Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Behavioral Sciences

Reinforcement learning: bringing together computation and cognition

Anne Gabrielle Eva Collins

A key aspect of human intelligence is our ability to learn very quickly. This ability is still lacking in artificial intelligence. This article will highlight recent research showing how bringing together the fields of artificial intelligence and cognitive science may benefit both. Ideas from artificial intelligence have provided helpful formal theories to account for aspects of human learning. In return, ideas from cognitive science and neuroscience can also inform artificial intelligence research with directions to make algorithms more human-like. For example, recent work shows that human learning can only be understood in the context of multiple separate, interacting memory systems, rather than as a single, complex learner. This insight is starting to show promise in improving artificial agents' learning efficiency.

Address
Department of Psychology and Helen Wills Neuroscience Institute, University of California, 2121 Berkeley Way, Berkeley, CA 94720, United States

Corresponding author: Collins, Anne Gabrielle Eva (annecollins@berkeley.edu)

Current Opinion in Behavioral Sciences 2019, **29**:63–68

This review comes from a themed issue on **Artificial intelligence**

Edited by **Matt Botvinick** and **Sam Gershman**

https://doi.org/10.1016/j.cobeha.2019.04.011

From the birth of the field of cognitive science, the study of machine and human intelligence have been tightly linked. Indeed, many scientists believe that they can only convincingly claim to have understood a process when they can reproduce it. Thus, creating a machine with human intelligence is the ultimate test of our understanding of human cognition. A current frontier in both human and artificial intelligence research is in the domain of learning. Human intelligence is characterized by an ability to learn and adapt efficiently to new environments. However, despite remarkable progress in the last 10 years, this ability is still lacking in most artificial intelligence (AI) agents. Here I will argue that this gap only partially reflects our weak understanding of how humans perform this feat, and that more cross-fertilization between cognitive and AI research could help bridge the gap. I will first show how input from AI has been useful for the cognitive science of learning, then argue that AI could benefit from greater input from cognitive science, and finally I will highlight specific domains where this is currently possible.

In many ways, the study of the human mind and brain is at the root of the field of AI. First, the very definition of intelligence relies on our intuition that humans have cognitive capacities qualitatively beyond those of non-human animals. The field of AI attempts to mimic or surpass these abilities, which include many independent aspects, such as language and reasoning (which were some of the early targets of AI), but also learning and autonomous decision-making in complex, changing environments. In addition to providing AI with benchmark for Intelligence, the study of the human mind and brain also inspired some early attempts at artificial intelligence, with the field of neural networks ('connectionism') directly modeled from our understanding of neurons and their information coding properties.

However, the field of AI also largely developed its own goals, methods, and results, without reference to the parallel development of cognitive science. This independent progress has proven extremely beneficial to the understanding of human intelligence in general, and human learning in particular, exemplified by AI algorithms that could model aspects of animal behavior.

How AI supports cognitive research

The field of reinforcement learning is often highlighted as an archetype of the success of theoretical approaches to cognitive science. Computational processes designed by mathematicians to have theoretical guarantees, are imported to model how animals modify their behavior when experiencing rewarding or aversive outcomes. Reinforcement learning (RL) algorithms are a class of algorithms that have a narrow definition: they attempt to maximize a specific cost function, the discounted sum of future expected reward [1]. Such a function would clearly be important for animal survival, and as such, researchers hypothesized that such RL algorithms might be implemented in the brain. Indeed, some simple RL algorithms were found to be related to learning behaviors characterized as the 'law of effect', and summarized mathematically into a 'Delta-rule' in early cognitive models of Pavlovian conditioning [2], where better than expected outcomes lead animals (including humans) to increase the strength of an association between a predictor and an outcome. A family of simple RL algorithms called 'model-free RL', such as Q-learning [3], have been very successful to explain learning in simple or stimulus-action learning in humans and animals [4].

RL algorithms have been so successful in cognitive neuroscience because they not only capture behavior, but also provide a quantitative theory for the underlying neural processes. Indeed, information carried by the neuromodulator dopamine can be modeled as a reward prediction error signal consistent with the RL framework of temporal difference learning [1,5]. The striatum, where dopamine modulates plasticity [6], encodes the choice value or policy [7,8]; it has even been suggested that distinct parts of the striatum may have distinct roles, as proposed by an RL actor-critic model [9]. In short, there is broad agreement that a striatal-dopaminergic system implements a form of model-free RL in the brain.

This theory has been refined since to better account for human learning, often taking inspiration from AI. For example, model-based RL algorithms [10], Dyna-RL offline replay [11], and successor representations [12] account for aspects of planning in human learning [13•]. Other models that capture more advanced aspects of human learning include hierarchical RL [14], PID controllers [15], and so on. In short, the influence of RL algorithms from AI remains essential and fruitful in understanding human learning and intelligence.

How cognitive science can support AI research

The reverse seems to be less true: The tremendous progress in cognitive science and neuroscience has only weakly influenced AI research in RL, despite the fact that RL itself was strongly inspired by research in animal conditioning [16]. Nevertheless, here are multiple ways in which research on human cognition can inform AI. First, it can provide tools for analyzing and thus improving agents' behavior. More importantly, it can provide inspiration for better algorithms by showing how humans perform tasks that artificial agents fail at. Indeed, hundreds of millions of years of evolution have sculpted highly complex, efficient, and effective nervous systems in the animal kingdom, and research on human behavior in particular reveals powerful natural learning algorithms. I will show examples of successful integration of knowledge from the cognitive sciences into AI, and suggest some directions where cognitive research is ahead of AI.

Tools

An important frontier in creating artificial agents is in fast learning, which in humans relies on the ability to transfer previous knowledge to new environments or problems. Despite recent progress in AI, humans still have a strong advantage in this domain, and most AI studies include human performance as a benchmark [17,18••]. This performance benchmark is often a rough aggregate measure of how well the agent is doing, for example, the number of points earned in a game. Cognitive scientists have already developed and can provide tools to make this more informative: First, they can provide more reliable benchmarks, by establishing a range of target performance across multiple human players, carefully controlling experimental factors (Tsividis *et al.*, unpublished). More importantly, analyzing fine-grained patterns of behavior, rather than aggregate measures, is an essential tool for cognitive scientists trying to deconstruct the algorithms that drive learning [19,20]. Such methods would also be informative to dissect artificial agents' behavior [21]. We could then better identify where specifically they fail to match human performance, and thus inform their improvement.

Inspiration

A more fundamental way, in which cognitive science can inform AI research, is by inspiring improved algorithms [22••]. Again, there is an evolutionary argument here: learning efficiently and adapting flexibly to changing environments is essential for animals' fitness and survival through evolution; as such, it is likely (though not certain) that the algorithms implemented by the brain for learning and decision-making are fine-tuned to support efficient learning. A strategy for improving AI in domains in which it currently lags behind human levels, such as fast learning and generalization, is to take what we understand from human cognition in that domain and apply it to artificial agents. In the next paragraphs, I will give examples where this has been done, and examples where this could be done more.

Multiple learning systems in parallel

Early improvements of AI algorithms for learning, such as deep reinforcement learning networks (e.g. DQN [17]) were successful, to a first approximation, by creating elaborate state spaces over which simple RL algorithms could operate using function approximation. However, a key insight from human cognitive neuroscience is that human learning can never be approximated by a single learning system, no matter how clever. Instead, at any time point, multiple learning, memory, and decision processes contribute to learning and choices in parallel, and sometimes interact with one another.

One example is *episodic memory*. In addition to the well-studied neural RL system, humans, use hippocampus-dependent memory to store, and recall when relevant, unique, precise events. Recent research has shown that the episodic memory system contributes essentially to learning [23,24,25••,26,27], in parallel with RL. This finding has inspired new AI agents. Recent research targeted such one-shot learning (putatively dependent on episodic memory) and developed deep-RL agents that were augmented with an ability to use external memory for fast learning [28]. These agents were able to perform one-shot imitation learning, taking them a step closer to human learning abilities [29,30].

Another cognitive process that crucially contributes to human learning is *working memory*, a process by which we actively hold in mind a limited amount of information for a short amount of time. Evidence for short-term maintenance of information, and sensitivity to information load, show that much of human learning is more dependent on working memory (WM) than RL [31–33]. Furthermore, recent work shows that the use of WM in learning interferes with computations performed by the RL process. Specifically, model-free RL algorithms are usually assumed to be closed-loop: they use their own predictions to estimate reward prediction errors and learn from them. Instead, recent work hints at more dependency between different systems, whereby WM provides inputs to RL computations [11,34,35•,36••].

The understanding of the role of WM in human learning has not inspired AI research yet, to our knowledge. Recurrent networks such as LSTMs [37] bear some resemblance to WM in that they keep a trace of past information in the network's state. However, this is a superficial resemblance: crucial characteristics of WM, both positive ones, such as the ability to manipulate information in WM, and apparently negative ones, such as the very limited amount of information it can hold, are not present [28]; these artificial agents instead have episodic-memory-like long-lasting, high-capacity memory stores. We should again assume that WM's characteristics may have an evolutionary purpose, and thus might improve artificial agents. A possibility is that WM's limited information capacity focuses attention and learning to a manageable, low-dimensional, prioritized state space [38]. Similarly, interactions between memory systems such as WM and RL may have beneficial computational properties that could improve AI agents. This type of interactions between memory systems is currently being explored for episodic memory and RL [39].

Executive functions provide structure for learning

Humans, as well as non-human primates and other animals, can learn to learn [40]. This ability is dependent on prefrontal cortex [41], and holds potential for improvement of artificial agents: recent work showed that when an agent is trained on families of problems that share the same structure, it can learn how to learn in these types of situations and become more and more efficient over time [42••,43].

An important way, in which humans learn to learn efficiently, is by integrating prior knowledge with new information [44,45]. Making intelligent agents will probably require an ability to integrate such priors into their computations, an area of active research [18••]. This ability manifests itself in humans by their tendency to search for structure in their interaction with the environment [46–49]. Humans even create structured representations that do not reflect objective structure in the environment (and thus provides no behavioral benefit), at a cost to their immediate performance [50–52]. This apparently suboptimal strategy is present not only in human adults, who may have extensive prior evidence that this is a useful long-term strategy, but also in human infants [53,54], highlighting an essential ingredient to the development of human intelligence.

One computational benefit of structure building in humans is that it provides representations for learned policies that are not strongly tied to a specific set of sensory inputs and motor outputs. Instead, from the start, policies are created to be broadly generalizable and transferable: A key behavioral marker of such structure learning in humans is the ability to later transfer and generalize learned knowledge and policies to new environments [55]. We learn how to behave appropriately in different contexts, reacting to the same cues differently: for example, we react to a colleague's statements or our phone ringing differently in an office during an interview and in a bar after a conference. Having learned these context-dependent policies, we can immediately reuse them in a different context (new office, different conference), or with different low-level choices (greeting someone in French or English, turning off a different phone). The ability to create reusable policies is related to the notion of 'state abstraction for lifelong RL' in AI [56]. However, it is unclear whether this research allows artificial agents to incur short-term costs for explicitly building more complex than necessary representations, in the hope that those will prove their worth in future interactions with new environments. The structure of human brain networks actively promotes structure creation [57], making us short-term suboptimal for increased long-term fitness and flexibility.

How neuroscience can inform AI research

This last point highlights the fact that inspiration for AI can be found not only by investigating human behavior, but also by understanding how the brain implements it. This benefit could come in one of two ways, often tightly intertwined. First, we investigate the 'hardware', and ask how brain networks that support a specific behavior are organized, assuming that evolutionary pressure may have constrained these processes to compress information efficiently. This approach has been successful in the domain of artificial vision for example, where convolutional networks were inspired by neuroscience [58•]. Second, we can investigate the 'software': how information related to a cognitive process is encoded.

As an example of learning about the brain's 'software', we can look at dopaminergic signals that encode reward prediction errors in the brain, which support policy learning. Finding non-reward related information that contributes to dopaminergic signals (such as novelty, uncertainty, or information [59]) suggests that this information is usefully integrated into a policy learning in humans, and thus potentially also for artificial agents. Similarly, recent findings about hippocampal replay, preplay, and mental simulation [60,61•], could inform how Dyna-like off-line learning [10] occurs and is prioritized in animals, and thus might inspire AI.

Regarding 'hardware', the organization of brain networks could also inspire how to structure artificial networks to support flexible learning. For example, the brain uses two separate but converging pathways for RL, coding highly anticorrelated information (a 'positive' policy: how much to approach a choice; and a 'negative' policy; how much to avoid a choice). Such redundancy is surprising in a biological system. Thus, it may instead reflect an adaptive role for flexible behavior, such as providing a method to flexibly adjust whether to prioritize costs versus benefits in decisions [62]. As another example, the brain networks that support RL computations exist in multiple parallel loops, originating in cortical regions that represent more or less abstract information [63–65]. This structure might indicate the usefulness of performing RL computations on multiple state/action spaces in parallel at different granularities of generalizability, with largely independent, but hierarchical dependencies [50].

Discussion

It is important to keep in mind that the human and artificial intelligence communities have different goals, and that what works *in silico* is not guaranteed to work *in vivo*, and vice-versa. There may be risks in trying to match too closely human and artificial intelligence. First, AI agents have different memory mechanisms and resources than animals, which could mean that algorithms designed to work within animal constraints might not be optimal for AI agents. In reverse, AI research may produce elegant and efficient frameworks that do not necessarily have to relate to human learning. Seeking to take a mechanism that is relevant for artificial agents and using it as a lens for investigating human learning might be productive, but also misleading, by artificially enforcing an interpretation that might not correspond to computations in brain and behavior. Both communities should remain aware of these drawbacks.

Nevertheless, here I have shown that research in human and artificial learning could benefit from more cross-talk between the two disciplines. Research in other domains, especially sensory processing, has already shown such bidirectional benefits by inspiring network structures (e.g. convolutional networks), and by helping model neural processing [58•]. The same process could be strengthened for the field of learning and decision making. To this end, researchers in AI might take inspiration from cognitive scientists' knowledge of human brain and behavior, and try to integrate it into better agents. For example, network structures could be informed by knowledge from brain connectivity. Algorithms could be constrained with knowledge about human learning, such as the existence of multiple parallel, interacting memory and decision-making processes. Researchers in human intelligence should also take inspiration from algorithms that improve artificial agents, and consider testing whether they might account for some aspects of human learning. This requires designing experiments that take into account the complexities of real-world learning, and thus probing intelligent learning more directly.

Conflict of interest statement

Nothing declared.

Acknowledgements

I am grateful to Maria Eckstein, Sarah Master, Sam McDougle and Victor Shnayder for helpful comments and suggestions on an early version of this draft. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

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. Sutton RS, Barto AG: *Reinforcement Learning*. MIT Press; 1998.
2. Wagner A, Rescorla R: **Inhibition in Pavlovian conditioning: application of a theory**. *Inhib Learn* 1972:301-336.
3. Watkins CJCH, Dayan P: **Q-learning**. *Mach Learn* 1992, **8**:279-292.
4. Huys QJM: **Bayesian approaches to learning and decision-making**. *Comput Psychiatry* 2018:247-271.
5. Montague PR, Dayan P, Sejnowski TJ: **A framework for mesencephalic dopamine systems based on predictive Hebbian learning**. *J Neurosci* 1996, **16**:1936-1947.
6. Reynolds JN, Hyland BI, Wickens JR: **A cellular mechanism of reward-related learning**. *Nature* 2001, **413**:67-70.
7. Samejima K, Ueda Y, Doya K, Kimura M: **Representation of action-specific reward values in the striatum**. *Science* 2005, **310**:1337-1340.
8. Tai L-H, Lee AM, Benavidez N, Bonci A, Wilbrecht L: **Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value**. *Nat Neurosci* 2012, **15**:1281-1289.
9. Joel D, Niv Y, Ruppin E: **Actor-critic models of the basal ganglia: new anatomical and computational perspectives**. *Neural Netw* 2002, **15**:535-547.
10. Gershman SJ, Markman AB, Otto AR: **Retrospective revaluation in sequential decision making: a tale of two systems**. *J Exp Psychol Gen* 2014, **143**:182-194.
11. Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ: **Model-based influences on humans' choices and striatal prediction errors**. *Neuron* 2011, **69**:1204-1215.

12. Dayan P: **Improving generalization for temporal difference learning: the successor representation**. *Neural Comput* 1993, **5**:613-624.

13. • Momennejad I *et al.*: **The successor representation in human reinforcement learning**. *Nat Hum Behav* 2017, **1**:680-692.
Momennejad *et al.* exemplify how cognitive science can learn from AI, showing that AI's successor representation framework accounts for some aspects of human learning.

14. Botvinick MM, Niv Y, Barto AC: **Hierarchically organized behavior and its neural foundations: a reinforcement-learning perspective**. *Cognition* 2009, **113**:262-280.

15. Ritz H, Nassar MR, Frank MJ, Shenhav A: **A control theoretic model of adaptive learning in dynamic environments**. *J Cogn Neurosci* 2018, **30**:1405-1421.

16. Sutton RS, Barto AG: **Toward a modern theory of adaptive networks: expectation and prediction**. *Psychol Rev* 1981, **88**:135-170.

17. Mnih V *et al.*: **Human-level control through deep reinforcement learning**. *Nature* 2015, **518**:529-533.

18. •• Lake BM, Ullman TD, Tenenbaum JB, Gershman SJ: **Building machines that learn and think like people**. *Behav Brain Sci* 2017, **40**:e253.
Lake *et al.* propose a framework for making artificial machines that think like humans, focusing on a few concrete directions to overcome current challenges.

19. Krakauer JW, Ghazanfar AA, Gomez-Marin A, Maciver MA, Poeppel D: **Neuron perspective neuroscience needs behavior: correcting a reductionist bias**. *Neuron* 2017, **93**:480-490.

20. Collins A, Frank M: **Cognitive control over learning: creating, clustering, and generalizing task-set structure**. *Psychol Rev* 2013, **120**:190.

21. Ritter S, Barrett DGT, Santoro A, Botvinick MM: *Cognitive Psychology for Deep Neural Networks: A Shape Bias Case Study*. 2017.

22. •• Hassabis D, Kumaran D, Summerfield C, Botvinick M: **Neuroscience-inspired artificial intelligence**. *Neuron* 2017, **95**:245-258.
Hassabis *et al.* highlight how findings from neuroscience can, and have, inspired better AI algorithms.

23. Vicario-feliciano R, Murray EA, Averbeck BB: **Ventral striatum lesions do not affect reinforcement learning with deterministic outcomes on slow time scales**. *Behav Neurosci* 2017, **131**:385-391.

24. Poldrack Ra *et al.*: **Interactive memory systems in the human brain**. *Nature* 2001, **414**:546-550.

25. •• Bornstein AM, Norman KA: **Reinstated episodic context guides sampling-based decisions for reward**. *Nat Neurosci* 2017, **20**:997-1003.
Bornstein *et al.* show that part of what is often interpreted as RL in humans is actually mediated by recall from episodic memory.

26. Bornstein AM, Khaw MW, Shohamy D, Daw ND: **Reminders of past choices bias decisions for reward in humans**. *Nat Commun* 2017, **8**:1-9.

27. Wimmer GE, Li JK, Gorgolewski KJ, Poldrack RA: **Reward learning over weeks versus minutes increases the neural representation of value in the human brain**. *J Neurosci* 2018, **38** http://dx.doi.org/10.1523/JNEUROSCI.0075-18.2018.

28. Graves A *et al.*: **Hybrid computing using a neural network with dynamic external memory**. *Nature* 2016, **538**:471-476.

29. Duan Y *et al.*: *One-Shot Imitation Learning*. 2017:1087-1098.

30. Santoro A, Bartunov S, Botvinick M, Wierstra D, Lillicrap T: *One-Shot Learning with Memory-Augmented Neural Networks*. 2016.

31. Collins AGE, Frank MJ: **How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis**. *Eur J Neurosci* 2012, **35**:1024-1035.

32. Collins AGE, Albrecht MA, Waltz JA, Gold JM, Frank MJ: **Interactions among working memory, reinforcement learning, and effort in value-based choice: a new paradigm and selective deficits in schizophrenia**. *Biol Psychiatry* 2017, **82**:431-439.

33. Viejo G, Girard B, Procyk E, Khamassi M: **Adaptive coordination of working-memory and reinforcement learning in non-human primates performing a trial-and-error problem solving task**. *Behav Brain Res* 2018, **355**:76-89.

34. Collins AGE, Ciullo B, Frank MJ, Badre D: **Working memory load strengthens reward prediction errors**. *J Neurosci* 2017, **37**:2700-2716.

35. • Collins AGE, Frank MJ: **Within- and across-trial dynamics of human EEG reveal cooperative interplay between reinforcement learning and working memory**. *Proc Natl Acad Sci U S A* 2018, **115**:2502-2507.
Collins and Frank show that human learning needs to be understood as a mixture of partially independent, partially interacting computational processes.

36. •• Collins AGE: **The tortoise and the hare: interactions between reinforcement learning and working memory**. *J Cogn Neurosci* 2018:1-12.
Collins shows that part of what is often interpreted as RL in humans is actually mediated by working memory, and that working memory use weakens RL function.

37. Hochreiter S, Schmidhuber J: **Long short-term memory**. *Neural Comput* 1997, **9**:1735-1780.

38. Todd MT, Niv Y, Cohen JD: **Learning to use working memory in partially observable environments through dopaminergic reinforcement**. *Neural Inf Process Syst* 2009:1689-1696.

39. Ritter S, Wang J, Kurth-Nelson Z, Botvinick M: **Episodic control as meta-reinforcement learning**. *bioRxiv* 2018 http://dx.doi.org/10.1101/360537.

40. Harlow HF: **The formation of learning sets**. *Psychol Rev* 1949, **56**:51.

41. Bhandari A, Badre D: **Learning and transfer of working memory gating policies**. *Cognition* 2018, **172**:89-100.

42. •• Wang JX *et al.*: **Prefrontal cortex as a meta-reinforcement learning system**. *Nat Neurosci* 2018, **21**:860-868.
Wang *et al.* exemplify how AI can learn from neuroscience, linking artificial agents' ability in meta-learning to prefrontal function.

43. Duan Y *et al.*: *RL^2: Fast Reinforcement Learning via Slow Reinforcement Learning*. 2016.

44. Tervo DGR, Tenenbaum JB, Gershman SJ: **Toward the neural implementation of structure learning**. *Curr Opin Neurobiol* 2016, **37**:99-105.

45. Tenenbaum JB, Kemp C, Griffiths TL, Goodman ND: **How to grow a mind: statistics, structure, and abstraction**. *Science* 2011, **331**.

46. Collins A, Koechlin E: **Reasoning, learning, and creativity: frontal lobe function and human decision-making**. *PLoS Biol* 2012, **10**:e1001293.

47. Donoso M, Collins AGE, Koechlin E: **Foundations of human reasoning in the prefrontal cortex**. *Science (80-)* 2014, **344**:1481-1486 http://dx.doi.org/10.1126/science.1252254.

48. Badre D, Kayser AS, Esposito MD: **Article frontal cortex and the discovery of abstract action rules**. *Neuron* 2010, **66**:315-326.

49. Franklin NT, Frank MJ: **Compositional clustering in task structure learning**. *PLoS Comput Biol* 2018, **14**:1-25.

50. Collins AGE, Frank MJMJ: **Cognitive control over learning: creating, clustering, and generalizing task-set structure**. *Psychol Rev* 2013, **120**:190-229.

51. Collins AGE: **The cost of structure learning**. *J Cogn Neurosci* 2017, **29**:1646-1655.

52. Collins AGE, Cavanagh JF, Frank MJ: **Human EEG uncovers latent generalizable rule structure during learning**. *J Neurosci* 2014, **34**:4677-4685.

53. Werchan DM, Collins AGE, Frank MJ, Amso D: **Role of prefrontal cortex in learning and generalizing hierarchical rules in 8-month-old infants**. *J Neurosci* 2016, **36**:10314-10322.

54. Werchan DM, Collins AGE, Frank MJ, Amso D: **8-month-old infants spontaneously learn and generalize hierarchical rules**. *Psychol Sci* 2015, **26**:805-815.

55. Collins AGE, Frank MJ: **Neural signature of hierarchically structured expectations predicts clustering and transfer of rule sets in reinforcement learning**. *Cognition* 2016, **152** http://dx.doi.org/10.1016/j.cognition.2016.04.002.

56. Abel D, Arumugam D, Lehnert L, Littman M: **State abstractions for lifelong reinforcement learning**. *Proceedings of the 35th International Conference on Machine Learning* 2018, **80**:10-19.

57. Collins AGE, Frank MJ: **Motor demands constrain cognitive rule structures**. *PLOS Comput Biol* 2016, **12**:e1004785.

58. Yamins DLK, DiCarlo JJ: **Using goal-driven deep learning • models to understand sensory cortex**. *Nat Neurosci* 2016, **19**:356-365.
Yamins and DiCarlo exemplify how deep neural networks can be leveraged to better understand the brain's computations.

59. Bromberg-Martin ES, Hikosaka O: **Midbrain dopamine neurons signal preference for advance information about upcoming rewards**. *Neuron* 2009, **63**:119-126.

60. Ambrose RE, Pfeiffer BE, Foster DJ: **Reverse replay of hippocampal place cells is uniquely modulated by changing reward**. *Neuron* 2016, **91**:1124-1136.

61. Foster DJ: **Replay comes of age**. *Annu Rev Neurosci* 2017, • **40**:581-602.
Foster reviews the important role of replay-like phenomena in human memory, and the role they may play in flexible learning and decision making.

62. Collins AGE, Frank MJ: **Opponent actor learning (OpAL): modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive**. *Psychol Rev* 2014, **121**:337-366.

63. Alexander GE, DeLong MR, Strick PL: **Parallel organization of functionally segregated circuits linking basal ganglia and cortex**. *Ann Rev Neurosci* 1986, **9**.

64. Summerfield C, Koechlin E: **A neural representation of prior information during perceptual inference**. *Neuron* 2008, **59**:336-347.

65. Badre D, D'Esposito M: **Functional magnetic resonance imaging evidence for a hierarchical organization of the prefrontal cortex**. *J Cogn Neurosci* 2007, **19**:2082-2099.