

Neural mechanisms and computations underlying stress effects on learning and memory

Gediminas Luksys^{1,2} and Carmen Sandi³

Stress has complex effects on memory function that can vary depending on the type of information that is learned and in relation to inter-individual characteristics. Recent work has also shown that stress can switch performance between memory systems, biasing it toward habit in detriment of spatial or goal-directed strategies. In addition, novel synaptic mechanisms have been implicated in the effects of stress in plasticity and memory. Computational modeling is emerging as a useful approach to integrate and to ascertain neural and cognitive computations underlying different effects of stress in memory. Having provided novel explanations for the inverted-U-shaped relationship between stress and cognitive performance, model-based analysis studies can improve our understanding of diverse effects of stress in cognition and psychopathology.

Addresses

¹ Division of Cognitive Neuroscience, University of Basel, Birnamngasse 8, CH-4055 Basel, Switzerland

² Division of Molecular Neuroscience, University of Basel, Missionsstrasse 60/62A, CH-4055 Basel, Switzerland

³ Laboratory of Behavioral Genetics, Brain Mind Institute, Ecole Polytechnique Federale de Lausanne (EPFL), Station 19, CH-1015 Lausanne, Switzerland

Corresponding author: Sandi, Carmen (carmen.sandi@epfl.ch)

Current Opinion in Neurobiology 2011, 21:502–508

This review comes from a themed issue on Behavioral and cognitive neuroscience
Edited by Ann Graybiel and Richard Morris

Available online 16th April 2011

0959-4388/\$ – see front matter

© 2011 Elsevier Ltd. All rights reserved.

DOI 10.1016/j.conb.2011.03.003

Introduction

Stress can have profound effects on memory function, both when chronically experienced and when acutely coupled with a cognitive challenge [1]. In this review, we mainly focus on the effects of acute stress on learning, memory, and underlying mechanisms. We will also examine new modeling approaches aiming to ascertain the neural computations that underlie stress effects in these cognitive processes.

It is generally accepted that acute stress or elevated stress hormones facilitate memory consolidation while impairing the retrieval of information [2]. Moreover, acute stress is generally highly disruptive for working memory pro-

cessing, while it can facilitate implicit learning [1,2,3*]. Recently, several animal and human studies showed that stress switches performance between memory systems, promoting a transition from ‘flexible’ spatial or goal-directed learning to ‘rigid’ habitual stimulus–response (S–R) cognitive strategies [4*,5**,6,7].

However, one should be cautious when summarizing the effects of stress on a particular memory process as either positive or negative. Stress is not a unitary process, with its duration, intensity, or timing with regards to the cognitive challenge being critical for its cognitive outcome [1]. In particular, stress intensity has been long recognized in the literature as highly relevant, with the predominant belief that an inverted-U-shaped function can explain the relationship between stress intensity and memory (i.e. low and high stress levels impairing memory, whereas intermediate levels facilitating it). Strikingly, despite the popularity of this hypothesis, experimental evidence supporting the existence of an inverted-U-shaped relationship for performance of rats in a hippocampal learning task under the same experimental conditions was presented only recently [8].

Stress mediators: glucocorticoids and norepinephrine

Glucocorticoids and the noradrenergic system have been identified as key mediators of the cognitive effects of stress [1,2,9,10]. Adrenal glucocorticoids, steroid hormones produced by the adrenal glands, can cross the blood–brain barrier, gaining access to the brain. Through the binding to specific receptors, glucocorticoids can affect brain function and cognition via slow genomic and rapid non-genomic actions [11]. Noradrenergic neurons in the locus coeruleus project to many forebrain regions. Animal studies showed that corticosterone and norepinephrine (NE) act together to facilitate memory formation through actions involving the amygdala, hippocampus, and prefrontal cortex (PFC) [10,12]. More recently, glucocorticoids and the noradrenergic system were shown to interact in the medial PFC in producing working memory impairment [13]. In addition, the capability of these systems to influence learning processes depending on other brain systems — such as the dorsal striatum — was also shown [14].

Recent work has also implicated the action of glucocorticoids and noradrenergic activity in the switches between memory systems described above for stress. The activation of these stress systems was shown in rodents to elicit the

switch from spatial to stimulus–response (S–R) learning [15] and, in humans, from goal-directed to habitual control [16]. For example, control human subjects submitted to an outcome devaluation paradigm following instrumental learning in which two actions led to two distinct food outcomes, chose the action that was associated with the devalued outcome significantly less than at training. However, subjects that received the synthetic glucocorticoid hydrocortisone along with the α 2-adrenergic receptor antagonist yohimbine — which increases noradrenergic activity — were completely insensitive to the change in the value of the outcome [16]. It is important to note that only one dose of each tested hormone-targeting drug was included in these studies and, therefore, potential differences related to variations in the degree of activation of these systems have been, so far, not explored.

Recent human studies provide information about the network level mechanisms. Psychological stress was shown to reduce working memory-related activity in the dorsolateral PFC [17] as well as reward-related responses in the medial PFC without affecting ventral striatal responses [18]. Similarly, the concerted activation of glucocorticoids and the noradrenergic system induced a strong deactivation of prefrontal areas [19]. These findings suggest that stress-induced biases toward habitual behaviors may be mostly prompted by a primary impairment of PFC-dependent cognitive control mechanisms.

Stress and synaptic plasticity

There is great interest in understanding how acute stress affects synaptic plasticity in different brain regions with the final goal of identifying mechanisms explaining the divergent effects of stress in different memory systems. In general, stress impairs long-term potentiation (LTP) in the hippocampus and PFC, while it facilitates it in the amygdala [20]. In the hippocampus, stress also facilitates long-term depression (LTD) while glucocorticoid effects show an inverted-U-shape [20], with low and high corticosterone levels impairing and intermediate levels facilitating LTP [21].

Recent evidence points to glutamatergic mechanisms as key mediators of the cognitive actions of acute stress. In particular, α -amino-3-hydroxy-5-methylisoxazole-4-propionic acid receptor (AMPA) trafficking has been highlighted among the mechanisms whereby stress and glucocorticoids facilitate spatial memory [22**] and LTP [23]. Recent findings have implicated the mechanisms that underlie LTD induction among those that mediate the impairing effects of stress and glucocorticoids in the retrieval of information [24], including the activation of extrasynaptic GluN2B subunit-containing N-methyl-D-aspartate receptors (NMDARs) and the endocytosis of the GluA2 AMPAR subunit [9,24]. Regarding other synaptic proteins, strong evidence has accumulated for the involvement of the neural cell adhesion molecules

(NCAM) in both facilitating [25] and impairing [26] effects of stress in memory function [27].

Computational approaches to plasticity and learning

Computational approaches are emerging as an important development in neuroscience. By employing biologically plausible mathematical algorithms, computational modeling is providing insight into the nature of synaptic plasticity and learning, with a few recent attempts addressed to model the impact of stress on these processes. LTP and LTD have been modeled based on the

Box 1 Temporal difference reinforcement learning and its parameters

In TDRL [32], environment is described by *states* s that an agent (animal, human, or robot) occupies (e.g. its location — see Figure 1), and *actions* a it can perform to move between states and obtain rewards ($r > 0$) or face punishments ($r < 0$). The agent's goal is to learn the reward value of each state $V(s)$ and/or of each action — so-called Q -value $Q(s, a)$. If *outcomes* r depend only on the last state–action pair, not their entire history, the optimal actions can be learned efficiently using the TDRL algorithm. To account for future outcomes (albeit discounted by a positive factor $\gamma < 1$), the value of action a performed under state s at time t can be written as

$$Q(s_t, a_t) = E[r_t + \gamma \times r_{t+1} + \gamma^2 \times r_{t+2} + \dots] \\ = E[r_t] + \gamma \times Q(s_{t+1}, a_{t+1}). \quad (1)$$

If the *discounting factor* γ is small, immediate rewards are favored to future ones, whereas in case of large γ values immediate and future rewards are considered similarly important.

If actual reward r_t differs from expected reward $E[r_t]$, their difference — the reward prediction error — can be used to update the Q -values:

$$\Delta Q(s_t, a_t) = \alpha(r_t - E[r_t]) \\ = \alpha(r_t - Q(s_t, a_t) + \gamma \cdot Q(s_{t+1}, a_{t+1})),$$

where α is the learning rate.

Once Q -values of different actions are learned, they can be used for choosing the best action (i.e. the one with the highest value). To allow for exploration, actions can be chosen probabilistically, that is using a sigmoidal relationship between their Q -values and the corresponding action probabilities:

$$p(a) = \frac{\exp(\beta \cdot Q(s, a))}{\sum_i (\exp(\beta \cdot Q(s, a_i)))}$$

where the sum \sum_i runs over all actions a_i available from state s , $p(a)$ is the probability of choosing the action a , \exp the exponential function, and β the *exploitation factor* (also called inverse temperature), indicating steepness of the sigmoid.

Learning rate α , exploitation factor β and discounting factor γ can influence agent's learning behavior not only quantitatively but also qualitatively (see Figure 1). It has been suggested [41] that to achieve efficient learning α should be gradually decreased and β increased. This allows sufficient flexibility at early stages of learning and preservation/exploitation of the acquired knowledge later. However, optimal parameter settings strongly depend on task structure and requirements.

ideas of Hebbian learning [28]: positively correlated firing of connected neurons should lead to an increase in synaptic weights (LTP), whereas negatively correlated firing to a decrease in synaptic weights (LTD). In computational terms, this can be expressed as follows: if x and y are the firing rates of two connected neurons, $\langle x \rangle$ and $\langle y \rangle$ the respective average firing rates, and parameter α controls the learning rate, then the change in synaptic weight between these 2 neurons could be the following [29]:

$$\Delta w_{xy} = \alpha(x - \langle x \rangle) \cdot (y - \langle y \rangle)$$

To provide a better fit to biological data, several modifications of this learning rule have been suggested. For example, according to the Bienenstock–Cooper–Munro (BCM) rule [30], the change in synaptic weight is proportional to the presynaptic firing rate x and to a nonlinear function ϕ of the postsynaptic rate y and a threshold θ_m separating potentiation (with $y > \theta_m$) and depression (with $y < \theta_m$):

$$\Delta w_{xy} = \alpha x \phi(y, \theta_m)$$

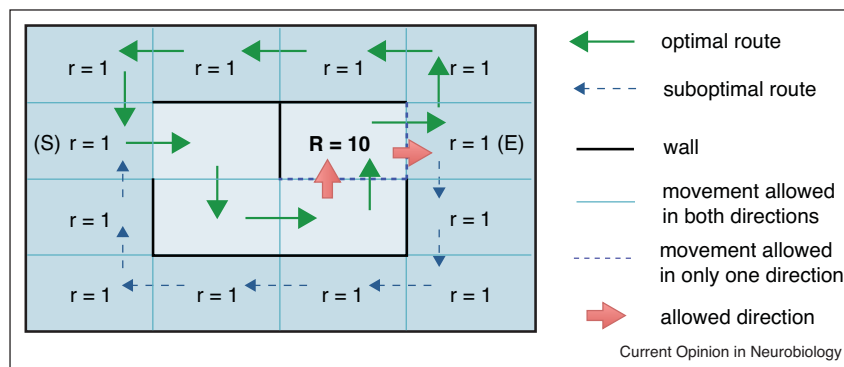
Threshold θ_m changes according to a slowly evolving time average of postsynaptic activity and can be influenced by various neurophysiological factors. For instance, in the hippocampus high levels of stress, through glucocorticoid actions, increase intracellular Ca^{2+} levels [21]; in BCM

theory it has been suggested [31] that these changes correspond to shifting the threshold θ_m to the right, which leads to a relative prevalence of LTD compared to LTP.

Behavior in many learning and memory tasks is guided by appetitive and aversive outcomes. The theory of *temporal difference reinforcement learning* (TDRL [32]; Box 1) was developed based on the assumption that intelligent agents learn reward-related consequences of actions they perform at different situations of their environment (so-called ‘states’), and based on this knowledge they can choose actions that lead to maximal expected reward. As reward-related consequences of actions can be both immediate and delayed, their relative importance is determined by the *discounting factor* (Box 1). Another key aspect in TDRL is the exploration–exploitation dilemma: should the seemingly best actions be chosen to gather maximal reward or should other (seemingly worse) actions be explored to obtain more accurate information about their reward-related consequences? How a TDRL model behaves here depends on the *exploitation factor* (Box 1).

In recent years key elements of TDRL models (such as reward prediction errors and action values) were shown to have neural correlates by numerous electrophysiological and neuroimaging studies in animals and humans [33–35]. While the simplest (discrete) version of TDRL was used in conditioning tasks [36**], it was also applied to spatial

Figure 1



Effects of TDRL parameters on agent's behavior. Imagine a maze, in whose center there is a platform with a large reward ($R = 10$). The environment is represented by 16 states (4×4). Actions correspond to movements between these states and are normally allowed in both directions, except in the case of walls (no crossing allowed) and around the platform (crossing allowed in only one direction — entry from the bottom and exit to the right). Generally the agent does not like to be in central area of the maze (e.g. aversion to open spaces), therefore locations near the outer wall are considered to have a small reward $r = 1$. The optimal strategy is to enter the central area from the left, visit the platform, and return along the top (a shorter route than along the bottom). In TDRL, however, the agent can learn and perform such a strategy only under certain parameter settings: **(a)** If discounting is steep, e.g. $\gamma = 0.5$, the agent will stick to the wall, because waiting for three steps to get the large reward has a lower discounted value ($\gamma^3 \cdot R_{t+3} = 0.5^3 \cdot 10 = 1.25$) than staying at the wall for 4 steps ($r_t + \gamma \cdot r_{t+1} + \gamma^2 \cdot r_{t+2} + \gamma^3 \cdot r_{t+3} = 1 + 0.5 + 0.5^2 + 0.5^3 = 1.875$, see Eq. (1) in Box 1). Under lesser discounting, e.g. $\gamma = 0.9$, going for the large reward ($0.9^3 \cdot 10 = 7.29$) becomes preferable to walking along the wall ($1 + 0.9 + 0.9^2 + 0.9^3 = 3.439$). **(b)** If exploitation β is very high already at the beginning, the agent will keep choosing the actions where it experienced (even small) rewards first, that is if at the starting location (S) it first chose to walk along the wall, it will never try going into the center because it *does not know* the reward is there (and is unlikely to discover it without sufficiently exploring the space). **(c)** If exploitation β remains low, the agent may learn going to the platform, but it may not always take the optimal (shorter) route along the top, because at the platform exit location (E) the difference between reward $R = 10$ discounted for nine steps (the optimal route) and for 11 steps (the suboptimal, longer route) is likely to be small. Such small differences in value can only be distinguished under high exploitation β and discounting factor γ . **(d)** Under explorative behavior (with low exploitation β), not only optimal performance is impossible, but because it takes longer to reach the goal, learning can be slower as well (with the same learning rates α).

learning by using hippocampal place cell-like continuous representation of states [37,38]. More recently, ‘three-factor’ rules that combine Hebbian learning with the reward-related term were also used to model spatial learning [39**].

Individual differences in learning under stress

Computational models are useful not only because they provide putative computational mechanisms of learning and memory, but also because they can account for individual differences. Performance of such models depends on certain parameters (e.g. exploitation and discounting factors in TDRL; Box 1 and Figure 1) that can be better interpreted in cognitive terms than classical behavioral measures or their principal components [40]. Such parameters can have neural correlates [41,42*] and differ between individuals [43,44*].

Considering the role of stress in learning and memory, trait anxiety has been identified as a key predictive factor for inter-individual differences in performance in learning tasks. As compared to low anxious rats, highly anxious ones display slower acquisition and worse retention in spatial versions of the water maze [8,45] as well as impaired place recognition memory in the Y-maze task [46]. Likewise, increasing anxiety levels — by injecting an anxiogenic drug either peripherally or directly into the amygdala and either pre-training or just before a probe trial — was shown to bias rats toward the use of (striatum-mediated) S–R strategies instead of (hippocampus-mediated) place strategies in navigational tasks that could be solved by either strategy [47]. When the same treatment was applied to animals trained in single-solution plus-maze tasks that require the use of either cognitive or habit learning, impaired (hippocampus-dependent) place memory and enhanced habit learning were observed [47].

Model-based analysis of stress and cognition

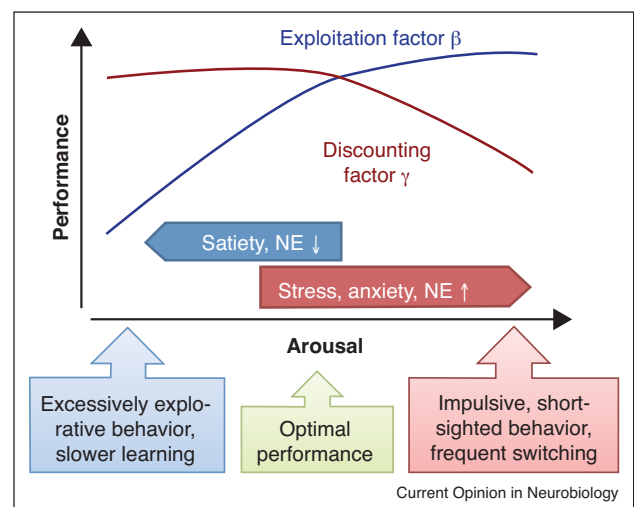
As opposed to classical neurobehavioral approaches, it has become increasingly popular to use computational models of behavior to infer underlying cognitive parameters and correlate them with experimental manipulations and neural changes [48**]. Motivated by the landmark discovery [33] of midbrain dopamine (DA) neurons in monkeys being responsive to the reward prediction error (i.e. the difference between actual and expected rewards), model-based analyses related discounting factors to the serotonin (5HT) levels [49], exploration to the frontopolar cortex activity [50], and learning rates to the anterior cingulate activity [51]. Polymorphisms in DARPP-32, DRD2 and COMT — genes that regulate different aspects of DA activity — were linked to differences in learning rates and exploration [43,44].

The model-based analysis approach has several benefits: firstly, while conventional behavioral measures (such as

escape latencies and response times) reflect a mix of cognitive variables (acquisition rate, memory strength, exploration vs. exploitation), computational model parameters can be interpreted without much ambiguity; secondly, if well-established models (e.g. TDRL) with identical or similar parameters are used to analyze data from different experiments, this enables comparability. However, certain requirements should be addressed to avoid faulty interpretation of results: firstly, experimental setup should not be oversimplified in modeling it, as even detailed aspects may substantially affect model’s performance; secondly, any essential model parameters should not be fixed but remain flexible — for instance, in TDRL fixing the exploitation factor may lead to attributing all behavioral differences to the learning rate, even if the former can explain them better.

The role of stress in learning was only recently addressed using computational approaches [36], where conditioning behavior of two genetic strains of mice (‘calm’ C57BL/6 and more anxious DBA/2) was modeled using TDRL. Individually estimated model parameters differed between genetic strains and stress conditions, and correlated

Figure 2



TDRL parameters and the inverted-U-shape. According to a recent computational model [36], the inverted-U-shaped relationship between stress/arousal and performance can be explained by changes in TDRL parameters: increasing exploitation β up to intermediate arousal levels and steeper reward discounting γ as a result of high stress. Stress, anxiety and increased NE can shift individual’s position on the x-axis to the right, whereas satiety or decreased NE can shift it to the left. In particular, frequent behavioral switches, observed at high levels of NE (or stress), can result from increasingly steep discounting combined with high exploitation. As Q-values of actions are defined as expected cumulative future reward (see Eq. (1) in Box 1), they become smaller not only if the actual reward diminishes but also if the discounting factor γ is decreased. As learning is applied for currently exploited actions, their Q-values become temporarily smaller compared to alternative actions. When these less favorable actions are exploited, their Q-values decrease as well, leading to a higher number of switches.

with individual's anxiety and motivation. The results showed that for less anxious animals stress increased exploitation, generally improving their performance, whereas for more anxious animals it led to steeper discounting, which impaired learning of delayed rewards. They suggest that both sufficient exploitation and not-too-steep discounting are necessary to achieve optimal performance at the middle of the inverted-U-shape (Figure 2). Consistent with this explanation, stress is linearly related to performance in simple tasks (which typically do not require learning over delays), and shows an inverted-U-shaped relationship in more complex tasks [20]. How exactly the state of high stress/norepinephrine should be interpreted is still unclear. The adaptive gain and optimal performance theory [52] proposed that labile attention and frequent switching observed under this state reflected exploration. However, recent model-based analysis studies [36,53*] found that increasing norepinephrine or stress levels did not lead to more exploration but rather the opposite. Thus, the observed behavior is possibly because of high exploitation combined with steeper discounting, which can result in relative devaluation of currently performed actions compared to their alternatives (Figure 2).

Conclusions

The active research field of stress and cognition has made substantial progress in recent years in characterizing a myriad of effects across learning types and stress conditions as well as network, cellular and molecular mechanisms involved in stress effects. Understanding how these different levels interact is difficult with conventional approaches. Computational modeling is an emerging approach to the field of stress and memory that allows revealing the fundamental cognitive computations affected by different degrees of stress in different individuals. Recent modeling results suggest that stress shifts cognitive operations from exploration predominant under low stress levels to increased exploitation at intermediate levels, and to steeper discounting of future rewards with increasing stress (and noradrenergic) levels. It is tempting to speculate that these operations can explain recent data indicating that stress switches performance between memory systems, from spatial or goal-directed learning to habitual stimulus–response cognitive strategies. Since using model-based analyses can help identify neural mechanisms underlying specific cognitive operations, we believe that their application to the field of stress and cognition can improve our understanding of the diversity of effects that stress impinges not only in learning and memory but also in various human psychopathologies.

Acknowledgements

The authors thank Eleni Vasilaki for her useful comments and suggestions about the manuscript. This work was supported by grants from the EU 7th Framework Program (MemStick), the Swiss National Science Foundation (310000-120791; Sinergia CRSIK0-122697 and CRSIK0-122691) and by Ecole Polytechnique Federale de Lausanne (EPFL).

References and recommended reading

Papers of particular interest, published within the period of review, have been highlighted as:

- of special interest
- of outstanding interest

1. Sandi C, Pinelo-Nava MT: **Stress and memory: behavioral effects and neurobiological mechanisms.** *Neural Plast* 2007. doi:10.1155/2007/78970.
2. de Quervain DJ, Aerni A, Schelling G, Roozendaal B: **Glucocorticoids and the regulation of memory in health and disease.** *Front Neuroendocrinol* 2009, **30**:358-370.
3. Luethi M, Meier B, Sandi C: **Stress effects on working memory, explicit memory, and implicit memory for neutral and emotional stimuli in healthy men.** *Front Behav Neurosci* 2009 doi: 10.3389/neuro.08.005.2008.
This study includes a wide battery of memory tests and shows, in humans, that stress can simultaneously have opposing effects on different learning and memory processes in the same individuals.
4. Schwabe L, Wolf OT: **Stress prompts habit behavior in humans.** *J Neurosci* 2009, **29**:7191-7198.
This study exposed human subjects to stress before being trained in instrumental actions with different reward outcomes. When, afterwards, one of the outcomes was devalued, stress rendered participants' behavior insensitive to the change in the value of the outcome. Thus, this was the first study showing that stress leads to habit performance in humans.
5. Dias-Ferreira E, Sousa JC, Melo I, Morgado P, Mesquita AR, Cerqueira JJ, Costa RM, Sousa N: **Chronic stress causes frontostriatal reorganization and affects decision-making.** *Science* 2009, **325**:621-625.
The study showed that in rats challenged in decision-making tasks, chronic stress leads to a bias in behavioral strategies toward habit, in parallel with corresponding structural changes in the corticostriatal circuits underlying different behavioral strategies. These data are the first highlighting neural mechanisms underlying stress effects in the switch between memory systems and behavioral strategies.
6. Schwabe L, Oitzl MS, Richter S, Schächinger H: **Modulation of spatial and stimulus–response learning strategies by exogenous cortisol in healthy young women.** *Psychoneuroendocrinology* 2009, **34**:358-366.
7. Schwabe L, Schächinger H, de Kloet ER, Oitzl MS: **Stress impairs spatial but not early stimulus–response learning.** *Behav Brain Res* 2010, **213**:50-55.
8. Salehi B, Cordero MI, Sandi C: **Learning under stress: the inverted-U-shape function revisited.** *Learn Mem* 2010, **17**:522-530.
9. Sandi C: **Glucocorticoids act through glutamatergic pathways to affect memory processes.** *Trends Neurosci* 2011, **34**:165-176.
10. Roozendaal B, McEwen BS, Chattarji S: **Stress, memory and the amygdala.** *Nat Rev Neurosci* 2009, **10**:423-433.
11. de Kloet ER, Karst H, Joëls M: **Corticosteroid hormones in the central stress response: quick-and-slow.** *Front Neuroendocrinol* 2008, **29**:268-272.
12. McGaugh JL, Roozendaal B: **Role of adrenal stress hormones in forming lasting memories in the brain.** *Curr Opin Neurobiol* 2002, **12**:205-210.
13. Barsegyan A, Mackenzie SM, Kurose BD, McGaugh JL, Roozendaal B: **Glucocorticoids in the prefrontal cortex enhance memory consolidation and impair working memory by a common neural mechanism.** *Proc Natl Acad Sci U S A* 2010, **107**:16655-16660.
14. Quirarte GL, de la Teja IS, Casillas M, Serafín N, Prado-Alcalá RA, Roozendaal B: **Corticosterone infused into the dorsal striatum selectively enhances memory consolidation of cued water-maze training.** *Learn Mem* 2009, **16**:586-589.
15. Schwabe L, Schächinger H, de Kloet ER, Oitzl MS: **Corticosteroids operate as a switch between memory systems.** *J Cogn Neurosci* 2010, **22**:1362-1372.

16. Schwabe L, Tegenthoff M, Höffken O, Wolf OT: **Concurrent glucocorticoid and noradrenergic activity shifts instrumental behavior from goal-directed to habitual control.** *J Neurosci* 2010, **30**:8190-8196.
17. Quin S, Hermans EJ, van Marie HJ, Luo J, Fernandez G: **Acute psychological stress reduces working memory-related activity in the dorsolateral prefrontal cortex.** *Biol Psychiatry* 2009, **66**:25-32.
18. Ossewaarde L, Qin S, van Marie HJ, van Wingen GA, Fernandez G, Hermans EJ: **Stress-induction reduction in reward-related prefrontal cortex.** *Neuroimage* 2011, **55**:345-352.
19. van Stegeren AH, Roozendaal B, Kindt M, Wolf OT, Joëls M: **Interacting noradrenergic and corticosteroid systems shift human brain activation patterns during encoding.** *Neurobiol Learn Mem* 2010, **93**:56-65.
20. Diamond DM, Campbell AM, Park CR, Halonen J, Zoladz PR: **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 Plast* 2007, **2007**:60803.
21. Joëls M: **Corticosteroid effects in the brain: U-shape it.** *Trends Pharmacol Sci* 2006, **27**:244-250.
22. Conboy L, Sandi C: **Stress at learning facilitates memory formation by regulating AMPA receptor trafficking through a glucocorticoid action.** *Neuropsychopharmacology* 2010, **35**:674-685.
- Using the water maze spatial task involving different stress levels, mice trained under high stressful conditions showed better learning and memory, higher corticosterone levels, and enhanced synaptic expression of GluA2 AMPARs, compared to those trained under lower stress. This is the first study that identified trafficking of AMPARs as a mechanism whereby stress and glucocorticoids facilitate memory processes.
23. Groc L, Choquet D, Chaouloff F: **The stress hormone corticosterone conditions AMPAR surface trafficking and synaptic potentiation.** *Nat Neurosci* 2008, **11**:868-870.
24. Wong TP, Howland JG, Robillard JM, Ge Y, Yu W, Titterness AK, Brebner K, Liu L, Weinberg J, Christie BR *et al.*: **Hippocampal long-term depression mediates acute stress-induced spatial memory retrieval impairment.** *Proc Natl Acad Sci U S A* 2007, **104**:11471-11476.
25. Lopez-Fernandez MA, Montaron MF, Varea E, Rougon G, Venero C, Abrous DN, Sandi C: **Upregulation of polysialylated neural cell adhesion molecule in the dorsal hippocampus after contextual fear conditioning is involved in long-term memory formation.** *J Neurosci* 2007, **27**:4552-4561.
26. Bisaz R, Schachner M, Sandi C: **Causal evidence for the involvement of the neural cell adhesion molecule, NCAM, in chronic stress-induced cognitive impairments.** *Hippocampus* 2011, **21**:56-71.
27. Bisaz R, Conboy L, Sandi C: **Learning under stress: a role for the neural cell adhesion molecule NCAM.** *Neurobiol Learn Mem* 2009, **91**:333-342.
28. Hebb DO: *The Organization of Behavior.* New York: Wiley; 1949.
29. Sejnowski TJ, Tesauro G: **The Hebb rule for synaptic plasticity: algorithms and implementations.** In *Neural Models of Plasticity.* Edited by Byrne JH, Berry WO. Academic Press; 1989:94-103.
30. Bienenstock EL, Cooper LN, Munro PW: **Theory of the development of neuron selectivity: orientation specificity and binocular interaction in visual cortex.** *J Neurosci* 1982, **2**:32-48.
31. Kim JJ, Yoon KS: **Stress: metaplastic effects in the hippocampus.** *Trends Neurosci* 1998, **21**:505-509.
32. Sutton RS, Barto AG: *Reinforcement Learning: An Introduction.* MIT Press; 1998.
33. Schultz W, Dayan P, Montague PR: **A neural substrate of prediction and reward.** *Science* 1997, **275**:1593-1599.
34. Samejima K, Ueda Y, Doya K, Kimura M: **Representation of action-specific reward values in the striatum.** *Science* 2005, **310**:1337-1340.
35. O'Doherty JP, Dayan P, Friston K, Critchley H, Dolan RJ: **Temporal difference models and reward-related learning in the human brain.** *Neuron* 2003, **28**:329-337.
36. Luksys G, Gerstner W, Sandi C: **Stress, genotype and norepinephrine in the prediction of mouse behavior using reinforcement learning.** *Nat Neurosci* 2009, **12**:1180-1186.
- In this study, effects of stress and other modulatory factors on learning in mice are analyzed using a TDRL model with individually estimated and dynamic parameters. The authors suggest a computational explanation for the inverted-U-shaped relationship between arousal and performance based on observed changes in exploitation and discounting factors.
37. Foster DJ, Morris RG, Dayan P: **A model of hippocampally dependent navigation, using the temporal difference learning rule.** *Hippocampus* 2000, **10**:1-16.
38. Strösslin T, Sheynikhovich D, Chavarriga R, Gerstner W: **Robust self-localisation and navigation based on hippocampal place cells.** *Neural Netw* 2005, **18**:1125-1140.
39. Vasilaki E, Frémaux N, Urbanczik R, Senn W, Gerstner W: **Spike-based reinforcement learning in continuous state and action space: when policy gradient methods fail.** *PLoS Comput Biol* 2009, **5**:e1000586.
- In this study a biologically realistic model that uses a reward-modulated synaptic learning rule for spiking neurons and a network architecture combining feed-forward with lateral connections shows learning in a Morris water maze-like task that is consistent with animal experiments. This paper also includes a review on three factor rules for spiking neuron models.
40. Clément Y, Joubert C, Kopp C, Lepicard EM, Venault P, Misslin R, Cadot M, Chapouthier G: **Anxiety in mice: a principal component analysis study.** *Neural Plast* 2007, **2007**:35457.
41. Doya K: **Metalearning and neuromodulation.** *Neural Netw* 2002, **15**:495-506.
42. Doya K: **Modulators of decision making.** *Nat Neurosci* 2008, **11**:410-416.
- This paper provides a comprehensive review of recent experimental, theoretical, and model-based studies addressing modulators of different aspects of decision making and their putative neural mechanisms.
43. Frank MJ, Moustafa AA, Haughey HM, Curran T, Hutchison KE: **Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning.** *Proc Natl Acad Sci U S A* 2007, **104**:16311-16316.
44. Frank MJ, Doll BB, Oas-Terpstra J, Moreno F: **Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation.** *Nat Neurosci* 2009, **12**:1062-1068.
- In this study the authors use a reinforcement learning model to relate polymorphisms in three dopaminergic genes to differences in exploitative learning based on positive and negative outcomes as well as in directed exploration.
45. Herrero AI, Sandi C, Venero C: **Individual differences in anxiety trait are related to spatial learning abilities and hippocampal expression of mineralocorticoid receptors.** *Neurobiol Learn Mem* 2006, **86**:150-159.
46. Hawley W, Grissom E, Keskitalo L, Hastings T, Dohanich G: **Sexual motivation and anxiety-like behaviors of male rats after exposure to a trauma followed by situational reminders.** *Physiol Behav* 2011, **102**:181-187.
47. Packard MG: **Anxiety, cognition, and habit: a multiple memory systems perspective.** *Brain Res* 2009, **1293**:121-128.
48. Corrado G, Doya K: **Understanding neural coding through the model-based analysis of decision making.** *J Neurosci* 2007, **27**:8178-8180.
- This paper introduces a novel methodology of analyzing behavior and its neural correlates using computational models whose internal parameters correspond to subjective decision variables. Basic requirements for such models and several examples of their application are discussed.
49. Schweighofer N, Bertin M, Shishida K, Okamoto Y, Tanaka SC, Yamawaki S, Doya K: **Low-serotonin levels increase delayed reward discounting in humans.** *J Neurosci* 2008, **28**:4528-4532.
50. Daw ND, O'Doherty JP, Dayan P, Seymour B, Dolan RJ: **Cortical substrates for exploratory decisions in humans.** *Nature* 2006, **441**:876-879.

51. Behrens TE, Woolrich MW, Walton ME, Rushworth MF: **Learning the value of information in an uncertain world.** *Nat Neurosci* 2007, **10**:1214-1221.
52. Aston-Jones G, Cohen JD: **An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance.** *Annu Rev Neurosci* 2005, **28**:403-450.
53. Jepma M, Te Beek ET, Wagenmakers EJ, van Gerven JM,
 - Nieuwenhuis S: **The role of the noradrenergic system in the**

exploration–exploitation trade-off: a psychopharmacological study. *Front Hum Neurosci* 2010, **4**:170.

This study addresses the question of whether pharmacologically increased norepinephrine levels lead to more explorative behavior in humans using two different experiments and model-based analysis in one of them. The reported results show no decrease in exploitation, suggesting that changes in other cognitive processes (such as the learning rate or future discounting) might be responsible for the impairment of focused performance at high NE levels.