Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Behavioral Sciences

The multiple effects of practice: skill, habit and reduced cognitive load

Adrian M Haith[1] and John W Krakauer[1,2]

When learning a new skill, even if we have been instructed exactly what to do, it is often necessary to practice for hours, weeks or months before we achieve proficient and fluid performance. Practice has a multitude of effects on behavior, including increasing the speed of performance, rendering the practiced behavior habitual and reducing the cognitive load required to perform the task. These effects are often collectively referred to as *automaticity*. Here, we argue that these effects can be explained as multiple consequences of a single principle: caching of the outcome of frequently occurring computations. We further argue that, in the context of more complex task representations, caching different intermediate computations can give rise to more nuanced behavioral signatures, including dissociation between skill, habit and cognitive load.

Addresses
[1] Department of Neurology, Johns Hopkins University, Baltimore, MD, USA
[2] Department of Neuroscience, Johns Hopkins University, Baltimore, MD, USA

Corresponding author: Haith, Adrian M (adrian.haith@jhu.edu), Haith, Adrian M (adrian.haith@jhu.edu)

Current Opinion in Behavioral Sciences 2018, **20**:196–201

This review comes from a themed issue on **Habits and skills**

Edited by **Barbara Knowlton** and **Jörn Diedrichsen**

https://doi.org/10.1016/j.cobeha.2018.01.015

Introduction

The multiple effects of practice

Acquiring any new motor skill requires practice. It is not enough to simply be instructed how to drive a car or how to play a new video game; many hours of practice are typically needed to achieve proficiency and fluidity. What attributes of performance are improved through practice? Most obviously, being skilled at a task involves being able to select appropriate actions [1,2], and to execute those actions accurately [3–5]. However, a further critical aspect of skill relates to the *speed* at which an appropriate action can be selected. A novice driver will know to press the brake pedal to slow down, but an experienced driver will be far quicker to hit the brakes in the face of an unexpected hazard.

It has long been recognized that practice not only promotes incremental performance gains, but also leads to a qualitative change in how behavior is generated. A given task seems to require less and less cognitive effort the more we practice it [6–8]. Driving a car for the first time can feel overwhelming but, after sufficient practice, we have no problem talking to a passenger or listening to the radio while we drive. This familiar experience occurs in almost any skill we practice; as our proficiency increases, the *cognitive load* decreases.

The diminishing *cognitive load* of a task, has mostly been studied through the use of dual tasks [9]. In this approach, participants are asked to perform the practiced task at the same time as a cognitively demanding secondary task, such as counting the number of vowels heard in a sequence of spoken letters, or counting backwards from 100 in increments of 7. Early in learning, performance on either or both tasks suffers when they are attempted concurrently. After sufficient practice, however, it usually becomes possible to perform both tasks simultaneously just as well as they can be performed in isolation [10–12].

Another important phenomenon associated with practice is the formation of habits. Habits are most commonly studied in operant learning tasks in which an animal must learn through experience which action to perform to earn a reward (e.g. which lever triggers delivery of a food pellet). In this context, habitual behavior is typically defined as behavior that is insensitive to changes in the goals of a task [13–15] (and, by definition, opposite from *goal-directed* behavior). If a rat has repeatedly pressed a particular lever to earn a food reward, it may continue to press it habitually even when it is not hungry. Such habits are often exposed in daily life when the environment changes. For instance, when driving abroad, if the steering wheel is on the opposite side of the car, you may find yourself habitually reaching toward the door when trying to shift gears or pull the handbrake. The habitual nature of skilled typing is similarly unmasked if one tries to type on a foreign keyboard, in which certain symbols might be mapped onto different keys. This key-switch manipulation is in fact often directly used in experiments in humans to assess whether a practiced visuomotor association has become habitual [16,17,18••].

Assessing whether or not behavior is habitual can be complicated by the fact that behavior is known to be generated through a combination of habitual and goal-directed processes [14,19–21] and a learned habit can be easily masked by more deliberate, goal-directed influences on behavior. One way to reveal latent habitual behavior is to limit the amount of time available to generate a response to a stimulus. For instance, if participants practice distinguishing between different stimuli, or categories of stimuli, by pressing particular buttons over multiple days or weeks, habitual errors following a button-switch are relatively scarce when participants are allowed to respond under self-paced conditions [16]. But if participants are forced to respond very rapidly, habitual errors can be elicited in a majority of trials [18••]. Similar low-latency expression of habits occurs when using hand movements to control an on-screen cursor. If the mapping between hand and cursor is distorted, for example, through a mirror-reversal of the position of the cursor on the screen, participants can quickly learn to generate accurate movements when allowed to take their time before moving. However, rapid corrective responses to a perturbation applied during the movement betray a persistent habitual tendency to generate baseline patterns of correction, even after extensive practice [22–24]. Thus, habits are most strikingly revealed when actions must be generated very rapidly.

In summary, practice leads to three distinct changes in behavior: First, it improves skill level, including the ability to select actions more rapidly. Second, it permits appropriate actions to be selected with less cognitive effort than before. Third, it leads action selection to become habitual.

Cached computation as a theory of automatic action selection

These three effects of practice, skill, habit and reduced cognitive load, are often viewed as alternative ways of observing a single underlying change in behavior: that it has become *automatic* [25–27] (Figure 1a). It has often been assumed that it is sufficient to study any one of these phenomena on its own as a proxy for assessing the assumed underlying property of 'automaticity'. However, this assumption has remained largely untested; skill, habit and cognitive load have rarely been measured together in the same experiment. Aside from the fact that they are all altered through practice, why would one expect them to relate to one another?

In order to answer this question, it is worthwhile to first consider how an action is selected in the first place. It has been argued that motor control, much like any other task we might engage in, can be framed as a decision problem [2]. In order to select appropriate actions, we need to weigh up the relative costs and benefits of each option in order to determine which is best. This process, however, is generally computationally intensive.

In practice, we often tend to encounter the same decisions over and over again. If so, repeating the same computations and obtaining the same outcomes each time is needlessly time-consuming. Instead, it should be possible to simply remember which action you ended up selecting last time you encountered the same choice. This trick of storing the outcome of frequently occurring computations for rapid use later is often referred to as *caching*, by analogy with a computer cache — a memory reserve that is dedicated for high-speed retrieval.

Dispensing with time-consuming computations by using a cached representation ought to allow actions to be selected much faster, that is, improving *skill*. Furthermore, simply looking up the solution in a cache frees up the resources that would have been required to compute the solution, reducing the *cognitive load* associated with performing the task. Alongside these advantages, however, behavior will be liable to become inflexible; if the task changes in a way that warrants a different action being selected, information in the cache will remain outdated, leading to persistence of old patterns of behavior. Behavior therefore, will become *habitual* [14,21].

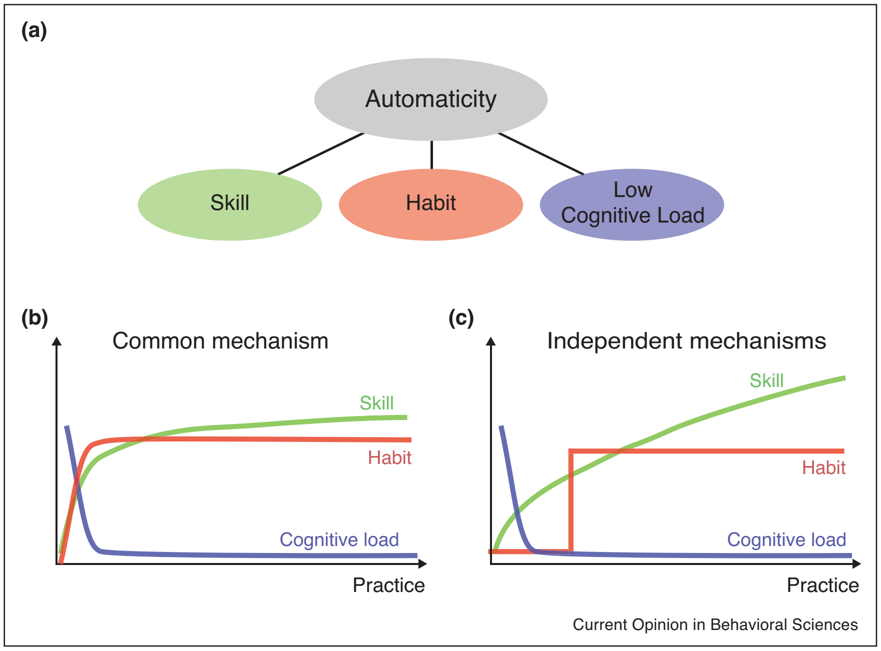
Importantly, the notion of caching is stronger than merely a process of retrieving earlier computations. Retrieval is, in general, a time-consuming process that itself can be improved through practice [28]. Caching should instead be understood as a means of storing direct associations between inputs and outputs in a way that is amenable to instant lookup. Retrieval can, in fact, be considered a form of cacheable computation and, indeed, many theories have argued that one-step retrieval should be considered a hallmark of automatic behavior [29].

The idea of cached versus computed action selection is related to a number of other important theories of learning (Box 1) (Table 1) [19]. We suggest that the principle of caching may be particularly valuable since it provides a potentially unifying and appealingly parsimonious theory of practice and automaticity: it accounts for improved speed of responding, reductions in cognitive load, and the tendency for behavior to become habitual. Considered in this way, the primary function of practice therefore might be to gather sufficient experience to establish this pre-computed cache. If so, changes in all three of these aspects should be closely correlated with one another (Figure 1b).

Hierarchical representations and intermediate computations

In principle, one can reduce any learned behavior, such as braking at a red light or hitting a tennis ball, to a single cached association between a stimulus and a response

Figure 1



Automaticity as a 'syndrome' of skill, habit and cognitive load. **(a)** Automaticity of well-practiced behaviors is typically characterized by a cluster of features: skill (faster performance), habit (behavior is expressed even if no longer appropriate) and low cognitive load (behavior can be produced with little attention or effort). **(b)** If these features are supported by a common underlying mechanism, they should exhibit a comparable time course. **(c)** Alternatively, if these facets of behavior are supported by independent mechanisms, they may follow dissociable time courses.

Box 1 Parallels with other learning dichotomies

The computed-versus-cached perspective is related to several other dichotomies that have been proposed in the context of learning [19] (Table 1). One such distinction is between declarative and procedural forms of memory. Since the time of H.M., the idea of declarative knowledge has come to be associated with dependence on the hippocampus and the medial temporal lobe [38]. In certain fields, the terms *declarative* and *procedural* have come to be associated with different types of learning that are either explicit or implicit [36]. However, the terms declarative and procedural originally derive from computer science, and theories of knowledge representation [39]. Declaratively represented knowledge is expressed in the form of facts about the world, which can be used to logically deduce answers to questions. Procedural knowledge, on the other hand, is expressed in terms of solutions to potential queries. Thus, a cached movement policy can be considered a form of procedural knowledge.

Another common dichotomy in theories of learning is the distinction between model-based and model-free reinforcement learning [19,21]. In *Model-based* reinforcement learning, individual agents maintain a rich model of the world, including the effects of their actions and availability of rewards, and use this *model* to compute the *value* (expected future reward) of each available state and/or action at a given point in time. In *model-free* reinforcement learning, by contrast, no such model is maintained. Instead, only the values of states and/or actions are stored. This model-free valuation can be considered a kind of cached representation. The key additional distinction of model-free reinforcement learning, however, is that it is associated with a means of updating these cached values directly, based on experience, rather than being informed by a model-based computation of value [40].

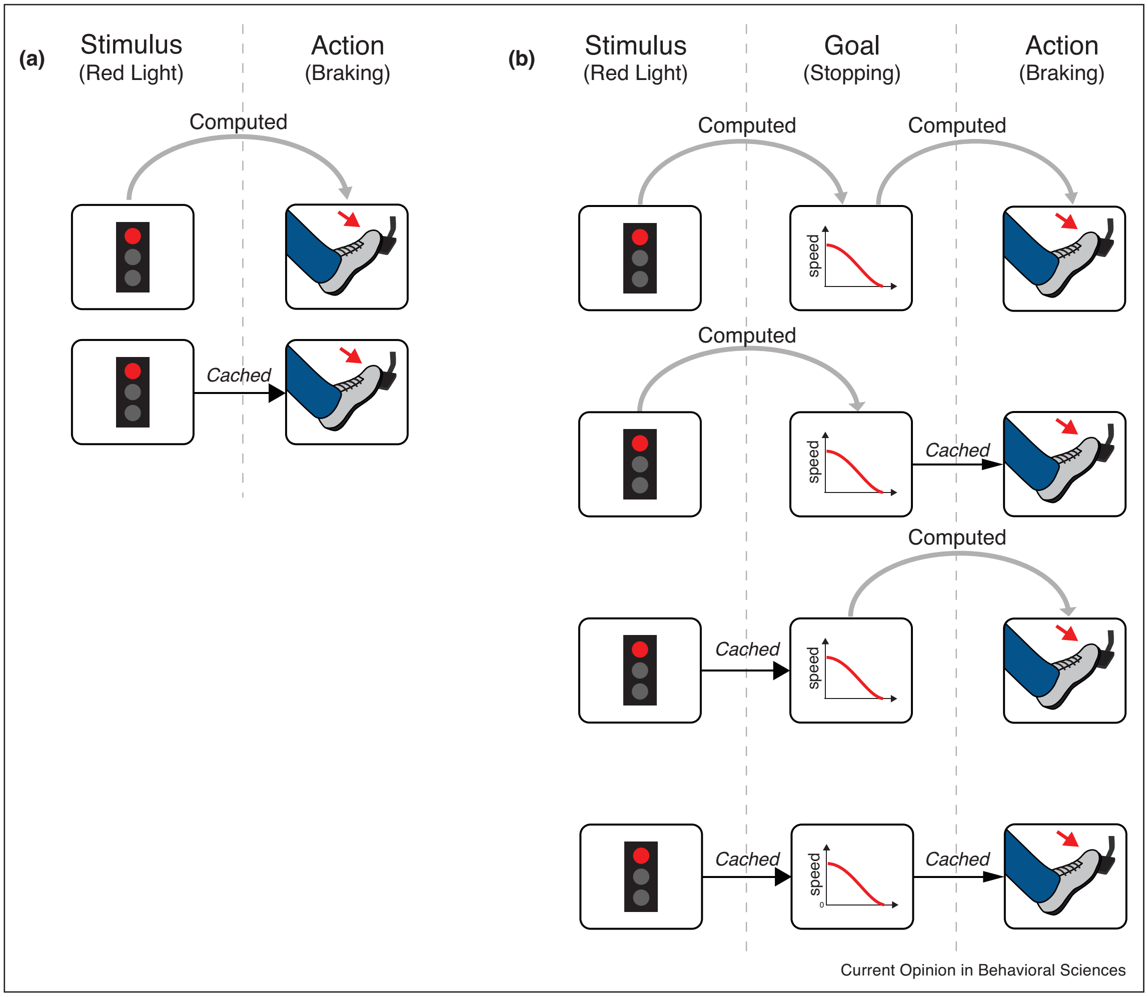
Table 1

Common dual-process dichotomies in learning

Defining feature	**Dominates early**	**Dominates late**
Response speed	Slow	Fast
Flexibility	Goal-directed	Habitual
Cognitive load	Effortful	Efficient
Evaluation	Computed	Cached
Representation of knowledge	Declarative	Procedural
Representation of value	Model-based	Model-free

(Figure 2a), which will naturally give rise to all three facets of the automaticity 'syndrome'. More generally, however, the process of determining an appropriate action may involve a more hierarchical process that entails intermediate computational steps. In the case of braking at a red light, seeing the light may prompt us to engage in the goal of stopping the car. We then select an appropriate action to satisfy this goal (Figure 2b). Both steps of this process might initially require time-consuming computations. With practice, however, either or both of these computational steps could, in principle, be cached.

Figure 2



Caching in the context of hierarchical selection. **(a)** In a simple stimulus-response scenario, only one computation is available to be cached. Consequently, skill, habit and cognitive load are likely to co-vary. **(b)** In a more complex scenario, action selection involves two distinct computations: firstly, behavioral goals must be determined from stimuli and secondly, actions must be selected to achieve these goals. Either of these computations could be cached, giving rise to a variety of ways that behavior might be generated. Each of these configurations may be associated with differing degrees of skill and with differing types of habit and cognitive load.

Furthermore, caching different steps will lead to different behavioral consequences that could potentially give rise to dissociations between skill, habit and automaticity (Figure 1c).

If you were to cache the association between the red light and the intermediate goal of stopping the car, you would habitually stop at red lights even if the rules of traffic lights were to change so that green now meant stop. But driving a car in which the brake and accelerator pedals were switched would pose no problem. Conversely, caching the association between stopping the car and braking would still allow you to be flexible to different traffic light rules, but would leave you unable to cope with a switch in pedals. Whichever computation is cached, skill level will improve due to faster responding. Thus it may be possible to obtain an apparent dissociation between skill and habit, depending on which specific notion of habit you have in mind.

These ideas are supported by findings from sequential decision-making tasks in which participants must select actions to navigate between states in pursuit of a reward. Cushman and Morris [30•] showed that human participants can be prone to habitual selection of which state they aim to eventually arrive at (based on which state historically yields a high reward). But participants were nevertheless able to engage in flexible goal-directed decisions to navigate towards that goal state. This behavior is consistent with caching the computation linking stimuli to goals in Figure 2b. Momennejad *et al.* [31••] observed behavior consistent with caching other intermediate computations. They found that participants could flexibly apply new information about which states earned reward to adjust their selection of high-level goals. However, they were less able to adapt their behavior following changes in which actions they needed to select in order to attain those goal states. This pattern of behavior is consistent with caching the computation linking goals to actions in Figure 2b. Thus each of the scenarios in Figure 2b seems plausible. Exactly which computations are cached likely depends on the nature of the task and how it is practiced.

The possibility of caching multiple component computations also has implications for cognitive load. Different component computations might depend on distinct cognitive resources (e.g. verbal working memory versus spatial working memory). Cognitive load might be affected by caching some computations more than others. Certain computations may be critical to the decision of which action to select, and caching this computation would render behavior habitual (depending, of course, on exactly which notion of habit you are interested in). Other computations might play a more auxiliary role that is nevertheless cognitively demanding and important for overall task accomplishment (for instance, deciding which stimuli to pay attention to or remembering which response options are available and how to execute them). Caching these latter computations would improve skill level and lighten cognitive load without necessarily making behavior habitual. Economides *et al.* [32•] showed just such an effect in a sequential decision task. Initially, performing a concurrent dual task interfered with goal-directed action selection. With practice, however, participants became able to select appropriate, goal-directed actions while also accurately performing the dual task. The dual-task improvement suggests that some computation had become cached, but not the critical computation that would have led to habitual behavior.

Conclusions

The principle of caching provides a compelling and parsimonious theory of automaticity following practice. Caching of a simple stimulus-response relationship accounts for the typically observed behavioral effects of practice: faster responding (skill), habitual behavior, and reduced cognitive load. Caching can, however, also be applied to intermediate steps of more complex computations, potentially giving rise to a multitude of different behavioral consequences.

This possibility leads to a richer and more nuanced view of automaticity view that raises the possibility that skill, habit and cognitive load might be dissociable (Figure 1c), depending on exactly which type of habit or cognitive load is being considered. If this is the case, it may help to account for why attempts to identify the neural basis of automaticity have yielded such disparate conclusions [10,12,33,34], which include claims that automaticity reflects recruitment of circuits in the cerebellum [10,35] or cortex [33], a shift in representation across subregions of the striatum [15,36], or a shift to cortical computations becoming independent of the basal ganglia altogether [37].

Ultimately, these considerations highlight the need for a richer understanding of the computations underlying action selection, which we suggest is most likely to be achieved through careful dissection of behavior.

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. Diedrichsen J, Kornysheva K: **Motor skill learning between selection and execution**. *Trends Cogn Sci* 2015, **19**:227-233.
2. Wolpert DM, Landy MS: **Motor control is decision-making**. *Curr Opin Neurobiol* 2012, **22**:996-1003.
3. Müller H, Sternad D: **Decomposition of variability in the execution of goal-oriented tasks: three components of skill improvement**. *J Exp Psychol Hum Percept Perform* 2004, **30**:212-233.
4. Reis J, Schambra HM, Cohen LG, Buch ER, Fritsch B, Zarahn E, Celnik PA, Krakauer JW: **Noninvasive cortical stimulation enhances motor skill acquisition over multiple days through an effect on consolidation**. *Proc Natl Acad Sci U S A* 2009, **106**:1590-1595.
5. Shmuelof L, Krakauer JW, Mazzoni P: **How is a motor skill learned? Change and invariance at the levels of task success and trajectory control**. *J Neurophysiol* 2012, **108**:578-594.
6. Anderson JR: **Acquisition of cognitive skill**. *Psychol Rev* 1982, **89**:369-406.
7. Fitts P: **Perceptual-motor skill learning**. *Categories of Human Learning*. Academic Press; 1964.
8. Taylor JA, Ivry RB: **The role of strategies in motor learning**. *Ann N Y Acad Sci* 2012, **1251**:1-12.
9. Pashler H: **Dual-task interference in simple tasks: data and theory**. *Psychol Bull* 1994, **116**:220-244.
10. Balsters JH, Ramnani N: **Cerebellar plasticity and the automation of first-order rules**. *J Neurosci* 2011, **31**:2305-2312.
11. Shiffrin RM, Schneider W: **Controlled and automatic human information processing: II. Perceptual learning, automatic attending and a general theory**. *Psychol Rev* 1977, **84**:127-190.
12. Wu T, Kansaku K, Hallett M: **How self-initiated memorized movements become automatic: a functional MRI study**. *J Neurophysiol* 2004, **91**:1690-1698.
13. Dickinson A: **Actions and habits: the development of behavioural autonomy**. *Philos Trans R Soc Lond B* 1985, **308**:67-78.
14. Dolan RJ, Dayan P: **Goals and habits in the brain**. *Neuron* 2013, **80**:312-325.
15. Yin HH, Knowlton BJ: **The role of the basal ganglia in habit formation**. *Nat Rev Neurosci* 2006, **7**:464-476.
16. Hélie S, Waldschmidt JG, Ashby FG: **Automaticity in rule-based and information-integration categorization**. *Atten Percept Psychophys* 2010, **72**:1013-1031.
17. Ashby FG, Ell SW, Waldron EM: **Procedural learning in perceptual categorization**. *Mem Cognit* 2003, **31**:1114-1125.

18. Hardwick RM, Forrence AD, Krakauer JW, Haith AM: **Skill acquisition and habit formation as distinct effects of practice**. *bioRxiv* 2017 http://dx.doi.org/10.1101/201095.
•• The authors propose a model of evolving competition between habitual and goal-directed action selection during the reaction time, and track how this is affected by practice.

19. Dayan P: **Goal-directed control and its antipodes**. *Neural Netw* 2009, **22**:213-219.

20. Gläscher J, Daw N, Dayan P, O'Doherty JP: **States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning**. *Neuron* 2010, **66**:585-595.

21. Daw ND, Niv Y, Dayan P: **Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control**. *Nat Neurosci* 2005, **8**:1704-1711.

22. Telgen S, Parvin D, Diedrichsen J: **Mirror reversal and visual rotation are learned and consolidated via separate mechanisms: recalibrating or learning de novo?** *J Neurosci* 2014, **34**:13768-13779.

23. Kasuga S, Telgen S, Ushiba J, Nozaki D, Diedrichsen J: **Learning feedback and feedforward control in a mirror-reversed visual environment**. *J Neurophysiol* 2015, **114**:2187-2193.

24. Gritsenko V, Kalaska JF: **Rapid online correction is selectively suppressed during movement with a visuomotor transformation**. *J Neurophysiol* 2010, **104**:3084-3104.

25. Cohen JD, Servan-Schreiber D, McClelland JL: **A parallel distributed processing approach to automaticity**. *Am J Psychol* 1992, **105**:239-269.

26. Hélie S, Cousineau D: **The cognitive neuroscience of automaticity: behavioral and brain signatures**. *Cogn Sci Hauppauge* 2011, **6**:35-53.

27. Moors A, De Hower J: **Automaticity: a theoretical and conceptual analysis**. *Psychol Bull* 2006, **132**:297-326.

28. Karpicke JD, Roediger HL: **The critical importance of retrieval for learning**. *Science* 2008, **319**:966-968.

29. Logan GD: **Skill and automaticity: relations, implications, and future directions**. *Can J Psychol Can Psychol* 1985, **39**:367-386.

30. Cushman F, Morris A: **Habitual control of goal selection in humans**. *Proc Natl Acad Sci U S A* 2015, **112**:13817-13822.
• The authors demonstrate the existence of a hierarchical task representation with habitual selection of goals, but deliberate and flexible pursuit of this goal.

31. Momennejad I, Russek EM, Cheong JH, Botvinick MM, Daw ND, Gershman SJ: **The successor representation in human reinforcement learning**. *Nat Hum Behav* 2017, **1**:680.
•• The authors show that human behavior in a reinforcement learning task is more flexible to changes in reward structure, than changes in state transitions. This behavior is consistent with caching predictions about future state occupancy.

32. Economides M, Kurth-Nelson Z, Lübbert A, Guitart-Masip M, Dolan RJ: **Model-based reasoning in humans becomes automatic with training**. *PLOS Comput Biol* 2015, **11**:e1004463.
• The authors demonstrate that an additional cognitive load initially renders behavior more habitual in a sequential decision task, but that this effect is eliminated following practice.

33. Grol MJ, Lange FPde, Verstraten FAJ, Passingham RE, Toni I: **Cerebral changes during performance of overlearned arbitrary visuomotor associations**. *J Neurosci* 2006, **26**:117-125.

34. Poldrack RA, Sabb FW, Foerde K, Tom SM, Asarnow RF, Bookheimer SY, Knowlton BJ: **The neural correlates of motor skill automaticity**. *J Neurosci* 2005, **25**:5356-5364.

35. Ramnani N: **Automatic and controlled processing in the corticocerebellar system**. *Prog Brain Res* 2014, **210**:255-285.

36. Ashby FG, Maddox WT: **Human category learning 2.0**. *Ann N Y Acad Sci* 2011, **1224**:147-161.

37. Helie S, Roeder JL, Ashby FG: **Evidence for cortical automaticity in rule-based categorization**. *J Neurosci* 2010, **30**:14225-14234.

38. Squire LR, Zola SM: **Structure and function of declarative and nondeclarative memory systems**. *Proc Natl Acad Sci U S A* 1996, **93**:13515-13522.

39. Winograd T: **Frame representations and the declarative/procedural controversy**. *Representation and Understanding: Studies in Cognitive Science*. 1975:185-210.

40. Sutton R, Barto A: *Reinforcement Learning: An Introduction*. MIT Press; 1998.