Available online at www.sciencedirect.com

ScienceDirect

Current Opinion in Neurobiology

Where learning paths meet: Convergence and divergence of statistical and reinforcement learning

Ambra Ferrari[1], Floris P. de Lange[2] and Athena Akrami[3]

Learning enables organisms to adapt to a dynamic world by forming and updating internal representations of their environment. Statistical Learning (SL) and Reinforcement Learning (RL) offer complementary perspectives on this process. RL is fundamentally goal-directed, focused on maximizing rewards through Reward Prediction Error (RPE). SL extracts the statistical structure of the environment without explicit instruction or reinforcement. Model-Based RL additionally incorporates State Prediction Error (SPE) to refine an internal model of the world, overlapping with SL, which may use SPE or associative mechanisms devoid of error computations to extract structure. Neurobiologically, current research shows that RL is linked to midbrain dopaminergic signaling, whereas SL is supported by cortical and subcortical networks including early sensory areas and the hippocampus. This review compares RL and SL across their historical foundations, objectives, computational principles, and neural implementation, suggesting ways to better delineate the boundaries and interconnections between these two fundamental forms of learning.

Addresses
[1] CIMeC, Center for Mind/Brain Sciences, University of Trento, 38068, Rovereto, Italy
[2] Radboud University Nijmegen, Donders Institute for Brain, Cognition and Behaviour, 6525 EN, Nijmegen, the Netherlands
[3] Sainsbury Wellcome Centre, University College London, W1T 4JG, London, United Kingdom

Corresponding authors: Ferrari, Ambra (ambra.ferrari@unitn.it); Akrami, Athena (athena.akrami@ucl.ac.uk)

Current Opinion in Neurobiology 2026, **98:**103181

This review comes from a themed issue on **Systems Neuroscience 2025**

Edited by **Robert Froemke**, **Máté Lengyel**, **József Fiser**, and **Livia de Hoz**

For complete overview of the section, please refer the article collection - Systems Neuroscience 2025

Available online 9 March 2026

https://doi.org/10.1016/j.conb.2026.103181

0959-4388/

Introduction

Learning is an essential feat of cognition, enabling animals to adapt flexibly to an ever-changing world by forming internal representations of their environment and using them to guide goal-directed behavior. These representations, shaped by prior experience, influence how information is processed and decisions are made, fundamentally impacting behavior [1,2]. Yet, the mechanisms that allow us to construct, refine, and update these internal models remain an open and active area of research. Among the diverse approaches to studying learning, two prominent frameworks—statistical learning (SL) and reinforcement learning (RL)—have emerged, offering complementary perspectives on how we internalize and respond to the structure of the world.

RL is a well-established field with a rich literature [3]. Briefly, the computational goal of RL is to discover relationships between states (sensory events and possibly their associated actions) and outcomes (rewards or punishments). To this aim, the observer computes the mismatch between actual and expected outcome (i.e. reward prediction error, RPE) or between actual and expected state (i.e. state prediction error, SPE [4]), or action (Action prediction error, APE [5]). This can be achieved through different algorithms that vary in computational complexity and flexibility [6]. Regarding neural implementation, mounting evidence suggests that midbrain dopamine neurons may signal RPEs [7–9] (but see Ref. [10]), or APEs [5].

Conversely, research on the computational principles, algorithmic mechanisms and neural implementation of SL is more equivocal. One factor hampering progress may be the lack of consensus on the definition of SL. Some have narrowly defined it as the parsing of the temporal structure in continuous streams of stimuli, based on early work on grammar learning in language acquisition [11–13]. More recent accounts prefer a broader view, where SL pertains to discovering statistical properties of and associations between environmental states [14,15]. Although SL and RL differ in their core objectives, i.e. SL aims to model environmental structure, while RL seeks to optimize actions for reward, they may converge on a partially overlapping computational goal: constructing internal models of the environment. We propose that RL may strongly benefit from an accurate world model internalized via SL, improving the learning of stimulus-action-outcome associations and supporting more flexible, model-based behavior.

Brief history and key characteristics of statistical and reinforcement learning

The roots of SL can be traced back to the mid-20th century [16]: the concept of "perceptual learning" was introduced by Gibson and Gibson as a primitive process where organisms become sensitive to information already inherent in their sensory input. During the 1960s, the term "implicit learning" gained traction, particularly within linguistics exploring artificial grammar learning, where participants learned to distinguish between grammatical and ungrammatical strings of symbols without being explicitly told the underlying rules. Researchers began investigating how individuals acquire knowledge without conscious awareness or explicit instruction [11,17]. The similarities between Gibsonian perceptual learning [16] and the implicit learning observed in these studies were quickly recognized. Both processes emphasize the extraction of statistical regularities from environmental input, suggesting a common underlying mechanism. This shared foundation led to the common term "statistical learning". In the subsequent decades, research on SL moved beyond the realm of linguistics to encompass a wide range of domains, including visual and auditory perception [15,18,19], music [20] and motor learning [21]. Collectively, research converges on the idea that SL is a process aimed at building internal models that recapitulate the statistical properties of the environment [15]. These models help understand relationships between variables, make sense of complex structures, and predict future states.

Parallel to the evolution of SL research, two key threads shaped the foundations of RL in the mid-20th century. The first stemmed from studies of animal learning through experience and trial-and-error, exemplified by Thorndike's Law of Effect [22], which proposed that animals are more likely to repeat actions that lead to satisfying outcomes. The second thread focused on the mathematical formalization of decision-making under uncertainty, marked by Bellman's Dynamic Programming [23], introducing the concept of value functions and optimal policies. By the late 20th century, researchers began to combine these ideas leading to various computational algorithms (see Section "Computational principles"). The integration of these computational frameworks with neuroscience gained momentum in the 1990s. In particular, it was discovered that midbrain dopaminergic neurons encode RPEs by modulating their firing rates according to positive or negative deviations between expected and received outcomes [7]. Demonstrating the neural implementation of RPEs provided a direct link between the computational and neurobiological levels.

Despite their independent histories, SL and RL exhibit a fundamental similarity: they use experience to discover associations among events (which could pertain to stimuli, actions, and outcomes). This poses a crucial question: are SL and RL fundamentally the same form of learning, or are there meaningful differences between them?

Divergent goals of statistical and reinforcement learning

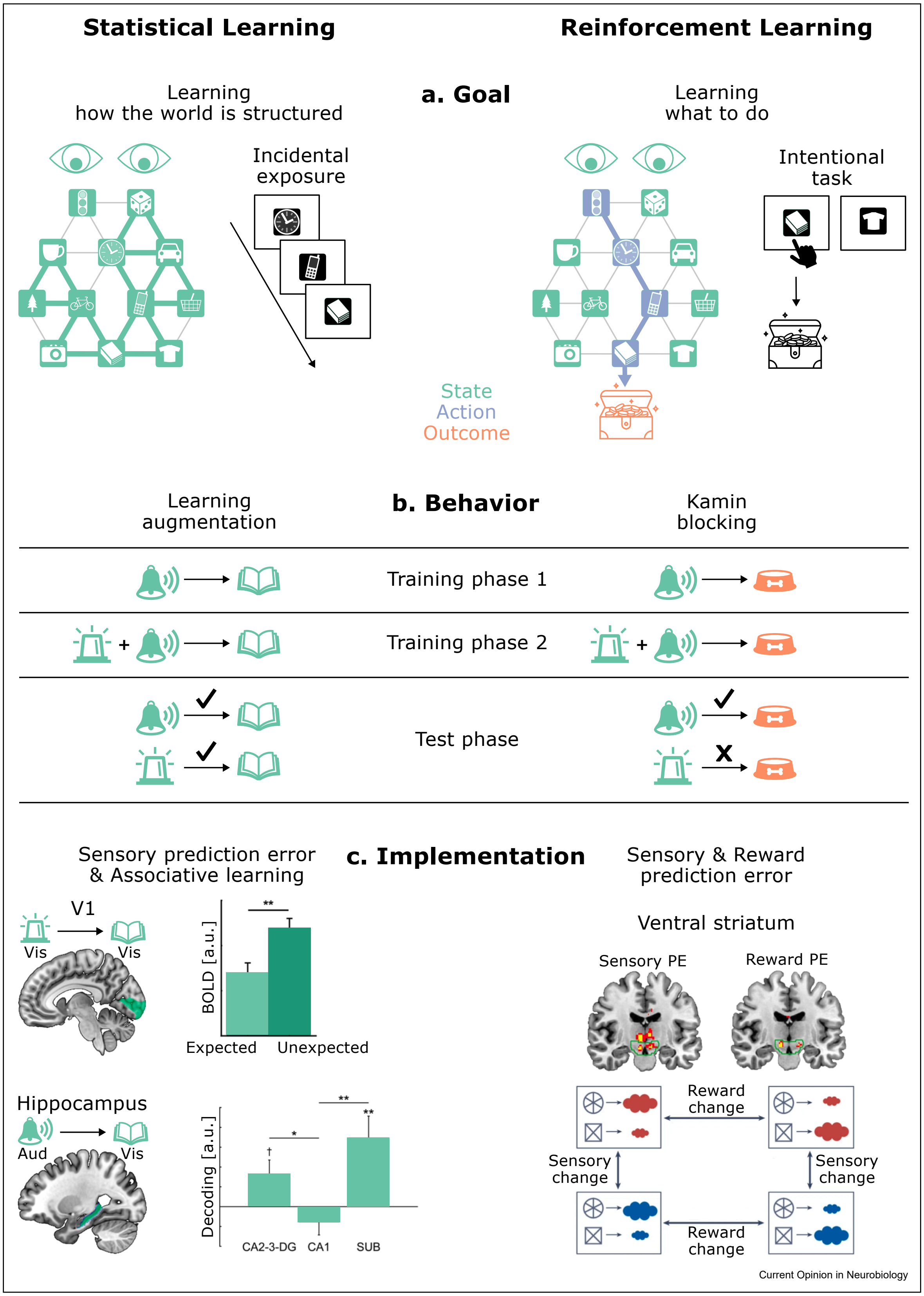
Learning can be broadly conceptualized as using past and current events to predict the future [27]. However, not all information is equally useful to learn. Event saliency modulates learning, as illustrated by the phenomenon of overshadowing, in which two concurrent cues compete for associative strength [28–30]. For example, when a bright and a faint light are paired with a subsequent sensory event (in SL) or an outcome (in RL), the brighter cue typically dominates learning, reducing or preventing learning about the fainter one. Yet, saliency alone is insufficient: we could, in principle, learn myriad salient regularities, but doing so indiscriminately would be inefficient. Learning must therefore be guided by goals that determine which information is worth acquiring and retaining (Figure 1a).

RL is fundamentally goal-directed, broadly characterized as learning *what to do*: how to map situations to actions to maximize reward and minimize punishment [3]. Across the broad family of RL algorithms (see Section "Computational principles"), all share a central feature: they aim to estimate which states and actions optimize future rewards.

In contrast, the goal of SL is to learn *how the world is structured*, by creating internal models based on the relationships within the input data [15,31]. One core challenge in SL is to uncover how higher-level structures, such as words or visually perceived objects, are constructed from lower-level features, such as phonemes or line segments. Also here, several algorithms have been proposed (see Section "Computational principles"). For example, individuals could learn simple transitions between adjacent events using simple associative mechanisms, or apply more complex learning algorithms to grasp more complex hierarchical structures [32].

These differences in goals between RL and SL are also reflected by the conditions under which learning happens. SL can occur when animals are experiencing their sensory environment without specific tasks or instructions. Indeed, learning can largely occur automatically and without the intention to learn, and consequently, the outcome of learning can remain implicit [14,33–35]. This sensitivity to environmental statistics is a pervasive property of the brain that operates over different timescales, both ontogenetically [36,37] and phylogenetically [38]. In contrast, RL is

Figure 1


Statistical Learning
Reinforcement Learning
a. Goal
Learning how the world is structured
Learning what to do
Incidental exposure
Intentional task
State
Action
Outcome
b. Behavior
Learning augmentation
Kamin blocking
Training phase 1
Training phase 2
Test phase
c. Implementation
Sensory prediction error & Associative learning
Sensory & Reward prediction error
V1
Vis
Vis
BOLD [a.u.]
Expected
Unexpected
Hippocampus
Aud
Vis
Decoding [a.u.]
CA2-3-DG
CA1
SUB
Ventral striatum
Sensory PE
Reward PE
Reward change
Sensory change
Sensory change
Reward change
Current Opinion in Neurobiology

inherently interactive: it requires the animal to engage with its environment, by choosing actions and evaluating outcomes [6].

In summary, while both RL and SL involve learning structures from experience, they pursue fundamentally different objectives. RL is driven by the goal of optimizing rewards through decision-making, with a strong emphasis on the relationship between state, action and outcome. Instead, SL seeks to understand the world by building internal models through passive observation and statistical pattern recognition, emphasizing knowledge acquisition over immediate behavioral goals.

What is learnt during statistical and reinforcement learning?

Because SL and RL pursue distinct goals, they may prioritize different types of information from the environment. That is, *what* is ultimately learnt strongly depends on *why* it is learned. A compelling dissociation between SL and RL has been observed in the context of Kamin blocking and learning augmentation (Figure 1b), which highlight how learning goals can shape the assimilation of information. In Kamin blocking [39,40], a classical finding in RL, a conditional stimulus A (e.g., a sound) is first associated with an outcome X (e.g., a food reward). When a second stimulus B (e.g., a light) is subsequently presented together with A (A + B→X), animals typically fail to learn the association between B and X. In other words, learning of B-X is blocked by the pre-existing A-X association. The influential Rescorla–Wagner model explains this by postulating that learning occurs only when outcomes violate predictions, i.e. when an RPE is present. Since A already fully predicts X, no RPE is generated, and therefore the B-X association is not learned. Blocking illustrates how error-driven learning mechanisms suppress the formation of associations that are redundant for predicting rewards. In contrast, learning augmentation refers to situations in which the addition of a novel cue (B) to an already predictive cue (A) enhances, rather than suppresses, subsequent learning about B and its association with other events [30,41]. This effect has been observed in SL paradigms involving the incidental learning of stimulus–stimulus associations, in the absence of any explicit reward. In such cases, the novel cue B remains surprising and informative about the structure of the environment, even if the subsequent event X is fully predicted by cue A. Given the goal of learning how the world is structured, learning the B-X association enriches the observer's knowledge of environmental structure despite the absence of reward prediction errors. Together, the contrast between Kamin blocking and learning augmentation suggests that systems driven by different learning goals may assimilate different kinds of information: while RL systems suppress learning about fully predicted events (consistent with error-driven updating), SL systems may continue to extract structure from novel stimuli to refine their internal model of the world.

Computational principles and neural implementation of statistical and reinforcement learning

Although both RL and SL involve updating internal models based on environmental inputs, they differ in how they use error signals, and in their underlying computational and neural mechanisms, which are specifically shaped by their objectives.

Any RL model iterates through three key steps: First, the agent chooses an action based on expected rewards (action selection). Second, any received feedback from the environment is encoded (value encoding). Third, the difference between the actual and predicted outcomes is computed and used to update the future estimates (value update; but see Policy Gradient algorithms in which a behavioral policy is optimized directly, without the intermediate step of value-learning [42]). The entire process is driven by error signals, which continuously correct discrepancies between expectations and outcomes, enabling adaptive learning and decision-making. Classical RL models, such as Rescorla-Wagner [43] and Temporal Difference learning [44], explicitly rely on the computation of RPE to adjust predictions iteratively. In these instances of Model-Free RL, the agent learns the expected value of actions by directly calculating the association strength between states and rewards. Conversely, in Model-Based RL, like Dyna algorithm [45], error signals extend beyond RPE to include State Prediction Error (SPE). This requires the agent to learn a model of the environment, including

Comparison of statistical and reinforcement learning. a. Statistical learning (SL) and Reinforcement Learning (RL) have different goals: in SL, observers estimate and predict relationships between environmental variables ('how the world is structured'), without any reward signal or explicit instruction; in RL, agents develop an optimal decision-making policy for reward maximization ('what to do'). **b.** In RL, agents selectively learn associations that are useful for predicting rewards: learning of new associations is blocked by pre-existing associations that already minimize reward prediction errors (Kamin blocking). In SL, observers can learn new associations to build an accurate model of the world (learning augmentation). **c.** In SL, regularities within a sensory modality are encoded in the respective sensory areas (e.g. via increased BOLD response to unexpected events, adapted from Ref. [24]); regularities across modalities are encoded in the hippocampus (which represents events predicted by cross-modal cues, adapted from Ref. [25]). These effects may rely on sensory prediction errors or associative plasticity devoid of error computations. In RL, dopaminergic transients in the human ventral striatum may signal not only reward prediction errors (e.g. changes in the amount of reward; see cloud size) but also sensory prediction errors (e.g. changes in the type of reward, with equal value; see cloud color) (adapted from Ref. [26]). Vis: visual; Aud: auditory; V1: primary visual cortex; CA1-3, DG, SUB: hippocampal subfields; PE: prediction error.

state transitions and rewards, through prediction error minimization. In biological systems (Figure 1c), striatal dopaminergic neurons are thought to encode the RPE, by responding to unexpected rewards and punishment, signaling discrepancies that influence learning and decision-making. Over time, RPE signals shift from unexpected reward outcomes to cues that reliably predict rewards [7]. The mechanistic implementation of SPE in the brain is less established, but mounting evidence shows trial-by-trial correlations of the Model-based SPE in the human striatum, lateral and medial prefrontal cortex, orbitofrontal cortex and posterior parietal cortex [4,26,46,47].

SL models, on the other hand, may or may not explicitly represent and use error signals. In error-driven models, signals similar to SPE, but outside reinforcement settings (i.e., without external rewards), enable systems to correct discrepancies and refine predictions. For example, simple recurrent network models learn sequential dependencies by learning to predict the next state, and iteratively updating weights based on prediction errors [48]. Other SL models operate without representing explicit error signals. In attractor networks [49], memory states evolve through dynamics shaped by prior inputs. Over time, the network's representations become influenced by the statistical distribution of previously encountered stimuli, effectively internalizing environmental regularities. In chunking models (e.g., Ref. [50]), statistical relationships within a data stream (e.g., transitional probabilities) can be learned by identifying high-probability co-occurrences through associative memory consolidation and forgetting, devoid of explicit computation of prediction errors. This contrasts with RL, where error signals like RPE or SPE are essential for updating values and driving goal-directed learning. Crucially, these 'errorless' SL models are not unbounded in learning. Capacity constraints and mechanisms such as normalization or forgetting (or decay) limit and stabilize learning without the need for error signals. Thus, the distinction between error-driven and errorless SL is mechanistic (based on whether an explicit error signal is computed and used), not necessarily algorithmic or mathematical. In biological systems (Figure 1c), statistical associations among sensory events (e.g. two adjacent auditory tones) may be encoded through exposure-driven plasticity between co-active neurons [35,51], with sensory cortices adapting to reflect expected stimulus statistics. This typically leads to perceptual facilitation [52,53] and reduced neural activation [24,54–56] to expected than unexpected stimuli [76], a phenomenon known as "expectation suppression" [57]. However, complex dependencies, such as cross-modal, non-adjacent or context-sensitive regularities, are unlikely to arise from local plasticity within sensory circuits [14]. Due to the limited temporal integration windows of early sensory areas [58], and their restricted access to multimodal inputs, the representations of higher-order structure likely depends on more integrative brain areas. The hippocampus shows a larger temporal receptive window than upstream sensory cortices [58], receives inputs from multiple sensory modalities [59], and is sensitive to contextual contingencies [60]. Accordingly, initial evidence suggests a dissociation between early sensory cortices and the hippocampus, with only the latter able to represent cross-modal associative predictions [25]. Furthermore, recent causal evidence shows that disrupting dorsal hippocampal activity in mice impairs their ability to learn auditory statistical sequences in an unsupervised setting, without affecting perceptual discrimination or motivation [77]. This demonstrates that hippocampal circuits are necessary for encoding structured temporal patterns, even in the absence of reward or task. Notably, midbrain striatal activity may act as the teaching signal that functionally connects task-relevant brain areas, for example, those responsible for processing stimuli across different sensory modalities [61,62]. Together, the hippocampus and striatum may be key regions for enabling the statistical learning of complex (non-adjacent, crossmodal, contextual) sensory regularities. In addition, posterior parietal, dorsolateral prefrontal and orbitofrontal cortex may guide hierarchical contextual learning [63–65], allowing the emergence of flexible action policies and schemas for adaptive behavior [47]. In summary, while local sensory plasticity, coupled with hippocampal and striatal signals, may determine the emergence of internal models via SL, the usage of these models to guide flexible inference and decision-making (including in Model-Based RL) may depend on downstream frontoparietal regions.

Convergence and divergence

While RL and SL share computational principles, such as the iterative refinement of internal models, they differ in their reliance on error signals and the types of learning they support. RL is inherently goal-directed, using scalar RPE or SPE signals to optimize decisions and actions. In contrast, SL encompasses a broader spectrum of mechanisms, from error-driven adjustments to associative learning devoid of error computation. Recent advancements in RL, such as Distributional RL [66,67], extend traditional approaches by modeling the full distribution of possible rewards rather than their expected value. By accounting for uncertainty and variability in outcomes, this approach allows agents to capture richer statistical properties of the reward landscape, similar to SL. However, even in these models, learning remains fundamentally reward-centric. Formally, within the Markov Decision Processes (MDP) framework, RL is about control: selecting actions (a) to maximize expected cumulative reward (R). This requires learning and using components such as the reward function $R(s, a)$, state transitions $P(s'|s, a)$, and optimal policies $\pi(a|s)$. RPEs and SPEs are used to improve value estimates or refine a model of the environment that supports planning and action selection. In contrast, SL is concerned with modeling the structure of the environment. Formally, this can be framed as learning components of the environment's generative model; for example, learning $P(s_{t+1}|s_t)$, or more generally $P(O_{1:T})$ over sequences of observations. This aligns with the world model used in Model-Based RL, but SL differs fundamentally in that there is no policy optimization or interaction with a task objective. SL often occurs through passive observation, without requiring action or feedback.

While Model-Free RL relies solely on RPEs, Model-Based RL also incorporates SPEs to build structured representations, bringing it conceptually closer to SL. In principle, the model acquired through SL could serve as the world model for Model-Based RL. Consistent with this idea, classic notions of cognitive maps [68] and recent work show that mice can form internal models of spatial layout through free exploration and later use them to rapidly learn the location of a water port [69], illustrating how structure acquired without explicit reward can subsequently support goal-directed behavior. However, whether the models learned by SL and Model-Based RL actually converge depends on the goals of learning and the nature of the agent–environment interaction. Crucially, Model-Based RL is shaped by the exploration–exploitation tradeoff: the agent must decide to explore the environment to learn more, or exploit current knowledge to gain reward. This tradeoff can systematically bias the learning toward task-relevant representations, omitting latent structure not immediately useful or relevant to the reward. As a result, Model-Based RL may implement only a discriminative model, i.e. a parsimonious representation tuned for mapping states (or state–action pairs) to outcomes or values. For instance, in a categorization task, learners may acquire a full generative model of category structure or instead converge on a more efficient discriminative mapping between stimuli and responses. Both strategies are viable, but only the former retains the full structure of the input space [70]. In contrast, SL aims to build a comprehensive model of the environment, independent of any task demands or reward contingencies. The resulting representations may preserve high-dimensional, richly structured information, yielding qualitatively different internal models from those acquired via RL, despite identical sensory experience. One RL algorithm that converges more closely with SL is the Successor Representation (SR) [71–73], which encodes the expected future state occupancy given the current state, discounted over time. Because it does not depend directly on rewards, SR offers a middle ground between Model-Free and Model-Based RL. SL-derived models may constrain or initialize such representations, providing a flexible and efficient substrate for policy learning. In this way, SL is not merely a subset of RL (i.e. Model-based RL) or redundant with it; it instead represents a distinct learning pathway, optimized for capturing structure rather than control.

Conclusions and future directions of research

While the goals of Statistical Learning (SL) and Reinforcement Learning (RL) differ—SL focusing on extracting environmental structure and RL on optimizing behavior for rewards—the two forms of learning are deeply interconnected. RL may strongly build on the knowledge internalized via SL, as an accurate world model acquired through SL can enhance Model-Based RL by facilitating the learning of stimulus-action-outcome associations. This interdependence blurs the boundaries between SL and RL, suggesting that they often operate in tandem rather than as entirely distinct processes.

Future research should focus on better delineating the boundaries between SL and RL and identify when they operate independently, synergistically or competitively. A key open question concerns the neural implementation of SL, particularly whether and how SPEs contribute to learning across cortical and subcortical networks. Importantly, it is unclear whether dopamine plays a role in SL, similar to its established function in RL, or whether SL relies on distinct neuromodulatory systems. Interestingly, dopamine signals appear to play a reinforcing role also during associative learning [26,74] or spontaneous behavior [75] that are devoid of any explicit reward. Investigating dopaminergic responses in SL paradigms may reveal shared error computation mechanisms across both learning systems. Additionally, examining the overlap between Model-Based RL and SL may clarify how predictive models of the world support both reward-driven and unsupervised learning. Computationally, understanding how SL-derived models support RL (and vice versa) may provide insight into the brain's capacity to flexibly shift between goal-free structure learning and goal-driven decision-making. This can also inform AI, where systems that flexibly integrate SL and RL principles may better adapt to complex, uncertain environments by combining reward optimization with structural pattern recognition. Ultimately, clarifying the relationships between these frameworks will deepen our understanding of how the brain adapts to complex environments.

Credit author statement

Ambra Ferrari: Conceptualization, Visualization, Writing - original draft, Writing - review & editing, Funding acquisition. Floris de Lange: Conceptualization, Writing - original draft, Writing - review & editing, Funding acquisition. Athena Akrami: Conceptualization, Writing - original draft, Writing - review & editing, Funding acquisition.

Funding

This work was supported by the MSCA PF Grant 2022 “MultiMeans” (Project 101105251) awarded to AF, the ERC Consolidator Grant 2020 “Surprise” (Project 101000942) awarded to FPdL, the Wellcome Trust (219627/Z/19/Z) and Gatsby Charitable Foundation (GAT3755) fundings to AA. This work was also partly supported by grant NSF PHY-2309135 to the Kavli Institute for Theoretical Physics (KITP).

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

The authors would like to thank the “Statistical Learning in the Brain” program at KITP (June 12 - July 21, 2023), where the first discussions on this topic began. We are also grateful to the members of the Predictive Brain Lab, LIMLab and Ryan Low for their valuable input and constructive discussions over the years, which have greatly contributed to shaping the ideas presented in this review.

Data availability

No data was used for the research described in the article.

References

Papers of particular interest, published within the period of review, have been highlighted as:

* of special interest
** of outstanding interest

1. Lange FP de, Heilbron M, Kok P: **How do expectations shape perception?** *Trends Cognit Sci* 2018, **22**:764–779.

2. Sherman BE, Graves KN, Turk-Browne NB: **The prevalence and importance of statistical learning in human cognition and behavior**. *Curr Opin Behav Sci* 2020, **32**:15.

3. Gershman SJ, Daw ND: **Reinforcement learning and episodic memory in humans and animals: an integrative framework**. *Annu Rev Psychol* 2017, **68**:101–128.

4. 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.

5. Greenstreet F, Vergara HM, Johansson Y, Pati S, Schwarz L, Lenzi SC, Geerts JP, Wisdom M, Gubanova A, Rollik LB, *et al.*: **Dopaminergic action prediction errors serve as a value-free teaching signal**. *Nature* 2025, **643**:1333–1342.

6. Sutton RS, Barto AG: *Reinforcement learning, second edition: an introduction*. MIT Press; 2018.

7. Schultz W, Dayan P, Montague PR: **A neural substrate of prediction and reward**. *Science* 1997, **275**:1593–1599.

8. Glimcher PW: **Understanding dopamine and reinforcement learning: the dopamine reward prediction error hypothesis**. *Proc Natl Acad Sci U S A* 2011, **108**(Suppl 3):15647–15654.

9. Cohen JY, Haesler S, Vong L, Lowell BB, Uchida N: **Neuron-type-specific signals for reward and punishment in the ventral tegmental area**. *Nature* 2012, **482**:85–88.

10. * Jeong H, Taylor A, Floeder JR, Lohmann M, Mihalas S, Wu B, Zhou M, Burke DA, Namboodiri VMK: **Mesolimbic dopamine release conveys causal associations**. *Science* 2022, **378**. eabq6740.

This study challenges the dominant view of dopamine signal as reflecting a reward prediction error signal, casting it instead as signaling meaningful causal targets (i.e., events whose cause should be learned).

11. Reber AS: **Implicit learning of artificial grammars**. *J Verb Learn Verb Behav* 1967, **6**:855–863.

12. Saffran JR, Aslin RN, Newport EL: **Statistical learning by 8-month-old infants**. *Science* 1996, **274**:1926–1928.

13. Aslin RN: **Statistical learning: a powerful mechanism that operates by mere exposure**. *Wiley Interdiscip Rev Cogn Sci* 2017, **8**.

14. Conway CM: **How does the brain learn environmental structure? Ten core principles for understanding the neurocognitive mechanisms of statistical learning**. *Neurosci Biobehav Rev* 2020, **112**:279–299.

15. Fiser J, Lengyel G: **Statistical learning in vision**. *Annu Rev Vis Sci* 2022, **8**:265–290.

16. Gibson JJ, Gibson EJ: **Perceptual learning: differentiation or enrichment?** *Psychol Rev* 1955, **62**:32–41.

17. Brown R, Fraser C: **The acquisition of syntax**. *Monogr Soc Res Child Dev* 1964, **29**:43–79.

18. Sohoglu E, Chait M: **Detecting and representing predictable structure during auditory scene analysis**. *eLife* 2016, **5**, e19113.

19. Garami L, Fiser J: **Temporal segmentation principles in vision and audition**. *J Vis* 2024, **24**:1117.

20. Pearce MT: **Statistical learning and probabilistic prediction in music cognition: mechanisms of stylistic enculturation**. *Ann N Y Acad Sci* 2018, **1423**:378–395.

21. Nemeth D, Hallgató E, Janacsek K, Sándor T, Londe Z: **Perceptual and motor factors of implicit skill learning**. *Neuroreport Rapid Commun Neurosci Res* 2009, **20**:1654–1658.

22. Thorndike EL: *The elements of psychology*. Routledge; 2013.

23. Bellman R: **The theory of dynamic programming**. *Bull Am Math Soc* 1954, **60**:503–515.

24. Richter D, Ekman M, de Lange FP: **Suppressed sensory response to predictable object stimuli throughout the ventral visual stream**. *J Neurosci* 2018, **38**:7452–7461.

25. Kok P, Turk-Browne NB: **Associative prediction of visual shape in the hippocampus**. *J Neurosci* 2018, **38**:6888–6899.

26. ** Kahnt T, Schoenbaum G: **The curious case of dopaminergic prediction errors and learning associative information beyond value**. *Nat Rev Neurosci* 2025, https://doi.org/10.1038/s41583-024-00898-8.

This paper offers a comprehensive review of studies in rodents and humans suggesting that midbrain dopamine transients implement not only reward prediction errors, but also sensory prediction errors.

27. Niv Y, Schoenbaum G: **Dialogues on prediction errors**. *Trends Cognit Sci* 2008, **12**:265–272.

28. Mackintosh NJ: **Overshadowing and stimulus intensity**. *Anim Learn Behav* 1976, **4**:186–192.

29. Boddez Y, Haesen K, Baeyens F, Beckers T: **Selectivity in associative learning: a cognitive stage framework for blocking and cue competition phenomena**. *Front Psychol* 2014, **5**:1–13.

30. * Nazlõ İ, Ferrari A, Huber-Huber C, de Lange FP: **Forward and backward blocking in statistical learning**. *PLoS One* 2024, **19**, e0306797.

This study shows that during incidental statistical learning of associations between sensory stimuli, learning is not ‘forward blocked’ by associations formed earlier. This suggests a potential contrast between SL and RL, possibly linked to their divergent goals.

31. Saffran JR, Kirkham NZ: **Infant statistical learning**. *Annu Rev Psychol* 2018, **69**:181–203.

32. Orbán G, Fiser J, Aslin RN, Lengyel M: **Bayesian learning of visual chunks by human observers**. *Proc Natl Acad Sci* 2008, **105**:2745–2750.

33. Sherman BE, Graves KN, Turk-Browne NB: **The prevalence and importance of statistical learning in human cognition and behavior**. *Curr Opin Behav Sci* 2020, **32**:15–20.

34. Frost R, Armstrong BC, Christiansen MH: **Statistical learning research: a critical review and possible new directions**. *Psychol Bull* 2019, **145**:1128–1153.

35. Batterink LJ, Paller KA, Reber PJ: **Understanding the neural bases of implicit and statistical learning**. *Top Cogn Sci* 2019, **11**:482–503.

36. Gopnik A, Tenenbaum JB: **Bayesian networks, Bayesian learning and cognitive development**. *Dev Sci* 2007, **10**:281–287.

37. Berkes P, Orbán G, Lengyel M, Fiser J: **Spontaneous cortical activity reveals hallmarks of an optimal internal model of the environment**. *Science* 2011, **331**:83–87.

38. Geisler WS, Diehl RL: **Bayesian natural selection and the evolution of perceptual systems**. *Philos Trans R Soc Lond B Biol Sci* 2002, **357**:419–448.

39. Kamin LJ: *Attention-like processes in classical conditioning*. 1967.

40. Aggarwal M, Wickens JR: **Behavioral determinants in the expression of the kamin blocking effect: implications for associative learning theory**. *Neurosci Biobehav Rev* 2021, **124**: 16–34.

41. Beesley T, Shanks DR: **Investigating cue competition in contextual cuing of visual search**. *J Exp Psychol Learn Mem Cogn* 2012, **38**:709–725.

42. Bennett D, Niv Y, Langdon AJ: **Value-free reinforcement learning: policy optimization as a minimal model of operant behavior**. *Curr Opin Behav Sci* 2021, **41**:114–121.

43. Rescorla R: **A theory of Pavlovian conditioning : variations in the effectiveness of reinforcement and nonreinforcement**. *Classsical Cond II Curr Res Theory* 1972, **2**:64–69.

44. Sutton RS: **Learning to predict by the methods of temporal differences**. *Mach Learn* 1988, **3**:9–44.

45. Sutton RS: **Integrated architectures for learning, planning, and reacting based on approximating dynamic programming**. In *Machine learning proceedings 1990*. Edited by Porter B, Mooney R, Morgan Kaufmann; 1990:216–224.

46. Klein-Flügge MC, Wittmann MK, Shpektor A, Jensen DEA, Rushworth MFS: **Multiple associative structures created by reinforcement and incidental statistical learning mechanisms**. *Nat Commun* 2019, **10**:1–15.

47. ** Bein O, Niv Y: **Schemas, reinforcement learning and the medial prefrontal cortex**. *Nat Rev Neurosci* 2025, https://doi.org/10.1038/s41583-024-00893-z.
This paper argues that schemas (knowledge structures about the typical unfolding of events in a context, e.g., a schema of a restaurant meal) are learned via RL. This is achieved by leveraging not only prediction errors, but also complementary algorithms such as dimensionality reduction and latent cause inference.

48. Elman JL: **Finding structure in time**. *Cogn Sci* 1990, **14**:179–211.

49. Boboeva V, Pezzotta A, Clopath C, Akrami A: **Unifying network model links recency and central tendency biases in working memory**. *eLife* 2024, **12**. RP86725.

50. Perruchet: PARSER: A model for word segmentation - google scholar. [date unknown].

51. Reber PJ: **The neural basis of implicit learning and memory: a review of neuropsychological and neuroimaging research**. *Neuropsychologia* 2013, **51**:2026–2042.

52. Kim R, Seitz A, Feenstra H, Shams L: **Testing assumptions of statistical learning: is it long-term and implicit?** *Neurosci Lett* 2009, **461**:145–149.

53. Bertels J, Franco A, Destrebecqz A: **How implicit is visual statistical learning?** *J Exp Psychol Learn Mem Cogn* 2012, **38**: 1425–1431.

54. Alink A, Schwiedrzik CM, Kohler A, Singer W, Muckli L: **Stimulus predictability reduces responses in primary visual cortex**. *J Neurosci* 2010, **30**:2960–2966.

55. Kaposvari P, Kumar S, Vogels R: **Statistical learning signals in macaque inferior temporal cortex**. *Cerebr Cortex* 2018, **28**: 250–266.

56. Richter D, de Lange FP: **Statistical learning attenuates visual activity only for attended stimuli**. *eLife* 2019, **8**, e47869.

57. Meyer T, Olson CR: **Statistical learning of visual transitions in monkey inferotemporal cortex**. *Proc Natl Acad Sci U S A* 2011, **108**:19401–19406.

58. Hasson U, Chen J, Honey CJ: **Hierarchical process memory: memory as an integral component of information processing**. *Trends Cognit Sci* 2015, **19**:304–313.

59. Eichenbaum H, Yonelinas AP, Ranganath C: **The medial temporal lobe and recognition memory**. *Annu Rev Neurosci* 2007, **30**:123–152.

60. Biderman N, Bakkour A, Shohamy D: **What are memories for? The hippocampus bridges past experience with future decisions**. *Trends Cognit Sci* 2020, **24**:542–556.

61. den Ouden HEM, Friston KJ, Daw ND, McIntosh AR, Stephan KE: **A dual role for prediction error in associative learning**. *Cerebr Cortex* 2009, **19**:1175–1185.

62. den Ouden HEM, Daunizeau J, Roiser J, Friston KJ, Stephan KE: **Striatal prediction error modulates cortical coupling**. *J Neurosci* 2010, **30**:3210–3219.

63. Niv Y, Daniel R, Geana A, Gershman SJ, Leong YC, Radulescu A, Wilson RC: **Reinforcement learning in multidimensional environments relies on attention mechanisms**. *J Neurosci* 2015, **35**:8145–8157.

64. Leong YC, Radulescu A, Daniel R, DeWoskin V, Niv Y: **Dynamic interaction between reinforcement learning and attention in multidimensional environments**. *Neuron* 2017, **93**:451–463.

65. Ferrari A, Richter D, de Lange FP: **Updating contextual sensory expectations for adaptive behavior**. *J Neurosci* 2022, **42**: 8855–8869.

66. Dabney W, Kurth-Nelson Z, Uchida N, Starkweather CK, Hassabis D, Munos R, Botvinick M: **A distributional code for value in dopamine-based reinforcement learning**. *Nature* 2020, **577**:671–675.

67. Lowet AS, Zheng Q, Matias S, Drugowitsch J, Uchida N: **Distributional reinforcement learning in the brain**. *Trends Neurosci* 2020, **43**:980–997.

68. Tolman EC: **Cognitive maps in rats and men**. *Psychol Rev* 1948, **55**:189–208.

69. Rosenberg M, Zhang T, Perona P, Meister M: **Mice in a labyrinth show rapid learning, sudden insight, and efficient exploration**. *eLife* 2021, **10**, e66175.

70. Menichini E, Pajot-Moric Q, Low R, Pedrosa V, Pourdehghan A, Vincent P, Zhou L, Teachen L, Akrami A: *Different learning algorithms achieve shared optimal outcomes in humans, rats, and mice*. 2025, https://doi.org/10.1101/2023.01.30.526119.

71. Momennejad I: **Learning structures: predictive representations, replay, and generalization**. *Curr Opin Behav Sci* 2020, **32**:155–166.

72. Momennejad I, Russek EM, Cheong JH, Botvinick MM, Daw ND, Gershman SJ: **The successor representation in human reinforcement learning**. *Nat Hum Behav* 2017, **1**:680–692.

73. Gershman SJ: **The successor representation: its computational logic and neural substrates**. *J Neurosci* 2018, **38**:151. 18.

74. Sharpe MJ, Batchelor HM, Mueller LE, Yun Chang C, Maes EJP, Niv Y, Schoenbaum G: **Dopamine transients do not act as model-free prediction errors during associative learning**. *Nat Commun* 2020, **11**:106.

75. ** Markowitz JE, Gillis WF, Jay M, Wood J, Harris RW, Cieszkowski R, Scott R, Brann D, Koveal D, Kula T, *et al.*: **Spontaneous behaviour**

is structured by reinforcement without explicit reward. *Nature* 2023, **614**:108–117.
This study shows that the neurotransmitter dopamine, which has a well-established role in reward-driven behavior, also shapes spontaneous behavior, in the absence of task structure, sensory cues or exogenous reward. This may suggest that dopamine also plays a role in spontaneous learning behavior.

76. Furutachi S, Franklin AD, Aldea AM, *et al*.: **Cooperative thalamocortical circuit mechanism for sensory prediction errors**. *Nature* 2024, **633**:398–406, https://doi.org/10.1038/s41586-024-07851-w.
**
This study identifies the circuit mechanisms underlying the generation of sensory prediction errors. Surprising visual inputs are associated with increased firing in layer 2/3 V1 neurons, caused by a cooperative mechanism requiring thalamic input from the pulvinar and cortical vasoactive-intestinal-peptide-expressing (VIP) inhibitory interneurons.

77. Onih A, Pentousi L, Shen X, Boboeva V, Akrami A: **The hippocampus enables abstract structure learning without reward**. *bioRxiv* 2026, https://doi.org/10.64898/2026.02.14.705916.