Abstract

Habits and Affects: Learning by an Associative Two-Process [†]

Robert Lowe

Department of Applied IT, University of Gothenburg, 412 96 Göteborg, Sweden; robert.lowe@gu.se

[†] Presented at the IS4SI 2017 Summit DIGITALISATION FOR A SUSTAINABLE SOCIETY, Gothenburg, Sweden, 12–16 June 2017.

In animal learning theory, the notion of habits is frequently employed to describe instrumental behaviour that is (among others): inflexible (i.e., slow to change), unconscious, insensitive to reinforcer devaluation [1,2]. It has also been suggested that learning using reinforcement learning algorithms somewhat reflects a transition from affect-based to more habit-based behaviour [2] where dual memory systems for affective working memory and standard (e.g., spatial) working memory systems exist [3,4].

Associative Two-Process theory has been proposed to explain phenomena emergent from differential outcomes training. In this procedure, animals (sometimes humans) are presented with stimuli/objects that uniquely identify differential outcomes, e.g., a circle stimulus precedes the presentation of a food outcome, a square stimulus precedes the presentation of a toy outcome. Outcomes are, in turn, mitigated by specific responses, e.g., press the right button to obtain the food, press the left button to obtain the toy. Manipulating these stimuli, response, outcome contingencies reveals the two types of memory, i.e., one that concerns 'standard' working memory of stimulus-response associations, the other that concerns 'prospective' memory, that stimulus-expectation-response follows in a sequence.

The neural dynamic relationship between the purported dual memory structures may vary depending on the stage of learning at which the animal/human (agent) has arrived at. Previously it has been suggested [5], and neural-computationally demonstrated, that a working memory route is critical in initial learning trials where the agent is presented sequentially with a given stimulus, action/behavioural options, and finally an outcome (e.g., rewarding stimulus or absence thereof). Subsequent trials lead to a dominance of affective (or otherwise prospective) memory that effectively scaffolds the learning of the outcome-achieving stimulus-response rules under conditions of relative uncertainty. Finally, during later stages of learning more 'habitual' responding may occur where the retrospective route becomes dominant and 'overshadows' the prospective memory.

In neural anatomical terms, candidate structures for implementing prospective memory include the orbitofrontal cortex (OFC), which is considered to enable fast, flexible and context-based learning (particularly important in studies of reversal learning [6]). This is in contrast to the amygdala, which is considered less flexible, i.e., resistant to unlearning, but, nevertheless, critical to learning valuations of stimuli [7]. Furthermore, the interplay between the basolateral division of the amygdala (BLA) and OFC may be crucial in differential reward evaluation [8,9]. Passingham and Wise [9] have suggested that medial prefrontal cortex (PFC) has a critical role in encoding outcome-contingent choice, whereas Watanabe et al. [4] have provided evidence for the lateral PFC integrating activation inputs from 'retrospective' (working memory) areas such as dorsal PFC and 'prospective' (outcome expectant) areas such as OFC and medial PFC.

A perspective of Urcuioli [10,11] is that outcome expectancies (from prospective memory) provide a means to effectively classify stimuli. Action selection can then be simplified through exploiting affordances of the subset of those actions already associated with the outcome expectancy classes. This is a reason why participants under certain forms of differential outcomes training can immediately select the unique action that leads to the desired outcome even though the stimulus-

action (response) contingency has previously not been experienced: Subjects have already classified the stimuli according to a given outcome expectancy previously associated with an action.

In this work, I discuss the associative two-process model in relation to (standard) working memory and 'affective working memory' [4] as providing a means to classify stimuli. I refer to a number of animal learning paradigms that demonstrate the potential for reward and reward omission anticipation to be associated with reward-promoting behaviour (cf. [11–15]) and neural computational aspects of the interplay of affective (prospective) and working (retrospective) memory that may yield more habitual behaviour. I show that, within an associative two-process context, habits can also be understood in terms of affective working memory—specifically in relation to reward acquisition expectation and reward omission expectation. Habits, in this context are considered behaviours that are inflexibly selected for in spite of reinforcer devaluation and their rigidity reflects the certainty/uncertainty of a particular rewarding outcome.

I discuss the implications for such learning of habits and affective mediations of behaviour particularly regarding memory and clinical conditions (e.g., Alzheimer's) and learning children. This may be informing of new digitized solutions for intervention approaches with senior citizens and pedagogy in relation to children development.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Dickinson, A. Actions and habits: The development of behavioural autonomy. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **1985**, *308*, 67–78, doi:10.1098/rstb.1985.0010.
2. Seger, C.A.; Spiering, B.J. A critical review of habit learning and the basal ganglia. *Front. Syst. Neurosci.* **2011**, *5*, doi:10.3389/fnsys.2011.00066.
3. Davidson, R.J.; Irwin, W. The functional neuroanatomy of emotion and affective style. *Trends Cogn. Neurosci.* **1999**, *3*, 11–21, doi:10.1016/S1364-6613(98)01265-0.
4. Watanabe, M.; Hikosaka, K.; Sakagami, M.; Shirakawa, S. Reward expectancy-related prefrontal neuronal activities: Are they neural substrates of ''affective'' working memory? *Cortex* **2007**, *43*, 53–64, doi:10.1016/S0010-9452(08)70445-3.
5. Lowe, R.; Sandamirskaya, Y.; Billing, E. The actor—Differential outcomes critic: A neural dynamic model of prospective overshadowing of retrospective action control. In Proceedings of the Fourth Joint IEEE Conference on Development and Learning and on Epigenetic Robotics, Genoa, Italy, 13–16 October 2014; pp. 440–447, doi:10.1109/DEVLRN.2014.6983021.
6. Delamater, A.R. The role of the orbitofrontal cortex in sensory-specific encoding of associations in pavlovian and instrumental conditioning. *Ann. N. Y. Acad. Sci.* **2007**, 1121, 152–173, 10.1196/annals.1401.030.
7. Schoenbaum, G.; Saddoris, M.; Stalnaker, T. Reconciling the roles of orbitofrontal cortex in reversal learning and the encoding of outcome expectancies. *Ann. N. Y. Acad. Sci.* **2007**, *1121*, 320–335, 10.1196/annals.1401.001.
8. Ramirez, D.; Savage, L. Differential involvement of the basolateral amygdala, orbitofrontal cortex, and nucleus accumbens core in the acquisition and use of reward expectancies. *Behav. Neurosci,* **2007**, *121,* 896–906, doi:10.1037/0735-7044.121.5.896.
9. Passingham, R.; Wise, S. *The Neurobiology of the Prefrontal Cortex: Anatomy, Evolution, and the Origin of Insight*; Oxford University Press: Oxford, UK, 2012; Volume 50.
10. Urcuioli, P.J. Behavioral and associative effects of differential outcomes in discriminating learning. *Learn. Behav.* **2005**, *33*, 1–21, doi:10.3758/BF03196047.
11. Urcuioli, P. Stimulus control and stimulus class formation. In *APA Handbook of Behavior Analysis*, 1st ed.; Madden, G.J., Dube, W.V., Hackenberg, T.D., Hanley, G.P., Lattal, K.A., Eds.; American Psychological Association: Washington, DC, USA, 2013; Volume 1, pp. 361–386.
12. Overmier, J.B.; Lawry, J.A. Pavlovian conditioning and the mediation of behavior. *Psychol. Learn. Motiv.* **1979**, *13*, 1–55, doi:10.1016/S0079-7421(08)60080-8.
13. Kruse, J.M.; Overmier, J.B. Anticipation of reward omission as a cue for choice behavior. *Learn. Motiv.* **1982**, *13*, 505–525, doi:10.1016/0023-9690(82)90007-8.

14. Lowe, R.; Almer, A.; Lindblad, G.; Gander, P.; Michael, J.; Vesper, C. Minimalist social-affective value for use in joint action: A neural-computational hypothesis. *Front. Comput. Neurosci.* **2016**, *10*, doi:10.3389/fncom.2016.00088.

15. Lowe, R.; Billing, E. Affective-Associative Two-Process theory: A neural network investigation of adaptive behaviour in differential outcomes training. *Adapt. Behav.* **2017**, *25*, 5–23, doi:10.1177/1059712316682999.