleasantness (0, not pleasant; 100, very pleasant) of the food outcomes before training as well as before and after the selective outcome devaluation. Subjects initially preferred the rewards (chocolate milk and orange juice) over the common and neutral liquids (peppermint tea and water, respectively). The selective satiation decreased the pleasantness rating for the eaten (devalued) food significantly compared with the food not eaten (valued). Data represent means \pm SEM.

satiety. The subjective pleasantness of the food eaten to satiety decreased sharply, whereas no such decrease was observed for the foods not eaten. This interpretation is supported by a mixed-design ANOVA showing a significant time (before vs after devaluation) \times value (devalued vs nondevalued) interaction effect ($F_{(1,56)} = 70.0$; $p < 0.001$). It is important to note that this pattern was affected by neither stress nor participants' sex (main and possible interaction effects, all $F < 2.5$; all $p > 0.14$).

Effects of outcome devaluation and stress on instrumental responses in the extinction test

The instrumental responses in the extinction test allowed assessing whether performance was goal directed or habitual. Choosing the high probability action associated with the devalued outcome

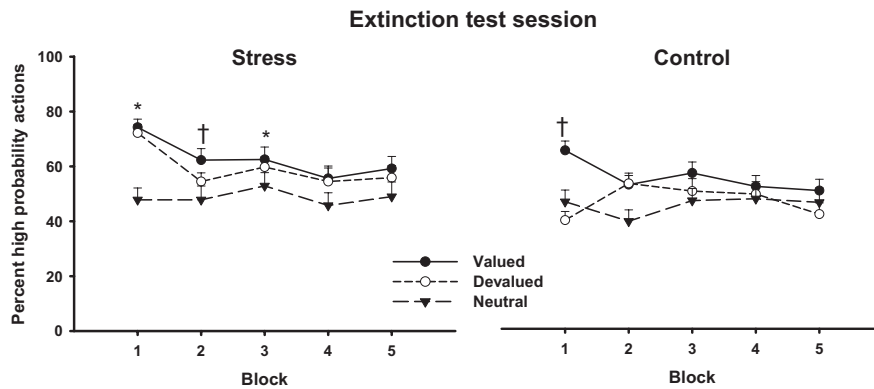


Figure 5. Percentage of high probability choices across the extinction test. Participants in the control group favored the valued high probability action but not the devalued high probability action in the first 15-trial block, suggesting that they altered their choice behavior as a function of the change in the value of the outcomes. In contrast, stressed subjects still favored both the valued and the devalued high probability action over the corresponding low probability actions, indicating habitual performance. Note that stressed subjects favored the valued and devalued high probability actions even in the third 15-trial block, despite that the valued and devalued outcomes were no longer presented during the extinction test session. * $p < 0.05$ (corrected), valued and devalued high probability actions favored over the corresponding low probability actions; † $p < 0.05$ (corrected), valued high probability action favored over its low probability counterpart; data represent means \pm SEM.

less often than the one associated with the valued (i.e., nondevalued) outcome indicated goal-directed learning (Valentin et al., 2007), whereas still favoring the high probability action associated with the devalued outcome (as much as the high probability action associated with the valued outcome) over its low probability counterpart indicated habit learning.

Participants in the control condition chose the valued high probability action significantly more frequently than the devalued high probability action across the extinction test trials ($F_{(1,30)} = 5.7$; $p = 0.02$). As shown in Figure 5, they still preferred the valued high probability action in the first 15-trial block ($t_{(30)} = 4.4$; $p < 0.001$), before they had the chance to learn that the valued outcome was no longer presented. On the contrary, they did not favor the devalued high probability action but even seemed to avoid the devalued outcome in the first 15-trial block, as reflected in a more frequent choice of the low probability action ($t_{(30)} = 3.3$; $p = 0.01$) (Fig. 5). In the remaining trials, the participants chose the low and high probability actions in all trial types at random, which suggests successful extinction learning.

Participants that were exposed to stress before learning showed a markedly distinct choice pattern (treatment \times time \times value interaction, $F_{(4,224)} = 5.5$; $p < 0.001$). They chose the devalued high probability action as often as the valued high probability action across the extinction test trials ($F_{(1,28)} = 1.3$; $p = 0.27$). Stressed subjects chose the high probability action associated with the devalued outcome significantly more often than the corresponding low probability action in the first and in the third 15-trial block (both $t_{(29)} > 2.6$; both $p < 0.05$, Bonferroni's corrected). Moreover, they still favored the valued high-probability action in the third 15-trial block, i.e., they continued to choose the valued high probability action that had not been associated with the valued outcome for >30 trials (blocks 1–3, all $t_{(29)} > 2.8$; all $p < 0.05$, Bonferroni's corrected) (Fig. 5).

The difference between the stress and control groups was most pronounced in the first 15-trial block of the extinction test. Therefore, we compared the change in their performance from the last training block to the first extinction test block. A mixed-design ANOVA with time (last 15 training trials vs first 15 extinction test trials) and value (valued vs devalued) as within-subjects factors and treatment as between-subjects factor yielded a signif-

icant three-way interaction ($F_{(1,56)} = 13.7$; $p < 0.001$), indicating decreased responding to the devalued high probability action after selective outcome devaluation in controls ($F_{(1,30)} = 24.9$; $p < 0.001$) but not in stressed participants ($F_{(1,28)} = 0.29$; $p = 0.59$) (Fig. 6). This underlines that participants in the control group performed goal directed, whereas participants in the stress group showed habit performance. There was no sex difference in the performance during the test session, nor did participants' sex interact with the treatment (all $F < 1.5$; all $p > 0.20$).

Effects of stress and outcome devaluation on reaction times

Mixed-design ANOVAs with the within-subjects factors time (five 15-trial blocks) and value (valued and devalued) as well as the between-subjects factors treatment (SECPT vs control) and sex (men vs women) on the reaction times in the train-

ing and test sessions revealed significant main effects of time (both $F_{(4,224)} > 4.3$; both $p < 0.01$). Participants responded increasingly faster with time during both the training and extinction test sessions. Men tended to respond faster than women during learning ($F_{(1,56)} = 3.1$; $p = 0.09$). We obtained no effect of the treatment or value, suggesting that reaction times were not affected by these factors (all $F < 1.2$; all $p > 0.29$).

Effects of stress on the awareness of action–outcome associations

Stress before learning had a detrimental effect on subjects' awareness of the action–outcome associations. Fifty-eight percent (18 of 31) of the controls but only 28% (8 of 29) of the stressed subjects could name the action–outcome associations in the three trial types correctly ($\chi^2_{(1)} = 5.7$; $p = 0.017$). The mean \pm SEM number of correctly named action–outcome associations was 2.5 ± 0.1 in the control group and 1.7 ± 0.2 in the stress group ($t_{(58)} = 3.7$; $p = 0.001$).

Interestingly, the number of correctly named action–outcome associations was negatively correlated with the percentage of devalued high probability choices in the first ($r = -0.31$; $p = 0.018$) and third ($r = -0.28$; $p = 0.034$) blocks of the extinction test. That is, reduced awareness of the action–outcome associations was associated with more habitual performance.

Discussion

This study examined the impact of stress on the coordination of habit and goal-directed instrumental learning in humans using a behavioral measure of habit formation that was previously used mainly in rodents. Overall, our findings provide strong evidence that stress favors habit performance, at the expense of goal-directed performance. In contrast to nonstressed controls, subjects that were exposed to stress continued to perform the action associated with a particular outcome after this outcome had been devalued. Moreover, stressed subjects stuck significantly longer to the acquired responses than controls. Interestingly, the effect of stress was not restricted to behavioral persistence but became also apparent in a reduced explicit knowledge of the action–outcome associations. The reduced awareness of action–outcome associations was associated with more habitual performance.

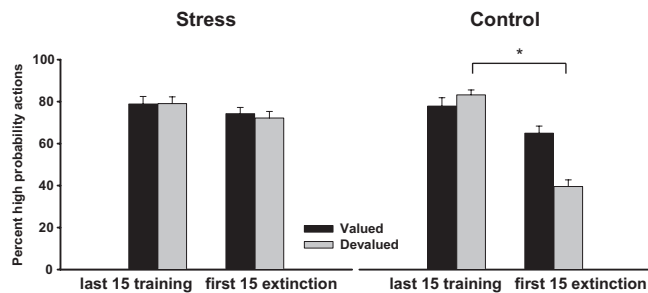


Figure 6. Comparison of valued and devalued high probability choices in the last 15-trial training block and the first 15-trial testing block after the selective satiation procedure. Although control subjects showed a significant decrease in the number of high probability actions associated with the devalued outcome ($*p < 0.01$, corrected), no such decrease was found in stressed subjects. Data represent means \pm SEM.

At a neural level, there is convincing evidence that goal-directed learning is mediated by prefrontal cortex areas (Corbit and Balleine, 2003; Dalley et al., 2004; Matsumoto and Tanaka, 2004; Valentin et al., 2007). The prefrontal cortex is characterized by a high density of glucocorticoid receptors, suggesting a high sensitivity to stress (Reul and de Kloet, 1985; McEwen et al., 1986). Indeed, electrophysiological studies show that stress reduces synaptic long-term potentiation in the prefrontal cortex (Maroun and Richter-Levin, 2003; Cerqueira et al., 2007; Diamond et al., 2007). These deficits in neuroplasticity are paralleled by impairments in prefrontal cortex-dependent memory functions (Lupien et al., 1999; Roozendaal et al., 2004; Schoofs et al., 2008). Moreover, other signaling pathways activated by stress, including the dopamine and noradrenaline systems, have been shown to induce prefrontal cortex impairments (Brennan and Arnsten, 2008). In the light of these findings, it could be argued that the stress-induced facilitation of habit performance we found in the present study is primarily attributable to impaired goal-directed learning. Given that learning is initially dependent on goal-directed processes whereas habit processes take over control as learning proceeds (Adams, 1982; Balleine and Dickinson, 1991; Dickinson et al., 1995), performance should be impaired early during training in stressed subjects if the beneficial effect of stress on habit learning is attributable to impaired goal-directed learning. We found no effect of stress on the learning curves. However, there was some very first (attributable to the small number of nonlearners statistically not significant) evidence that stress might have a negative influence on the acquisition of the instrumental task, which would be consistent with the suggested deficit in goal-directed processes guiding early learning. This picture, however, is complicated by two issues. First, although it appears to be widely accepted that the transition from goal-directed to habit learning can occur with overtraining, there is also evidence for intact sensitivity to outcome devaluation even after extensive training (Colwill and Rescorla, 1985). Second and maybe even more important, there is considerable evidence that goal-directed and habit learning processes depend not solely on the prefrontal cortex and dorsolateral striatum, respectively, but rather on networks of different neuronal structures. In rats, lesions of the mediodorsal thalamus render instrumental behavior insensitive to changes in outcome value (Corbit et al., 2003). Furthermore, goal-directed actions necessitate an intact dorsomedial striatum (Yin et al., 2005), and habits are promoted by amphetamine exposure, which leads to reduced spine density in the dorsomedial part of the striatum (Robinson and Kolb, 2004; Nelson and Killcross, 2006). The latter findings suggest a func-

tional heterogeneity within the dorsal striatum, with the dorso-lateral striatum being relevant for habit learning whereas the dorsomedial striatum supports goal-directed learning. A comparable double dissociation has been found in the medial prefrontal cortex in which the prelimbic region has been shown to control goal-directed behavior, whereas the infralimbic region has been suggested to mediate habit learning (Killcross and Coutureau, 2003). Future studies using functional magnetic resonance imaging are clearly needed to unravel the neuronal correlates of the stress-induced promotion of habit performance reported here.

Another brain structure that has been assigned an important role in instrumental learning is the amygdala (Balleine and Killcross, 2006). Lesions of the basolateral nucleus of the amygdala rendered rats' behavior insensitive to changes in the value of an outcome and thus abolished goal-directed performance (Hatfield et al., 1996; Blundell et al., 2001; Balleine et al., 2003). Stress and stress hormones, however, lead to increased amygdala activity rather than to a deactivation of the amygdala (Fallon and Ciofi, 1992; Shepard et al., 2000; van Stegeren et al., 2007). In line with a recent model of amygdala functioning (McGaugh, 2002; Roozendaal et al., 2008), we suggest that the amygdala exerts a modulating influence on other brain systems and coordinates habit and goal-directed behavior via its connections with the prefrontal cortex and striatum, respectively (Smith and Bolam, 1990; Goldstein et al., 1996).

Although responding to the devalued high probability action indicated goal-directed vs habit performance, responding to the valued high probability action provided information about memory extinction. Nonstressed controls showed a decrease in the frequency of high probability actions associated with the valued outcome after they noticed that this was no longer presented, which indicates successful extinction learning. In contrast, stressed subjects favored the valued high probability action in the first 45 trials of the test session, although it was never reinforced by the valued outcome. This is another sign of habitual performance after stress. At the same time, it might suggest reduced extinction learning. Stress effects on extinction learning and habit formation can hardly be disentangled because habits imply persistence. Nevertheless, there is recent evidence that stress hormones impair the extinction of fear memories in mice (Brinks et al., 2009) (for reports of enhanced fear extinction, see Barrett and Gonzalez-Lima, 2004; Yang et al., 2006). Interestingly, these effects were genotype dependent. Whether the genetic background may also account for some of the individual variability in habit formation is a challenge for future research.

Previous studies demonstrated that stress modulates multiple anatomically and functionally distinct memory systems in favor of neostriatum-dependent habit (S–R) learning and at the expense of hippocampus-dependent cognitive (spatial) learning (Kim et al., 2001; Schwabe et al., 2007). In these studies, cognitive memory was conceptualized as a declarative (explicit) system that allows flexible use of knowledge, whereas habit memory was seen as a rather rigid, nondeclarative (at least partly implicit) system. The kinds of instrumental learning investigated in the present study fit well in this terminology. This notion is supported by the fact that the stress-induced shift toward habit performance was accompanied by a significant decrease in explicit knowledge of action–outcome contingencies. The finding that stressed subjects improved over learning although they had relatively poor knowledge of the action–outcome associations is in line with reports indicating that habit learning does not require awareness for what is learned (Bayley et al., 2005). Furthermore, the decrease in explicit knowledge in stressed participants sug-

gests impaired hippocampus- and prefrontal cortex-dependent memory and is consistent with a number of studies showing a reduction in episodic memory after stress (Buchanan et al., 2006; Lupien et al., 2007; Payne et al., 2007; Wolf, 2008). These studies, however, focused on a single memory system and did not control for the use of different learning systems. To date, the effect of stress on the transition between multiple memory systems has been shown solely in the domain of spatial navigation (Kim et al., 2001; Schwabe et al., 2007). The present results indicate that the modulating effect of stress is not limited to one particular domain. Rather, they suggest that stress favors habitual over cognitive learning and memory in general.

It is to be noted that, in the face of the discriminative cues used here and the reduced ability of stressed participants to describe which symbol had to be selected for which outcome, it cannot be ruled out that the performance of control subjects was, at least partly, mediated by stimulus–outcome learning. Another limitation of the present study can be seen in the fact that both the training and the extinction test session were given within 90 min after the stress exposure and cortisol levels were still higher in stressed than in control subjects after training (i.e., before extinction testing). Thus, based on the present study, it cannot be decided whether stress affected the instrumental processes involved in either task acquisition (e.g., attention or initial encoding) or performance (e.g., retrieval processes or response inhibition). These possible effects need to be disentangled in future studies by varying the timing of the stress exposure in the learning process.

To summarize, this study shows that stress promotes habit performance in humans. The present findings provide novel insights into the effects of stress on learning processes and the modulation of multiple memory systems. Furthermore, they may have significant implications for our understanding of the development of compulsive behavior and addiction, which have been related to the aberrant engagement of habitual processes in instrumental behavior (Berke and Hyman, 2000; Everitt et al., 2001; Everitt and Robbins, 2005).

References

- Adams C (1982) Variations in the sensitivity of instrumental responding to reinforcer devaluation. *Q J Exp Psychol* 34B: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 (1998) The role of incentive learning in instrumental outcome revaluation 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, Killcross AS, Dickinson A (2003) The effect of lesions of the basolateral amygdala on instrumental conditioning. *J Neurosci* 23:666–675.
- Barrett D, Gonzalez-Lima F (2004) Behavioral effects of metyrapone on pavlovian extinction. *Neurosci Lett* 371:91–96.
- Bayley PJ, Frascino JC, Squire LR (2005) Robust habit learning in the absence of awareness and independent of the medial temporal lobe. *Nature* 436:550–553.
- Berke JD, Hyman SE (2000) Addiction, dopamine, and the molecular mechanisms of memory. *Neuron* 25:515–532.
- Blundell P, Hall G, Killcross S (2001) Lesions of the basolateral amygdala disrupt selective aspects of reinforcer representation in rats. *J Neurosci* 21:9018–9026.
- Brennan AR, Arnsten AF (2008) Neuronal mechanisms underlying attention deficit hyperactivity disorder: the influence of arousal on prefrontal cortical function. *Ann N Y Acad Sci* 1129:236–245.
- Brinks V, de Kloet ER, Oitzl MS (2009) Corticosterone facilitates extinction of fear memory in BALB/c mice but strengthens cue related fear in C57BL/6 mice. *Exp Neurol* 216:375–382.
- Buchanan TW, Tranel D, Adolphs R (2006) Impaired memory retrieval correlates with individual differences in cortisol response but not autonomic response. *Learn Mem* 13:382–387.
- Cerqueira JJ, Mailliet F, Almeida OF, Jay TM, Sousa N (2007) The prefrontal cortex as a key target of the maladaptive response to stress. *J Neurosci* 27:2781–2787.
- Colwill RM, Rescorla RA (1985) Postconditioning devaluation of a reinforcer affects instrumental responding. *J Exp Psychol Anim Behav Process* 11:120–132.
- Corbit LH, Balleine BW (2003) The role of prefrontal cortex in instrumental conditioning. *Behav Brain Res* 146:145–157.
- 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.
- Dalley JW, Cardinal RN, Robbins TW (2004) Prefrontal executive and cognitive functions in rodents: neural and neurochemical substrates. *Neurosci Biobehav Rev* 28:771–784.
- Diamond DM, Campbell AM, Park CR, Halonen J, Zoladz PR (2007) The temporal dynamics model of emotional memory processing: a synthesis on the neurobiological basis of stress-induced amnesia, flashbulb and traumatic memories, and the Yerkes-Dodson law. *Neural Plasticity* 2007:60803.
- Dickinson A (1985) Actions and habits: the development of behavioral autonomy. *Philos Trans R Soc Lond B Biol Sci* 308:67–78.
- Dickinson A, Balleine BW, Watt A, Gonzalez F, Boakes RA (1995) Motivational control after extended instrumental training. *Anim Learn Behav* 23:197–206.
- Everitt BJ, Robbins TW (2005) Neural systems of reinforcement for drug addiction: from actions to habits to compulsion. *Nat Neurosci* 8:1481–1489.
- Everitt BJ, Dickinson A, Robbins TW (2001) The neuropsychological basis of addictive behaviour. *Brain Res Brain Res Rev* 36:129–138.
- Fallon JL, Ciofi P (1992) Distribution of monoamines within the amygdala. In: *The amygdala: neurobiological aspects of emotion, memory, and mental dysfunction* (Aggleton J, ed), pp 94–114. New York: Wiley.
- Goldstein LE, Rasmusson AM, Bunney BS, Roth RH (1996) Role of the amygdala in the coordination of behavioral, neuroendocrine, and prefrontal cortical monoamine responses to psychological stress in the rat. *J Neurosci* 16:4787–4798.
- Hatfield T, Han JS, Conley M, Gallagher M, Holland P (1996) Neurotoxic lesions of basolateral, but not central, amygdala interfere with Pavlovian second-order conditioning and reinforcer devaluation effects. *J Neurosci* 16:5256–5265.
- Hull CL (1943) *Principles of behavior*. New York: Appleton-Century-Crofts.
- Killcross S, Coutureau E (2003) Coordination of actions and habits in the medial prefrontal cortex of rats. *Cereb Cortex* 13:400–408.
- Kim JJ, Lee HJ, Han JS, Packard MG (2001) Amygdala is critical for stress-induced modulation of hippocampal long-term potentiation and learning. *J Neurosci* 21:5222–5228.
- Kirschbaum C, Kudielka BM, Gaab J, Schommer NC, Hellhammer DH (1999) Impact of gender, menstrual cycle phase, and oral contraceptives on the activity of the hypothalamus-pituitary-adrenal axis. *Psychosom Med* 61:154–162.
- Knowlton BJ, Mangels JA, Squire LR (1996) A neostriatal habit learning system in humans. *Science* 273:1399–1402.
- Lupien SJ, Gillin CJ, Hauger RL (1999) Working memory is more sensitive than declarative memory to the acute effects of corticosteroids: a dose-response study in humans. *Behav Neurosci* 113:420–430.
- Lupien SJ, Maheu F, Tu M, Fiocco A, Schramek TE (2007) The effects of stress and stress hormones on human cognition: implications for the field of brain and cognition. *Brain Cogn* 65:209–237.
- Maroun M, Richter-Levin G (2003) Exposure to acute stress blocks the induction of long-term potentiation of the amygdala-prefrontal cortex pathway *in vivo*. *J Neurosci* 23:4406–4409.
- Matsumoto K, Tanaka K (2004) The role of the medial prefrontal cortex in achieving goals. *Curr Opin Neurobiol* 14:178–185.
- McEwen BS, De Kloet ER, Rostene W (1986) Adrenal steroid receptors and actions in the nervous system. *Physiol Rev* 66:1121–1188.
- McGaugh JL (2002) Memory consolidation and amygdala: a systems perspective. *Trends Neurosci* 25:456–461.
- Mendelson JH, Sholar MB, Goletiani N, Siegel AJ, Mello NK (2005) Effects

- of low- and high-nicotine cigarette smoking on mood states and the HPA axis in men. *Neuropsychopharmacology* 30:1751–1763.
- Nelson A, Killcross S (2006) Amphetamine exposure enhances habit formation. *J Neurosci* 26:3805–3812.
- Packard MG, Wingard JC (2004) Amygdala and “emotional” modulation of the relative use of multiple memory systems. *Neurobiol Learn Mem* 82:243–252.
- Payne JD, Jackson ED, Hoscheidt S, Ryan L, Jacobs WJ, Nadel L (2007) Stress administered prior to encoding impairs neutral but enhances emotional long-term memories. *Learn Mem* 14:861–868.
- Reul JM, de Kloet ER (1985) Two receptor systems for corticosterone in rat brain: microdistribution and differential occupation. *Endocrinology* 117:2505–2511.
- Robinson TE, Kolb B (2004) Structural plasticity associated with exposure to drugs of abuse. *Neuropsychopharmacology* 47 [Suppl 1]:33–46.
- Roosendaal B, McReynolds JR, McGaugh JL (2004) The basolateral amygdala interacts with the medial prefrontal cortex in regulating glucocorticoid effects on working memory impairment. *J Neurosci* 24:1385–1392.
- Roosendaal B, Barsegyan A, Lee S (2008) Adrenal stress hormones, amygdala activation, and memory for emotionally arousing experiences. In: *Progress in brain research* (de Kloet ER, Oitzl MS, Vermetten E, eds). Amsterdam: Elsevier.
- Rose J, Otto T, Dittrich L (2008) The biopsychology-toolbox: a free, open-source Matlab-toolbox for the control of behavioral experiments. *J Neurosci Methods* 175:104–107.
- Schoofs D, Preuss D, Wolf OT (2008) Psychosocial stress induces working memory impairments in an n-back paradigm. *Psychoneuroendocrinology* 33:643–653.
- Schwabe L, Oitzl MS, Philippson C, Richter S, Bohringer A, Wippich W, Schachinger H (2007) Stress modulates the use of spatial and stimulus-response learning strategies in humans. *Learn Mem* 14:109–116.
- Schwabe L, Dalm S, Schächinger H, Oitzl MS (2008a) Chronic stress modulates the use of spatial and stimulus-response learning strategies in mice and man. *Neurobiol Learn Mem* 90:495–503.
- Schwabe L, Haddad L, Schachinger H (2008b) HPA axis activation by a socially evaluated cold pressor test. *Psychoneuroendocrinology* 33:890–895.
- Schwabe L, Oitzl MS, Richter S, Schächinger H (2009a) Modulation of spatial and stimulus-response learning strategies by exogenous cortisol in healthy young women. *Psychoneuroendocrinology* 34:358–366.
- Schwabe L, Schächinger H, de Kloet ER, Oitzl MS (2009b) Corticosteroids operate as switch between memory systems. *J Cogn Neurosci*. Advance online publication. Retrieved May 19, 2009. doi: 10.1162/jocn.2009.21278.
- Shepard JD, Barron KW, Myers DA (2000) Corticosterone delivery to the amygdala increases corticotropin-releasing factor mRNA in the central amygdaloid nucleus and anxiety-like behavior. *Brain Res* 861:288–295.
- Smith AD, Bolam JP (1990) The neural network of the basal ganglia as revealed by the study of synaptic connections of identified neurons. *Trends Neurosci* 13:259–265.
- 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.
- van Stegeren AH, Wolf OT, Everaerd W, Scheltens P, Barkhof F, Rombouts SA (2007) Endogenous cortisol level interacts with noradrenergic activation in the human amygdala. *Neurobiol Learn Mem* 87:57–66.
- Wolf OT (2008) The influence of stress hormones on emotional memory: relevance for psychopathology. *Acta Psychol (Amst)* 127:513–531.
- Yang YL, Chao PK, Lu KT (2006) Systemic and intra-amygdala administration of glucocorticoid agonist and antagonist modulate extinction of conditioned fear. *Neuropsychopharmacology* 31:912–924.
- 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.