Not all that glittered is gold: neural mechanisms that determine when reward will enhance or impair memory

David V. Clewett[1] and Mara Mather[1,2,3]*

[1] Neuroscience Graduate Program, University of Southern California, Los Angeles, CA, USA
[2] Davis School of Gerontology, University of Southern California, Los Angeles, CA, USA
[3] Psychology Department, University of Southern California, Los Angeles, CA, USA
*Correspondence: mara.mather@usc.edu

A commentary on

High reward makes items easier to remember, but harder to bind to a new temporal context
by Madan, C. R., Fujiwara, E., Gerson, B. C., and Caplan, J. B. (2012). Front. Integr. Neurosci. 6:61. doi: 10.3389/fnint.2012.00061

Rewards optimize future behavior by enhancing memory of events relevant to wellbeing. But how do rewards received in one context affect subsequent learning about the information previously associated with rewards? Updating existing knowledge is important to maximize rewards: when high-value stimuli are reinforced by reward in a new situation, contextual binding will support optimal decision-making later; however, contextual *un*binding can also improve wellbeing by preventing low utility associations from being formed.

Previous work indicates that initial emotional associations can impair later novel associations (Mather and Knight, 2008; Novak and Mather, 2009). To explore whether monetary rewards elicit memory impairments or enhancements, Madan et al. (2012) used a series of value-learning tasks to examine the effects of reward on explicit memory, implicit memory, and contextual binding. Consistent with earlier studies using emotional stimuli, high-value words were better recalled than low-value words on both implicit and explicit memory tests. Moreover, high-value words were more difficult to bind

to a new temporal context, as measured by relatively later word output during free recall of a list. Interestingly, though free recall was equal for high- and low-value items, high-value words tended to proactively interfere with current list memory. These findings add to a growing literature showing that arousal can sometimes enhance and sometimes impair memory for context (see Mather and Sutherland, 2011), yet to date there is no neural model that can account for these dual effects. Below we discuss potential neural mechanisms that may help explain Madan et al.'s finding that rewards, much like other forms of arousal, impair subsequent context binding and proactively interfere with memory.

Motivationally significant events trigger release of the catecholamines, dopamine (DA) and norepinephrine (NE), which optimize central resource allocation to ensure that such events are remembered (McGaugh and Roozendaal, 2002; Adcock et al., 2006). However, how these brain systems sometimes lead to memory impairment is less understood. One possibility is that the mnemonic effect of arousal depends on the predictive value or goal-relevance of a salient item's context (Mather and Sutherland, 2011). Catecholamines are thought to broadcast prediction errors across the brain, providing a mechanism by which arousal can either enhance or impair memory (Hollerman and Schultz, 1998; Harley, 2004). Prediction errors arise from perceived discrepancies between expected and actual outcomes, and are therefore

sensitive to changes in the context in which salient stimuli occur. In Madan et al.'s list-discrimination task, participants re-learned the high- and low-value words (along with novel words) in unrewarded lists. Thus, their finding that previous rewards impair rather than enhance memory binding to a new temporal context may reflect the influence of prediction errors on memory-updating processes driven by catecholamines.

The hippocampus represents associative information and selectively binds or unbinds contextual information in a manner consistent with prediction errors (Mizumori, 2013). Thus, in the study by Madan et al. (2012), reduced DA in the hippocampus may have impaired contextual binding due to a lack of reward reinforcement during re-learning (i.e., negative prediction error). Contextual mismatches also engage the orbitofrontal cortex (OFC), which helps update memory representations to account for shifting reward contingencies, such as during reversal learning (O'Doherty et al., 2001; Rolls, 2004). Similarly, in monkeys, phasic NE release signals changes in stimulus-reward associations (Aston-Jones et al., 1997), a process that is likely driven by context-specific inputs from the OFC (Aston-Jones and Cohen, 2005). Taken together, these findings suggest that arousal's influence on associative memory may be determined by catecholaminergic modulation of brain structures that track the predictive value of contextual information.

An interesting implication of this prediction error account is that punishment-related incentives should enhance associative binding during the list-discrimination task. Updating previously threatening stimuli to represent a safer environment would support memory of cues that also promote wellbeing. However, this account contradicts the prevailing view that arousal-induced memory impairments are invariant to stimulus valence (Mather and Knight, 2008). Thus, one alternative explanation is that arousal's selective influence on binding is not determined by valence but rather by the priority of the context in which an arousing item is encoded.

According to the *arousal-biased competition* (ABC) theory, arousal will amplify the effects of priority such that memory of goal-relevant stimuli is enhanced, while memory of less important information is suppressed (Mather and Sutherland, 2011). Since there was no instruction to pay special attention to the new lists in Madan et al.'s study, any arousal induced by remembering a prior reward would impair context binding. However, if the experimenters explicitly directed participants to prioritize learning the current list context, ABC theory would predict that arousal should enhance rather than impair binding of high-value words to their current list. Consistent with this account, one recent study demonstrated that harbinger cues that predict emotional stimuli enhance rather than impair associative binding when participants are made aware of their contingencies (Sakaki et al., 2014). We propose that catecholamines also play a key role in ABC, since they mediate arousal's selective influence on cognitive processing (Robbins and Arnsten, 2009).

In summary, Madan et al.'s findings underscore that adaptive behavior not only relies on the ability of rewards to enhance memory of significant stimuli but to also impair binding of new contextual associations that do not predict such events. We propose that DA and NE help determine reward's divergent effects on memory by interacting with regions that update the salience and priority of a stimulus. Whereas such neuromodulation optimizes wellbeing under normal circumstances, acute stress appears to hijack these

mechanisms in ways harmful to behavior. Acute stress amplifies reward salience (Mather and Lighthall, 2012), resulting in a shift from goal-directed processes toward habitual responding (Schwabe and Wolf, 2011). This stress-induced shift appears to be driven by elevated levels of catecholamines and glucocorticoids in the OFC (Robbins and Arnsten, 2009), which reduce its sensitivity to reward devaluation (van Eimeren et al., 2009; Schwabe et al., 2011, 2012). Under normal circumstances, the updating of emotional but not neutral associations relies on greater OFC activation (Sakaki et al., 2011; Nashiro et al., 2012, 2013). Thus, the impairing effects of elevated stress hormones on OFC function may be one mechanism by which devalued items proactively interfere with memory. Ultimately, understanding how arousal affects associative memory has important implications for other pathologies, such as post-traumatic stress disorder, in which the inability to restrict fear to the appropriate context leads to intrusive thoughts and impairs day-to-day life.

REFERENCES

Adcock, R. A., Thangavel, A., Whitfield-Gabrieli, S., Knutson, B., and Gabrieli, J. D. (2006). Reward-motivated learning: mesolimbic activation precedes memory formation. *Neuron* 50, 507–517. doi: 10.1016/j.neuron.2006.03.036

Aston-Jones, G., and Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: adaptive gain and optimal performance. *Annu. Rev. Neurosci.* 28, 403–450. doi: 10.1146/annurev.neuro.28.061604.135709

Aston-Jones, G., Rajkowski, J., and Kubiak, P. (1997). Conditioned responses of monkey locus coeruleus neurons anticipate acquisition of discriminative behavior in a vigilance task. *Neuroscience* 80, 697–715. doi: 10.1016/S0306-4522(97)00060-2

Harley, C. W. (2004). Norepinephrine and dopamine as learning signals. *Neural Plast.* 11, 191–204. doi: 10.1155/NP.2004.191

Hollerman, J. R., and Schultz, W. (1998). Dopamine neurons report an error in the temporal prediction of reward during learning. *Nat. Neurosci.* 1, 304–309. doi: 10.1038/1124

Madan, C. R., Fujiwara, E., Gerson, B. C., and Caplan, J. B. (2012). High reward makes items easier to remember, but harder to bind to a new temporal context. *Front. Integr. Neurosci.* 6:61. doi: 10.3389/fnint.2012.00061

Mather, M., and Knight, M. (2008). The emotional harbinger effect: poor context memory for cues that previously predicted something arousing. *Emotion* 8, 850–860. doi: 10.1037/a0014087

Mather, M., and Lighthall, N. R. (2012). Risk and reward are processed differently in decisions made under stress. *Curr. Dir. Psychol. Sci.* 21, 36–41. doi: 10.1177/0963721411429452

Mather, M., and Sutherland, M. R. (2011). Arousal-biased competition in perception and memory. *Perspect. Psychol. Sci.* 6, 114–133. doi: 10.1177/1745691611400234

McGaugh, J. L., and Roozendaal, B. (2002). Role of adrenal stress hormones in forming lasting memories in the brain. *Curr. Opin. Neurobiol.* 12, 205–210. doi: 10.1016/S0959-4388(02)00306-9

Mizumori, S. J. (2013). Context prediction analysis and episodic memory. *Front. Behav. Neurosci.* 7:132. doi: 10.3389/fnbeh.2013.00132

Nashiro, K., Sakaki, M., Nga, L., and Mather, M. (2012). Differential brain activity during emotional vs. non-emotional reversal learning. *J. Cogn. Neurosci.* 24, 1794–1805. doi: 10.1162/jocn_a_00245

Nashiro, K., Sakaki, M., Nga, L., and Mather, M. (2013). Age-related similarities and differences in brain activity underlying reversal learning. *Front. Integr. Neurosci.* 7:37. doi: 10.3389/fnint.2013.00037

Novak, D. L., and Mather, M. (2009). The tenacious nature of memory binding for arousing negative items. *Mem. Cognit.* 37, 945–952. doi: 10.3758/MC.37.7.945

O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., and Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nat. Neurosci.* 4, 95–102. doi: 10.1038/82959

Robbins, T. W., and Arnsten, A. F. (2009). The neuropsychopharmacology of fronto-executive function: monoaminergic modulation. *Annu. Rev. Neurosci.* 32, 267. doi: 10.1146/annurev.neuro.051508.135535

Rolls, E. T. (2004). The functions of the orbitofrontal cortex. *Brain Cogn.* 55, 11–29. doi: 10.1016/S0278-2626(03)00277-X

Sakaki, M., Niki, K., and Mather, M. (2011). Updating existing emotional memories involves the frontopolar/orbitofrontal cortex in ways that acquiring new emotional memories does not. *J. Cogn. Neurosci.* 23, 3498–3514. doi: 10.1162/jocn_a_00057

Sakaki, M., Ycaza, A. E., and Mather, M. (2014). Association learning for emotional harbinger cues: when do previous emotional associations impair and when do they facilitate subsequent learning of new associations? *Emotion.* 14, 115–129. doi: 10.1037/a0034320

Schwabe, L., Höffken, O., Tegenthoff, M., and Wolf, O. T. (2011). Preventing the stress-induced shift from goal-directed to habit action with a β-adrenergic antagonist. *J. Neurosci.* 31, 17317–17325. doi: 10.1523/JNEUROSCI.3304-11.2011

Schwabe, L., Tegenthoff, M., Höffken, O., and Wolf, O. T. (2012). Simultaneous glucocorticoid and noradrenergic activity disrupts the neural basis of goal-directed action in the human brain. *J. Neurosci.* 32, 10146–10155. doi: 10.1523/JNEUROSCI.1304-12.2012

Schwabe, L., and Wolf, O. T. (2011). Stress-induced modulation of instrumental behavior: from goal-directed to habitual control of action. *Behav. Brain Res.* 219, 321–328. doi: 10.1016/j.bbr.2010.12.038

van Eimeren, T., Ballanger, B., Pellecchia, G., Miyasaki, J. M., Lang, A. E., and Strafella, A. P. (2009). Dopamine agonists diminish value

sensitivity of the orbitofrontal cortex: a trigger for pathological gambling in parkinson's disease? *Neuropsychopharmacology* 34, 2758–2766. doi: 10.1038/npp.2009.124

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.