

COGNITIVE NEUROSCIENCE

COMMENTARY

Walking bundles of habits (and response–outcome associations) (Commentary on Liljeholm *et al.*)



Aaron M. Bornstein, Amitai Shenhav and Kevin J. Miller
Princeton Neuroscience Institute, Princeton University, Princeton, NJ, USA

In the shapeliness of a life, habit plays its sovereign role... Mary Oliver

Distinct behaviors imply distinct neural processes. Goal-directed and habitual control have been repeatedly shown to have starkly different behavioral signatures and neural substrates. The distinction critically hinges on the necessary underlying representations: stimulus–response (S–R) associations subserve habits, while response–outcome (R–O) associations allow flexible, goal-directed control. But another separation remains relatively underexplored: that between acquisition and expression. Are the structures that learn S–R and R–O representations also responsible for acting on them?

Lesion studies in animal models have fractionated cortical and striatal subregions based on separable contributions to learning and performance (Ostlund & Balleine, 2005; Atallah *et al.*, 2007). However, while these studies offer temporal specificity and causal manipulation, they have necessarily been restricted to select brain regions. Studies employing lesions offer temporal specificity and causal manipulation, and have been conducted in many brain regions (for a review, see Balleine & O’Doherty, 2010), but suffer from logical limitations: the finding that a lesion disables expression does not exclude a structure’s role in acquisition. Previous human functional magnetic resonance imaging (fMRI) studies have examined goal-directed and habitual control systems across the whole brain, but were not designed to examine the acquisition/expression distinction (Tricomi *et al.*, 2009; Gläscher *et al.*, 2010; Liljeholm *et al.*, 2012).

A new study by Liljeholm *et al.* (2015) attempts to fill this gap by studying human behavior and neural activity in a novel behavioral task while undergoing fMRI. Their task addresses a key confound in the typical practice of studying habit acquisition through overtraining (extensive experience with a given S–R pairing): overtraining conflates acquisition with the performance improvements that come with practice. Instead, the conditions in this experiment were designed to produce habits or goal-directed behaviors with minimal training. As expected based on the previous literature, following training habits proved less sensitive to a change in outcome contingencies in the task, referred to as devaluation. This design allowed the authors to distinguish neural correlates of S–R/R–O representations ‘early’ vs. ‘late’ in their development (putatively capturing acquisition vs. expression).

Using this task, the authors were able to replicate and extend many previous findings. Specifically, correlates of habit acquisition were found in the posterior caudate and cerebellum; whereas the former is to be expected, the latter is rarely reported in these types of studies. Similarly, correlates of habit expression in the ventral striatum and subgenual anterior cingulate cortex (ACC) both affirmed established knowledge (in the former case) and offered tantalizing clues to new functional relationships (in the latter case). Ventral striatal involvement accords with previous reports of its centrality to habitual control, but the role of subgenual ACC in this process remains a matter of much debate. Here, the authors suggest it as a functional homolog to rodent infralimbic cortex, a structure that has been implicated in both learning and expression of habits (Smith *et al.*, 2012; Smith & Graybiel, 2013).

Indeed, the question of cross-species homology remains a persistent source of tension in the instrumental control literature. The authors found that putamen activity decreased across the experiment, interpreted as suggesting a role in acquisition rather than expression. This appears to conflict with reports from perturbations of the rodent homolog, dorsolateral striatum (DLS; Furlong *et al.*, 2014). Both results are, however, consistent with the observation that DLS representations ‘sharpen’ with training (Smith & Graybiel, 2013) – and thus fMRI measures of population activity might decrease. An important question for future research is whether divergent results across species arise from interpretational nuances of different methodologies, or actual differences in the functional anatomy.

As this design is novel in the field, further work is needed to confirm that it faithfully distinguishes habitual and goal-directed behavior. One potential concern is that the current task may generate these behaviors by biasing attention towards or away from devaluation-relevant stimuli rather than by distinctly training S–R vs. R–O associations. Moreover, unlike classical manifestations of habitual behavior in the devaluation literature, habits in the current study were not necessarily maladaptive: giving the incorrect (devalued) response required completing a three-button sequence, sequences that were begun but rarely completed when subjects performed a (devaluation-insensitive) habit. Whether these features had bearing on the desired distinction between habits and goal-directed actions can be tested in future studies, for instance by using shorter response sequences and decoupling the focus of attention from the devalued target. Nevertheless, that these conditions distinguished behaviors and brain networks largely along expected lines should be seen as promising, if not yet conclusive, evidence that the manipulation was successful.

Finally, though this study begins to clarify a previously muddled distinction, it also hints at new divisions yet to be substantiated. In the computational reinforcement learning literature, the question arises whether goal-directed and habit systems operate largely in parallel, or

within an explicitly serial architecture with habits built by ‘training’ from the goal-directed system (Sutton, 1991). It remains to be seen which approach obtains in biology. The current observation of distinct neural correlates of acquisition and expression in both systems on its face seems to exclude a serial architecture, but on closer look the question remains open. It may be that certain regions perform functions intermediate to these two phases. One candidate midpoint structure is the ventral striatum. This area, widely reported in studies of habit learning and expression, is also modulated by the goal-directed system (Daw *et al.*, 2011) and may be necessary for some goal-directed behaviors (McDannald *et al.*, 2011). Liljeholm and colleagues found greater activity in this region for the S–R (vs. R–O) condition predicted increased habitual behavior – but if this region was active in both the acquisition and expression of R–O associations (and would therefore have been excluded from this paper’s analyses, which focused explicitly on circuits that were uniquely involved in one phase), then it may have acted as a site of learning from goal-directed representations. Future studies may build on the findings presented here to examine interactions between learning and expression, and whether interplay between goal-directed and habitual systems during learning impacts their relative prominence during expression.

Acknowledgement

This publication was made possible through the support of grants from the John Templeton Foundation and the CV Starr Foundation. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.

Abbreviations

ACC, anterior cingulate cortex; DLS, dorsolateral striatum; fMRI, functional magnetic resonance imaging; R–O, response–outcome; S–R, stimulus–response.

References

- Atallah, H.E., Lopez-Paniagua, D., Rudy, J.W. & O’Reilly, R.C. (2007) Separate neural substrates for skill learning and performance in the ventral and dorsal striatum. *Nat. Neurosci.*, **10**, 126–131.
- Balleine, B.W. & O’Doherty, J.P. (2010) Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology*, **35**, 48–69.
- Daw, N.D., Gershman, S.J., Seymour, B., Dayan, P. & Raymond, J. (2011) Model-based influences on humans’ choices and striatal prediction errors. *Neuron*, **69**, 1204–1215.
- Furlong, T.M., Jayaweera, H.K., Balleine, B.W. & Corbit, L.H. (2014) Binge-like consumption of a palatable food accelerates habitual control of behavior and is dependent on activation of the dorsolateral striatum. *J. Neurosci.*, **34**, 5012–5022.
- Gläscher, J., Daw, N., Dayan, P. & O’Doherty, J.P. (2010) States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, **66**, 585–595.
- Liljeholm, M., Molloy, C.J. & O’Doherty, J.P. (2012) Dissociable brain systems mediate vicarious learning of stimulus–response and action–outcome contingencies. *J. Neurosci.*, **32**, 9878–9886.
- Liljeholm, M., Dunne, S. & O’Doherty, J.P. (2015) Differentiating neural systems mediating the acquisition vs. expression of goal-directed and habitual behavioral control. *Eur. J. Neurosci.*, **41**, 1358–1371.
- McDannald, M.A., Lucantonio, F., Burke, K.A., Niv, Y. & Schoenbaum, G. (2011) Ventral Striatum and Orbitofrontal Cortex Are Both Required for Model-Based, But Not Model-Free, Reinforcement Learning. *J. Neurosci.*, **31**, 2700–2705.
- Ostlund, S.B. & Balleine, B.W. (2005) Lesions of medial prefrontal cortex disrupt the acquisition but not the expression of goal-directed learning. *J. Neurosci.*, **25**, 7763–7770.
- Smith, K. & Graybiel, A. (2013) A dual operator view of habitual behavior reflecting cortical and striatal dynamics. *Neuron*, **79**, 361–374.
- Smith, K.S., Virkud, A., Deisseroth, K. & Graybiel, A.M. (2012) Reversible online control of habitual behavior by optogenetic perturbation of medial prefrontal cortex. *Proc. Natl. Acad. Sci. USA*, **109**, 18932–18937.
- Sutton, R.S. (1991) Dyna, an integrated architecture for learning, planning, and reacting. *ACM SIGART Bull.*, **2**, 160–163.
- Tricomi, E., Balleine, B.W. & O’Doherty, J.P. (2009) A specific role for posterior dorsolateral striatum in human habit learning. *Eur. J. Neurosci.*, **29**, 2225–2232.