Version of Record: https://www.sciencedirect.com/science/article/pii/S2352154621001236
Manuscript_e86f1dc984f3badb47debf00eda4f1f4

What do Reinforcement Learning Models Measure? Interpreting Model Parameters in Cognition and Neuroscience

Maria K. Eckstein[c], Linda Wilbrecht[c,d], Anne G.E. Collins[c,d]

[c]*Department of Psychology, UC Berkeley, 2121 Berkeley Way West, Berkeley, 94720, CA, USA*
[d]*Helen Wills Neuroscience Institute, UC Berkeley, 175 Li Ka Shing Center, Berkeley, 94720, CA, USA*

1. Introduction

Reinforcement learning (RL) is an exploding field. In the domain of machine learning, it has led to tremendous progress in the last decade, ranging from the creation of artificial agents that can beat humans at complex games, such as Go [1] and StarCraft [2], to successful deployment in industrial settings, such as the autonomous navigation of internet balloons in the stratosphere [3]. In cognitive neuroscience, RL models have been used successfully to capture a broad range of latent learning-related phenomena, at the level of both behavior [4, 5] and neural signals [6]. However, the impression that RL can help us identify reasonable and predictive latent variables hides heterogeneity in what RL variables reflect, even within cognitive neuroscience. The success of RL has fed a notion of omniscience that RL can peer into the brain and behavior and surgically isolate and measure essential functions. As this notion grows with the popular uptake of RL methods, it sometimes leads to overgeneralization and overinterpretation of findings.

Here, we argue that a more nuanced view is better supported empirically and theoretically. We first discuss how RL is used in distinct subfields, highlighting shared and distinct components. Then, we examine where cognitive neuroscience may be overstepping in its interpretation, and conclude that, when properly contexualized, RL models retain great value for the field.

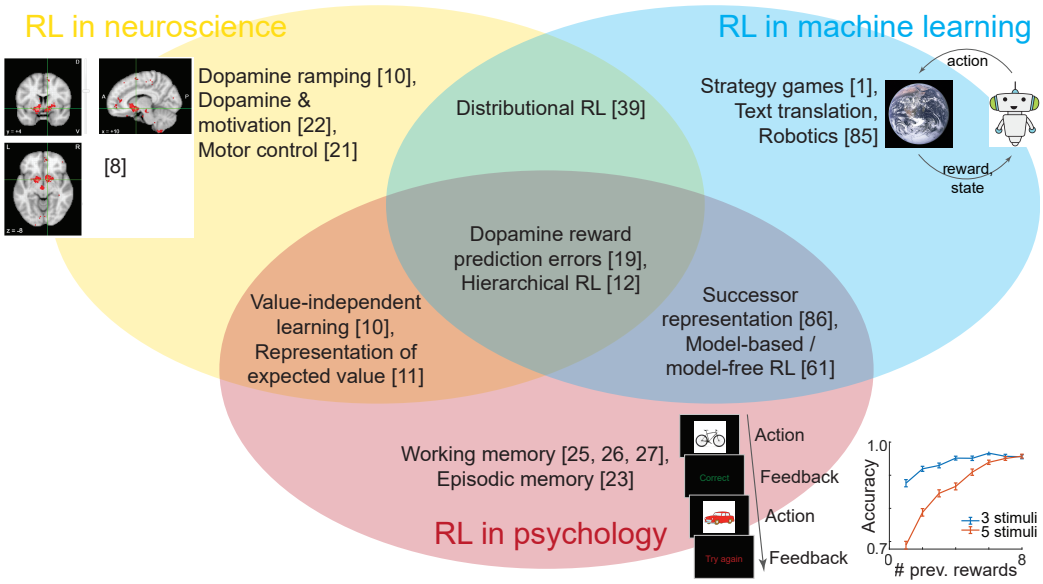

Figure 1: The meaning of "RL" differs between neuroscience, machine learning, and psychology, reflecting a specific brain network, a family of problems and algorithms, and a type of learning, respectively. The concepts are related: RL models successfully capture aspects of RL behavior and brain signals, and some RL behaviors rely on the RL brain network. The dopamine reward prediction error hypothesis combines ideas from all three fields. However, there are also significant discrepancies in what RL means across fields, such that activity in the brain's RL network might not relate to RL behavior and might not be captured by RL models (e.g., dopamine ramping in neuroscience). Importantly, RL behavior may rely on non-RL brain systems and may or may not be captured by RL algorithms. Recent trends have aimed to increase communication between fields and emphasize areas of mutual benefits [7, 8]. RL in neuroscience inset shows the neurosynth automated meta-analysis for "reinforcement learning" (x=10, y=4, z=-8), highlighting striatal function [9]. RL in cognition inset shows that participants become more likely to select a rewarded choice the more previous rewards they have experienced (data replotted from [5]). RL in machine learning shows the agent-environment loop at the basis of RL theory [10].

2. RL in Machine Learning, Psychology, and Neuroscience

In machine learning, RL is defined as a class of learning problems and a family of algorithms that solve these problems. An RL agent can be in any of a set of states, take actions to change states, and receive rewards/punishments (Fig. 1, top-right). RL agents are designed to optimize a specific objective: the expected sum of discounted future rewards. A wide family of RL algorithms offers solutions that achieve this objective [10], for example model-free RL, which estimates the values of actions based on reward prediction errors (Fig. 2A, top).

In psychology, RL defines a psychological process and a method for its study. RL occurs when an organism learns to make choices (or predict outcomes) directly based on experienced rewards/punishments (rather than indirectly through instructions, for example). This includes simple situations, such as those historically studied by behaviorists (classical [6, 11] and instrumental conditioning [12]), as well as more complex ones, such as learning over longer time horizons [13, 14], meta-learning [15], and learning across multiple contexts [16, 17].

Neuroscientists investigating RL usually focus on a well-defined network of regions that implements value learning. These include cortico-basal-ganglia loops, and in particular the striatum (Fig. 1A), thought to encode RL values, and dopamine neurons, thought to signal temporal-difference reward-prediction errors (*RPEs*; Fig. 2A) [6, 9, 18, 19, 20, 21].

The meaning of "RL" overlaps in these three communities (Fig. 1), and RL algorithms from AI have been successful at capturing biological RL behavior and neural function. However, there are also important discrepancies. For example, many functions of the brain's RL network do not relate to RL behavior, such as dopamine's role in motor control [22] or cognitive effort [23]. On the other hand, some RL brain functions that do relate to RL behavior are poorly explained by classic RL models, such as dopamine's role in value-independent learning [11]. Furthermore, many aspects of learning from reward do not depend on the brain's RL network, whether they are captured by RL algorithms or not. For example, hippocampal episodic memory [24, 25] and prefrontal working memory [26, 27, 28] contribute to RL behavior, but are often not explicitly modeled in RL, obscuring the contribution of non-RL neural processes to learning.

Because of these differences in meaning, the term "RL" can cause ambiguity and lead to misinterpretations. Fig. 2 provides an example in which an

RL model leads to conflicting conclusions as to how RL parameters change with age when applied to two slight variants of the same task. This conflict is reconciled, however, by recognizing that working memory contributes most learning in one variant, whereas RL does in the other [5].

Because RL's meaning is ambiguous, it is often unclear how RL model variables (e.g., parameters such as learning rates or decision noise; reward prediction errors; RL values) should be interpreted in models of human and animal learning. In the following, we show that the field often optimistically assumes that model variables are readily interpretable and naturally generalize between studies. We then show that these beliefs are oftentimes not well supported, and offer an alternative interpretation.

3. Interpretability and Generalizability of RL Model Variables

3.1. What do "Cognitive" Models Measure?

RL models attempt to approximate behavior by fitting free parameters [30, 31, 32, 33], and are used by most researchers to elucidate cognitive and/or neural function (Box 1): RL "has emerged as a key framework for modeling and understanding *decision-making*"[1] [34]. The reason why models of behavior are used as "cognitive models" is that they implement hypotheses about cognition. Therefore, the good fit of a model to behavior implies that participants could have employed the modeled algorithm cognitively. Nevertheless, stronger conclusions are often drawn: For example, the good fit of inference algorithms to human behavior and brain function has been taken as evidence that human brains implement inference [17]. However, there always is an infinite number of alternative algorithms that would fit behavior equally well, such that inferring participants' cognitive algorithms through model fitting is impossible [33, 35, 36].

3.2. Interpretability and Generalizability

This notion that computational models—astonishingly—isolate and measure intrinsic (neuro)cognitive processes from observable behavior has contributed to their attractiveness as a research method. However, we believe we need to temper our optimism in two areas: *interpretability* and *generalizability* (Fig. 3).

[1] emphasis added

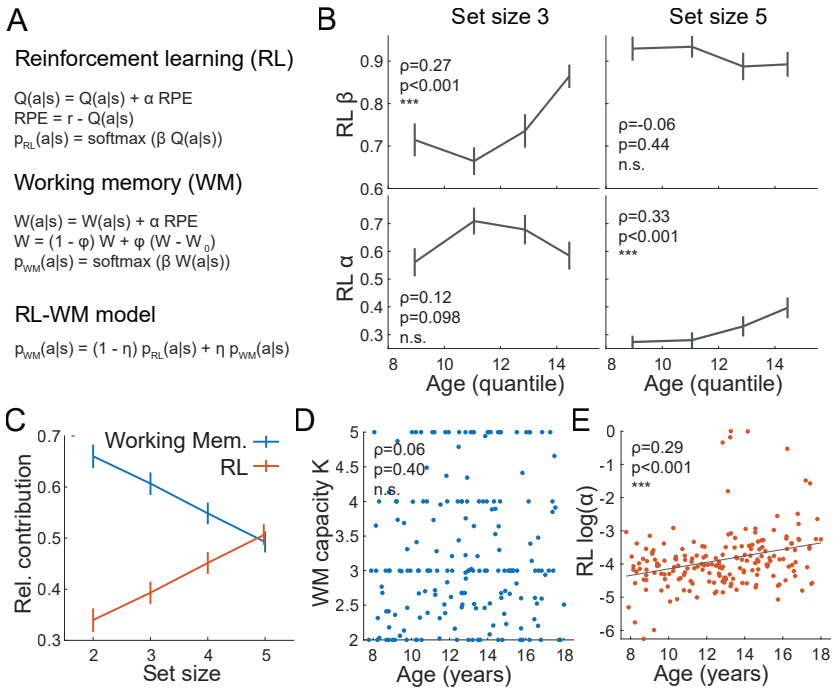

Figure 2: Fitting standard RL models can lead to the wrong impression that cognitive processing is purely based on RL. (A) Update equations for the RL-WM model. $Q(a|s)$ indicates the RL state-action value of action a in state s, which is updated based on the reward prediction error RPE. $W(a|s)$ is the working-memory weight of a in s, and ϕ is a forgetting parameter, β is the decision noise, and η the mixing parameter combining RL and working memory processes. For model details, see [5, 29]. (B) When separate standard RL models are fit to different contexts within the same task (here, the number of stimuli [6]), they provide different answers as to how age affects RL model parameters (decision noise β, top; learning rate α, bottom). Contexts with fewer stimuli ("Set size 3", left) suggest that age does not affect learning rates, whereas contexts with more stimuli ("Set size 5", right) suggest that learning rates increase with age. Inset statistics show non-parametric Spearman correlation coefficients ρ and p-values (N=187, * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$). (C)-(E) When using a model that fits all contexts jointly by combining RL processes with working memory ("RL-WM" model), these discrepancies are resolved [5]. (C) The RL-WM model reveals that the relative contributions of RL compared to working memory differ between contexts. A standard RL model would falsely attribute working-memory processes in contexts with small sizes to the RL system, in this example suggesting that learning rates do not change with age (A, set size 3). (D) Working memory capacity in the RL-WM model was not related to participants' ages, explaining why learning rates did not increase with age in (A, set size 3), in which working memory contributed most to learning. (E) RL learning rates in the RL-WM model increased with age. Since RL contributed more to learning in set size 5 (C), this was detected in the standard RL model of only set size 5 (B). Data reanalyzed from [5].

Interpretability means that model variables (e.g., parameters, reward prediction errors) isolate specific, fundamental, and invariant elements of (neuro)cognitive processing: Decomposing behavior into model variables is seen as a way of carving cognition at its joints, producing model variables that are of essential nature. Generalizability means that model variables capture inherent individual characteristics (e.g., a person with a high learning rate), such that we can robustly infer the same parameter for the same person across different contexts, tasks, and model variants.

Though rarely stated explicitly, assumptions about interpretability and generalizability lie at the heart of much current computational cognitive research (including our own), as we show in the literature survey below (Box 1), and play a consequential role in interpreting and guiding future research. However, we also show that empirical support for interpretability and generalizability is ambivalent at best, and often negative. We highlight a recent multi-task within-participants study from our group that explores precisely when model parameters do and do not generalize between tasks, and how dissimilar the cognitive processes are they capture (interpretability).

Box 1: Representative statements from the literature that imply interpretability and generalizability.[a]

- **Interpretability**: Computational models have been described as "illuminating [...] *cognitive processes* or *neural representations* that are otherwise difficult to tease apart" [37]; clarifying "the *neural processes* underlying *decision-making*" [18]; and revealing "what computations are performed in *neuronal populations* that support a particular *cognitive process*"[b] [38]. This highlights the common assumption that computational models can reveal cognitive and neural processes and identify specific, "theoretically meaningful" [39] elements of (neuro)cognitive function.

  Models are thereby often expected to provide the "linking propositions" [40] between cognition and neural function, "*mapping* latent decision-making processes onto dissociable neural substrates" [41] and "*link[ing]* cognitive mechanisms to [clinical] symptoms" [38].

  These links are often assumed to be specific one-to-one mappings: "Dopamine neurons code an error in the prediction of reward" [20]; "corticostriatal loops enable state-dependent value-based choice"

[27]; "striatal areas [...] support reinforcement learning, and frontoparietal attention areas [...] support executive control processes" [42]; "individual differences in DA clearance and frontostriatal coordination may serve as markers for RL" [43]; and "BOLD activity in the VS, dACC, and vmPFC is correlated with learning rate, expected value, and prediction error, respectively" [44]. This shows that computational variables are often interpreted as specific (neuro)cognitive functions, revealing an assumption of *interpretability*.

- **Generalizability**: Empirical parameter distributions obtained in one task were described as "fairly *transferable*" [45] and used as priors when fitting parameters to a new task [46], revealing the belief that model parameters generalize between studies, tasks, and models.

  Developmental research has aimed to illuminate "the tuning of the learning rate parameter across development" and the "developmental change in the inverse temperature parameter" [37], suggesting that parameters are person-specific but task-independent.

  Many have aimed to find regularities in parameter findings between studies: "[D]ifferential learning rates *tend to be* biased in the direction of learning from positive RPEs" [47]; "this finding *[supports] previous results* on decreased involvement of the reinforcement learning system when cortical resources [...] support task execution" [42]; from our own work: "there was [...] a bias towards learning from positive feedback, which is *consistent with other work*" [5].

[a] We acknowledge that these statements may not represent the full complexity of researchers' knowledge, as many are aware of modeling limitations.

[b] All emphases added.

3.2.1. Interpretability

Many research practices are deeply invested in the interpretability of RL (Box 1). The computational neurosciences, for example, aim to link computational variables to specific neural functions, searching for one-to-one mappings that would allow the inference of one from the other [6, 12, 43, 48]. Prominent examples of interpretable mappings are the links between the

midbrain-dopamine system and RL reward prediction errors [20, 49, 50], and between striatal function and value learning [19, 51, 52, 53]. Computational psychiatry aims to map model variables onto psychiatric diagnoses or symptoms, in an effort to obtain diagnostic tools and causal explanations of aberrant processing [38, 39, 41, 54]. Developmental research aims to map age-related changes in model variables onto developing neural function and real-world behavior [37, 55, 56]. In sum, the conviction in model interpretability is evident in the practice of interpreting model variables as specific cognitive processes, unique neural substrates, and well-delineated psychiatric symptoms.

3.2.2. Generalizability

Assumptions about parameter generalizability are also widespread. In computational neuroscience, model variables are routinely expected to measure the same latent neural substrates, even when the underlying task, model, or participant samples differ [18, 19, 20, 57, 58, 59, 60]. For example, fields studying individual differences, such as clinical [38, 39] and developmental psychology [37, 55, 56], aim to identify how model variables covary with other variables of interest (e.g., age, traits, symptoms) in a systematic way across studies, and review articles and discussion sections confidently compare modeling variables between studies.

3.3. Evidence Against Interpretability and Generalizability

However, meta-reviews suggest that interpretability and generalizability might be overassumptions, common in classic psychological research [61] and RL modeling [41]. RL appears interpretable because multiple studies have replicated mappings between RL variables and specific neural function. However, these mappings are not as consistent as expected: The famous mapping between dopamine / striatal activity and reward prediction errors, for example, supported by classic and recent research [6, 20], varies considerably between studies based on details of the experimental protocol, as shown in several recent meta-analyses [57, 59, 62].

Discrepancies are also evident in the mapping between RL variables and cognitive function. For example, learning rates are often interpreted as incremental updating (dopamine-driven neural plasticity) in classical conditioning [20], but also as reward sensitivity [63], sampling from (hippocampal) episodic memory [25], the ability to optimally weigh decision outcomes [64], or approximate inference [4], in other tasks. There is substantial variance between

studies in terms of which neural and which cognitive processes underlie the same RL variables, contradicting the notion of interpretability.

Evidence for generalizability is also weak: Similar adult samples have differed strikingly in terms of their average estimated RL learning rates (0.05-0.7) [44, 63, 65, 66] and "positivity bias" [47, 67, 68], depending on the underlying task and model parameterization. In developmental samples, the trajectories of RL learning rates have shown increases [5, 63, 69], decreases [70], U-shaped trajectories [4], or no change [71] in the same age range. Similar discrepancies have also arisen in the computational psychiatry literature [38, 39, 72, 73]. These inconsistencies would not be expected if model variables were an inherent property of participants that could be assessed independently of study specifics, i.e., if models were generalizable.

Many in our community have noticed such discrepancies and invoked methodological differences between studies to explain them [12, 37, 44, 62, 74, 75]. However, this insight has rarely been put into practice, and model variables keep being compared between studies (Box 1). To remedy this, we assessed interpretability and generalizability empirically, comparing RL parameters from three tasks performed by the same subjects in a developmental sample (291 subjects aged 8-30; Fig. 3B) [4, 5, 69, 76]. We found generalizability but poor interpretability for decision noise, and a fundamental lack of both interpretability and generalizability for learning rates (Fig. 3C).

A likely reason why generalizability and interpretability are lacking in many cases is that computational models are fundamentally models of behavior, and not cognition. Because participants—reasonably—behave differently in different tasks (e.g., repeating non-rewarded actions in stochastic, but not deterministic tasks [76]), estimated parameters (e.g., learning rates) differ as well. Such differences do not necessarily reflect a failure of computational models to measure intrinsic processes, but likely the fact that the same parameters capture different behaviors and different cognitive processes when applied to different tasks (Fig. 3B, 3C) [76].

Another reason for lacking generalizability and interpretability is that the design of computational models, a researcher degree of freedom [35, 36], can impact parameters severely, as recent research has highlighted [47, 67, 68]. Because the same models can be parameterized differently [77], and models with different equations can approximate similar processes [4], model differences are a ubiquitous feature of computational modeling.

To explain parameter discrepancies, others have argued that participants adapt their parameter values to tasks based on optimality [37], or that task

characteristics (e.g., uncertainty) influence neural processing (e.g., dopamine function), which is reflected in differences in model variables (e.g., reward prediction errors) [78, 79]. Whether choices are aligned with participants' goals also fundamentally impacts neural RL processes [80], and so do other common task characteristics [59]. This shows that small task differences impact behavior, neural processing, and computational variables. Even though RL models might successfully capture behavior in each task, parameters likely capture different aspects each time, leading to a lack of interpretability and generalizability.

4. Conclusion and Outlook

A tremendous literature has shown RL's potential and successes—this opinion piece emphasizes some caveats, showing that RL is not a single concept and that RL models are a broad family that reflects a range of cognitive and neural processes.

A lack of interpretability and generalizability has major implications for the comparison of model variables between tasks, a practice that forms the basis for many review articles, meta-analyses, introduction and discussion sections of empirical papers, and for directing future research. Evidence suggests that in many cases, parameters cannot directly be compared between studies, and capture different (neuro)cognitive processes depending on task characteristics. Future research needs to determine which model variables do and do not generalize, over which domain, and what the determining factors are. In the meantime, researchers should be more nuanced when comparing results between studies, and acknowledge contextual factors that might limit generalizability. Lastly, what model variables measure might differ for each task, and researchers should provide additional validation on a task-by-task basis, relating variables to behavioral measures or individuals' traits, and using simulations to determine the role of model variables in specific tasks.

Another solution is to explicitly model variability between features that should be generalized over, including task characteristics (Fig. 2), models, participants, and potentially even neural processes [61]. Several studies have made strides in this direction, incorporating features that are intrinsic to participants (working memory [5, 29], attention [28], development [37, 81, 56]), or extrinsic (task time horizon [13, 14], context changes [16]), thus broadening the domain over which models generalize. However, infinitely many features likely affect RL processes, rendering entirely general models

infeasible. Researchers therefore need to select a domain of interest for each model, and acknowledge this choice. As authors, reviewers, and editors, we should balance our excitement about general statements with our knowledge about the inherent limitations of all models, including RL. Future research needs to determine whether similar issues arise for other model families, such as sequential sampling [82, 83], Bayesian inference [4, 28, 84], and others.

We hope that this explicit discussion of assumptions and overassumptions will help our field solve the mysteries of the brain as modeling—with its limitations—is embraced by a growing audience.

5. Acknowledgements

This work was in part supported by National Science Foundation grant 1640885 SL-CN: Science of Learning in Adolescence to AGEC and LW, NIH grant 1U19NS113201 to LW, and NIMH RO1MH119383 to AGEC.

References

[1] D. Silver, J. Schrittwieser, K. Simonyan, I. Antonoglou, A. Huang, A. Guez, T. Hubert, L. Baker, M. Lai, A. Bolton, Y. Chen, T. Lillicrap, F. Hui, L. Sifre, G. van den Driessche, T. Graepel, D. Hassabis, Mastering the game of Go without human knowledge, Nature 550 (7676) (2017) 354–359, number: 7676 Publisher: Nature Publishing Group. doi:10.1038/nature24270.
URL https://www.nature.com/articles/nature24270

[2] O. Vinyals, I. Babuschkin, W. M. Czarnecki, M. Mathieu, A. Dudzik, J. Chung, D. H. Choi, R. Powell, T. Ewalds, P. Georgiev, J. Oh, D. Horgan, M. Kroiss, I. Danihelka, A. Huang, L. Sifre, T. Cai, J. P. Agapiou, M. Jaderberg, A. S. Vezhnevets, R. Leblond, T. Pohlen, V. Dalibard, D. Budden, Y. Sulsky, J. Molloy, T. L. Paine, C. Gulcehre, Z. Wang, T. Pfaff, Y. Wu, R. Ring, D. Yogatama, D. Wünsch, K. McKinney, O. Smith, T. Schaul, T. Lillicrap, K. Kavukcuoglu, D. Hassabis, C. Apps, D. Silver, Grandmaster level in StarCraft II using multi-agent reinforcement learning, Nature 575 (7782) (2019) 350–354, number: 7782 Publisher: Nature Publishing Group. doi:10.1038/s41586-019-1724-z.
URL https://www.nature.com/articles/s41586-019-1724-z

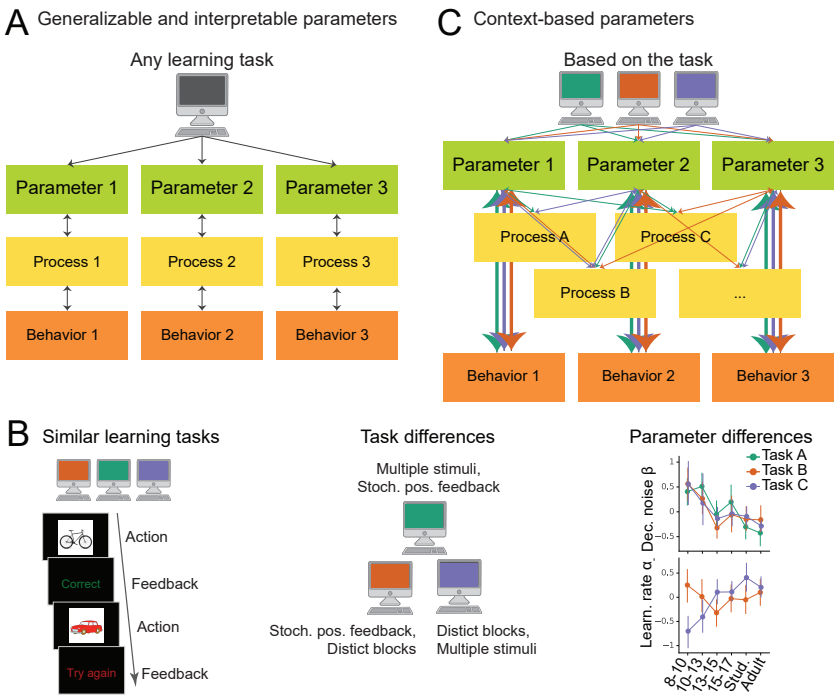

Figure 3: What model variables (e.g., parameters) measure in psychology and neuroscience. (A) View based on interpretability and generalizability. In this view—implicitly taken by much current research—models are fitted in order to reveal individuals' intrinsic characteristics, whereby model parameters reflect clearly delineated, separable, and unique (neuro)cognitive processes. This concept of *interpretability* is shown in the figure in that every model parameter captures one specific cognitive process (bidirectional arrows between parameter and process), and that cognitive processes are separable from each other (no connections between processes). Specific task characteristics are neglected as irrelevant, a concept we call *generalizability*, which is evident in that parameters of "any learning task" (within reason) are expected to capture the same cognitive processes. (B) In our empirical study [76], participants worked on three learning tasks with similar structure (left), but slight differences (middle), reflecting the literature. We created three RL models that captured the behavior in each task [4, 5, 69]. Compared between tasks but within participants, learning rate parameters showed poor interpretability and generalizability [76]: Both absolute values and age trajectories (right, bottom) differed vastly, and individual differences in one task could not be predicted by those in other tasks, as would be expected if they were interpretable as the same (neuro)cognitive substrate. Other parameters, most notably decision noise (right, top), were more generalizable and interpretable, in accordance with emerging patterns in the literature [37], even though they also lacked a shared core of variance across tasks (more for more dissimilar tasks). In contrast, the mappings between parameters and behavioral features were consistent across tasks, suggesting that parameters generalized in terms of behavioral processes, but not cognitive ones. (C) Updated view that acknowledges the role of context in computational modeling (e.g., task characteristics, model parameterization, participant characteristics). Which cognitive processes are captured by each model parameter is influenced by the task (green, orange, blue), as shown by distinct connections between parameters and cognitive processes. Different parameters within the same task can capture overlapping cognitive processes (not interpretable), and the same parameters can capture different processes depending on the task (not generalizable). However, parameters likely capture consistent behavioral patterns across tasks (thick vertical arrows).

[3] M. G. Bellemare, S. Candido, P. S. Castro, J. Gong, M. C. Machado, S. Moitra, S. S. Ponda, Z. Wang, Autonomous navigation of stratospheric balloons using reinforcement learning, Nature 588 (7836) (2020) 77–82, number: 7836 Publisher: Nature Publishing Group. doi:10.1038/s41586-020-2939-8.
URL https://www.nature.com/articles/s41586-020-2939-8

[4] M. K. Eckstein, S. L. Master, R. E. Dahl, L. Wilbrecht, A. G. E. Collins, Understanding the Unique Advantage of Adolescents in Stochastic, Volatile Environments: Combining Reinforcement Learning and Bayesian Inference, bioRxiv (2020) 2020.07.04.187971Publisher: Cold Spring Harbor Laboratory Section: New Results. doi:10.1101/2020.07.04.187971.
URL https://www.biorxiv.org/content/10.1101/2020.07.04.187971v1

[5] S. L. Master, M. K. Eckstein, N. Gotlieb, R. Dahl, L. Wilbrecht, A. G. E. Collins, Disentangling the systems contributing to changes in learning during adolescence, Developmental Cognitive Neuroscience 41 (2020) 100732. doi:10.1016/j.dcn.2019.100732.
URL http://www.sciencedirect.com/science/article/pii/S1878929319303196

[6] E. J. P. Maes, M. J. Sharpe, A. A. Usypchuk, M. Lozzi, C. Y. Chang, M. P. H. Gardner, G. Schoenbaum, M. D. Iordanova, Causal evidence supporting the proposal that dopamine transients function as temporal difference prediction errors, Nature Neuroscience 23 (2) (2020) 176–178, number: 2 Publisher: Nature Publishing Group. doi:10.1038/s41593-019-0574-1.
URL https://www.nature.com/articles/s41593-019-0574-1

[7] E. O. Neftci, B. B. Averbeck, Reinforcement learning in artificial and biological systems, Nature Machine Intelligence 1 (3) (2019) 133–143, number: 3 Publisher: Nature Publishing Group. doi:10.1038/s42256-019-0025-4.
URL https://www.nature.com/articles/s42256-019-0025-4

[8] A. G. E. Collins, Reinforcement learning: bringing together computation and cognition, Current Opinion in Behavioral Sciences 29 (2019) 63–68. doi:10.1016/j.cobeha.2019.04.011.
URL http://www.sciencedirect.com/science/article/pii/S235215461830175X

[9] T. Yarkoni, R. A. Poldrack, T. E. Nichols, D. C. Van Essen, T. D. Wager, Large-scale automated synthesis of human functional neuroimaging data, Nature Methods 8 (8) (2011) 665–670, number: 8 Publisher: Nature Publishing Group. doi:10.1038/nmeth.1635.
URL https://www.nature.com/articles/nmeth.1635

[10] R. S. Sutton, A. G. Barto, Reinforcement Learning: An Introduction, 2nd Edition, MIT Press, Cambridge, MA; London, England, 2017.

[11] M. J. Sharpe, H. M. Batchelor, L. E. Mueller, C. Yun Chang, E. J. P. Maes, Y. Niv, G. Schoenbaum, Dopamine transients do not act as model-free prediction errors during associative learning, Nature Communications 11 (1) (2020) 106. doi:10.1038/s41467-019-13953-1.
URL http://www.nature.com/articles/s41467-019-13953-1

[12] A. Mohebi, J. R. Pettibone, A. A. Hamid, J.-M. T. Wong, L. T. Vinson, T. Patriarchi, L. Tian, R. T. Kennedy, J. D. Berke, Dissociable dopamine dynamics for learning and motivation, Nature 570 (7759) (2019) 65–70, number: 7759 Publisher: Nature Publishing Group. doi:10.1038/s41586-019-1235-y.
URL https://www.nature.com/articles/s41586-019-1235-y

[13] M. Botvinick, Hierarchical reinforcement learning and decision making, Current Opinion in Neurobiology 22 (6) (2012) 956–962. doi:10.1016/j.conb.2012.05.008.
URL http://linkinghub.elsevier.com/retrieve/pii/S0959438812000876

[14] L. Xia, A. G. E. Collins, Temporal and state abstractions for efficient learning, transfer and composition in humans, Psychological Review (2021).

[15] J. X. Wang, Z. Kurth-Nelson, D. Kumaran, D. Tirumala, H. Soyer, J. Z. Leibo, D. Hassabis, M. Botvinick, Prefrontal cortex as a meta-reinforcement learning system, Nature Neuroscience 21 (6) (2018) 860–868. doi:10.1038/s41593-018-0147-8.

[16] M. K. Eckstein, A. G. E. Collins, Computational evidence for hierarchically structured reinforcement learning in humans, Proceedings of the National Academy of Sciences 117 (47) (2020) 29381–29389.

doi:10.1073/pnas.1912330117.
URL https://www.pnas.org/content/117/47/29381

[17] C. Findling, N. Chopin, E. Koechlin, Imprecise neural computations as a source of adaptive behaviour in volatile environments, Nature Human Behaviour 5 (1) (2021) 99–112, number: 1 Publisher: Nature Publishing Group. doi:10.1038/s41562-020-00971-z.
URL https://www.nature.com/articles/s41562-020-00971-z

[18] Y. Niv, Reinforcement learning in the brain, Journal of Mathematical Psychology 53 (3) (2009) 139–154.

[19] M. J. Frank, E. D. Claus, Anatomy of a decision: Striato-orbitofrontal interactions in reinforcement learning, decision making, and reversal., Psychological Review 113 (2) (2006) 300–326. doi:10.1037/0033-295X.113.2.300.
URL http://doi.apa.org/getdoi.cfm?doi=10.1037/0033-295X.113.2.300

[20] W. Schultz, A. Dickinson, Neuronal Coding of Prediction Errors, Annual Review of Neuroscience 23 (1) (2000) 473–500. doi:10.1146/annurev.neuro.23.1.473.
URL http://www.annualreviews.org/doi/10.1146/annurev.neuro.23.1.473

[21] Y. Wang, O. Toyoshima, J. Kunimatsu, H. Yamada, M. Matsumoto, Tonic firing mode of midbrain dopamine neurons continuously tracks reward values changing moment-by-moment, eLifePublisher: eLife Sciences Publications Limited (Mar. 2021). doi:10.7554/eLife.63166.
URL https://elifesciences.org/articles/63166/figures

[22] D. Meder, D. M. Herz, J. B. Rowe, S. Lehéricy, H. R. Siebner, The role of dopamine in the brain - lessons learned from Parkinson's disease, NeuroImage 190 (2019) 79–93. doi:10.1016/j.neuroimage.2018.11.021.
URL https://www.sciencedirect.com/science/article/pii/S1053811918320925

[23] A. Westbrook, R. v. d. Bosch, J. I. Määttä, L. Hofmans, D. Papadopetraki, R. Cools, M. J. Frank, Dopamine promotes cognitive effort by biasing the benefits versus costs of cognitive work, Science 367 (6484) (2020) 1362–1366, publisher: American Association for the Advancement of Science Section: Report. doi:10.1126/science.aaz5891.
URL https://science.sciencemag.org/content/367/6484/1362

[24] O. M. Vikbladh, M. R. Meager, J. King, K. Blackmon, O. Devinsky, D. Shohamy, N. Burgess, N. D. Daw, Hippocampal Contributions to Model-Based Planning and Spatial Memory, Neuron 102 (3) (2019) 683–693.e4. doi:10.1016/j.neuron.2019.02.014.
URL https://www.sciencedirect.com/science/article/pii/S0896627319301230

[25] A. M. Bornstein, K. A. Norman, Reinstated episodic context guides sampling-based decisions for reward, Nature Neuroscience 20 (7) (2017) 997–1003. doi:10.1038/nn.4573.
URL https://www.nature.com/articles/nn.4573

[26] A. G. E. Collins, M. J. Frank, Within- and across-trial dynamics of human EEG reveal cooperative interplay between reinforcement learning and working memory, Proceedings of the National Academy of Sciences 115 (10) (2018) 2502–2507, publisher: National Academy of Sciences Section: Biological Sciences. doi:10.1073/pnas.1720963115.
URL https://www.pnas.org/content/115/10/2502

[27] M. Rmus, S. McDougle, A. G. E. Collins, The role of executive function in shaping reinforcement learning, Current Opinion in Behavioral Sciences 38 (2021) 66–73. doi:10.1016/j.cobeha.2020.10.003.

[28] A. Radulescu, Y. Niv, I. Ballard, Holistic Reinforcement Learning: The Role of Structure and Attention, Trends in Cognitive Sciences 23 (4) (2019) 278–292. doi:10.1016/j.tics.2019.01.010.
URL https://www.sciencedirect.com/science/article/pii/S1364661319300361

[29] A. G. E. Collins, M. J. Frank, How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis: Working memory in reinforcement learning, European Journal of Neuroscience 35 (7) (2012) 1024–1035. doi:10.1111/j.1460-9568.2011.07980.x.

[30] R. C. Wilson, A. G. Collins, Ten simple rules for the computational modeling of behavioral data, eLife 8 (2019) e49547, publisher: eLife Sciences Publications, Ltd. doi:10.7554/eLife.49547.
URL https://doi.org/10.7554/eLife.49547

[31] S. Palminteri, V. Wyart, E. Koechlin, The Importance of Falsification in Computational Cognitive Modeling, Trends in Cognitive Sciences

21 (6) (2017) 425–433. doi:10.1016/j.tics.2017.03.011.
URL https://linkinghub.elsevier.com/retrieve/pii/S1364661317300542

[32] O. Guest, A. E. Martin, How Computational Modeling Can Force Theory Building in Psychological Science, Perspectives on Psychological Science (2021) 1745691620970585Publisher: SAGE Publications Inc. doi:10.1177/1745691620970585.
URL https://doi.org/10.1177/1745691620970585

[33] G. Blohm, K. P. Kording, P. R. Schrater, A How-to-Model Guide for Neuroscience, eNeuro 7 (1), publisher: Society for Neuroscience Section: Research Article: Methods/New Tools (Jan. 2020). doi:10.1523/ENEURO.0352-19.2019.
URL https://www.eneuro.org/content/7/1/ENEURO.0352-19.2019

[34] C. Diuk, A. Schapiro, N. Córdova, J. Ribas-Fernandes, Y. Niv, M. Botvinick, Divide and Conquer: Hierarchical Reinforcement Learning and Task Decomposition in Humans, in: Computational and Robotic Models of the Hierarchical Organization of Behavior, Springer, Berlin, Heidelberg, 2013, pp. 271–291. doi:10.1007/978-3-642-39875-$9_1$2.

[35] W. R. Uttal, On some two-way barriers between models and mechanisms, Perception & Psychophysics 48 (2) (1990) 188–203. doi:10.3758/BF03207086.
URL https://doi.org/10.3758/BF03207086

[36] D. J. Navarro, Between the Devil and the Deep Blue Sea: Tensions Between Scientific Judgement and Statistical Model Selection, Computational Brain & Behavior 2 (1) (2019) 28–34. doi:10.1007/s42113-018-0019-z.
URL https://doi.org/10.1007/s42113-018-0019-z

[37] K. Nussenbaum, C. A. Hartley, Reinforcement learning across development: What insights can we draw from a decade of research?, Developmental Cognitive Neuroscience 40 (2019) 100733. doi:10.1016/j.dcn.2019.100733.
URL http://www.sciencedirect.com/science/article/pii/S1878929319303202

[38] T. U. Hauser, G.-J. Will, M. Dubois, R. J. Dolan, Annual Research Review: Developmental computational psychiatry, Journal of Child Psychology and Psychiatry 60 (4) (2019) 412–426. doi:https://doi.org/10.1111/jcpp.12964.

[39] Q. J. M. Huys, T. V. Maia, M. J. Frank, Computational psychiatry as a bridge from neuroscience to clinical applications, Nature neuroscience 19 (3)

(2016) 404–413. doi:10.1038/nn.4238.
URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5443409/

[40] D. Y. Teller, Linking propositions, Vision Research 24 (10) (1984) 1233–1246. doi:10.1016/0042-6989(84)90178-0.
URL https://www.sciencedirect.com/science/article/pii/0042698984901780

[41] V. M. Brown, J. Chen, C. M. Gillan, R. B. Price, Improving the Reliability of Computational Analyses: Model-Based Planning and Its Relationship With Compulsivity, Biological Psychiatry: Cognitive Neuroscience and Neuroimaging 5 (6) (2020) 601–609. doi:10.1016/j.bpsc.2019.12.019.
URL https://www.sciencedirect.com/science/article/pii/S2451902220300161

[42] R. Daniel, A. Radulescu, Y. Niv, Intact Reinforcement Learning But Impaired Attentional Control During Multidimensional Probabilistic Learning in Older Adults, The Journal of Neuroscience 40 (5) (2020) 1084–1096. doi:10.1523/JNEUROSCI.0254-19.2019.
URL https://www.jneurosci.org/lookup/doi/10.1523/JNEUROSCI.0254-19.2019

[43] R. H. Kaiser, M. T. Treadway, D. W. Wooten, P. Kumar, F. Goer, L. Murray, M. Beltzer, P. Pechtel, A. Whitton, A. L. Cohen, N. M. Alpert, G. El Fakhri, M. D. Normandin, D. A. Pizzagalli, Frontostriatal and Dopamine Markers of Individual Differences in Reinforcement Learning: A Multi-modal Investigation, Cerebral Cortex 28 (12) (2018) 4281–4290. doi:10.1093/cercor/bhx281.
URL https://doi.org/10.1093/cercor/bhx281

[44] A. H. Javadi, D. H. K. Schmidt, M. N. Smolka, Adolescents adapt more slowly than adults to varying reward contingencies, Journal of Cognitive Neuroscience 26 (12) (2014) 2670–2681. doi:10.1162/jocn$_{a0}$0677.

[45] S. J. Gershman, Empirical priors for reinforcement learning models, Journal of Mathematical Psychology 71 (2016) 1–6. doi:10.1016/j.jmp.2016.01.006.
URL http://www.sciencedirect.com/science/article/pii/S0022249616000080

[46] W. Kool, F. A. Cushman, S. J. Gershman, When Does Model-Based Control Pay Off?, PLOS Computational Biology 12 (8) (2016) e1005090, publisher: Public Library of Science. doi:10.1371/journal.pcbi.1005090.
URL https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.100509

[47] T. Harada, Learning From Success or Failure? – Positivity Biases Revisited, Frontiers in Psychology 11 (Jul. 2020). doi:10.3389/fpsyg.2020.01627.
URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC7396482/

[48] R. T. Gerraty, J. Y. Davidow, K. Foerde, A. Galvan, D. S. Bassett, D. Shohamy, Dynamic Flexibility in Striatal-Cortical Circuits Supports Reinforcement Learning, Journal of Neuroscience 38 (10) (2018) 2442–2453, publisher: Society for Neuroscience Section: Research Articles. doi:10.1523/JNEUROSCI.2084-17.2018.
URL https://www.jneurosci.org/content/38/10/2442

[49] M. Watabe-Uchida, N. Eshel, N. Uchida, Neural circuitry of reward prediction error, Annual review of neuroscience 40 (2017) 373–394. doi:10.1146/annurev-neuro-072116-031109.
URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6721851/

[50] W. Dabney, Z. Kurth-Nelson, N. Uchida, C. K. Starkweather, D. Hassabis, R. Munos, M. Botvinick, A distributional code for value in dopamine-based reinforcement learning, Nature 577 (7792) (2020) 671–675. doi:10.1038/s41586-019-1924-6.
URL http://www.nature.com/articles/s41586-019-1924-6

[51] A. G. E. Collins, M. J. Frank, Opponent actor learning (OpAL): Modeling interactive effects of striatal dopamine on reinforcement learning and choice incentive., Psychological Review 121 (3) (2014) 337–366. doi:10.1037/a0037015.
URL http://doi.apa.org/getdoi.cfm?doi=10.1037/a0037015

[52] L.-H. Tai, A. M. Lee, N. Benavidez, A. Bonci, L. Wilbrecht, Transient stimulation of distinct subpopulations of striatal neurons mimics changes in action value, Nature Neuroscience 15 (9) (2012) 1281–1289. doi:10.1038/nn.3188.

[53] J. Cox, I. B. Witten, Striatal circuits for reward learning and decision-making, Nature Reviews. Neuroscience 20 (8) (2019) 482–494. doi:10.1038/s41583-019-0189-2.

[54] S. Rupprechter, L. Romaniuk, P. Series, Y. Hirose, E. Hawkins, A.-L. Sandu, G. D. Waiter, C. J. McNeil, X. Shen, M. A. Harris, A. Campbell, D. Porteous, J. A. Macfarlane, S. M. Lawrie, A. D. Murray, M. R. Delgado, A. M. McIntosh, H. C. Whalley, J. D. Steele, Blunted medial prefrontal cortico-limbic

reward-related effective connectivity and depression, Brain 143 (6) (2020) 1946–1956. doi:10.1093/brain/awaa106.
URL https://doi.org/10.1093/brain/awaa106

[55] W. van den Bos, R. Bruckner, M. R. Nassar, R. Mata, B. Eppinger, Computational neuroscience across the lifespan: Promises and pitfalls, Developmental Cognitive Neuroscience (Oct. 2017). doi:10.1016/j.dcn.2017.09.008.
URL http://linkinghub.elsevier.com/retrieve/pii/S1878929317301068

[56] F. Bolenz, A. M. F. Reiter, B. Eppinger, Developmental Changes in Learning: Computational Mechanisms and Social Influences, Frontiers in Psychology 8, publisher: Frontiers (2017). doi:10.3389/fpsyg.2017.02048.
URL https://www.frontiersin.org/articles/10.3389/fpsyg.2017.02048/full

[57] Z. A. Yaple, R. Yu, Fractionating adaptive learning: A meta-analysis of the reversal learning paradigm, Neuroscience & Biobehavioral Reviews 102 (2019) 85–94. doi:10.1016/j.neubiorev.2019.04.006.
URL http://www.sciencedirect.com/science/article/pii/S0149763418308996

[58] J. P. O'Doherty, S. W. Lee, D. McNamee, The structure of reinforcement-learning mechanisms in the human brain, Current Opinion in Behavioral Sciences 1 (2015) 94–100. doi:10.1016/j.cobeha.2014.10.004.

[59] J. Garrison, B. Erdeniz, J. Done, Prediction error in reinforcement learning: A meta-analysis of neuroimaging studies, Neuroscience & Biobehavioral Reviews 37 (7) (2013) 1297–1310. doi:10.1016/j.neubiorev.2013.03.023.
URL http://www.sciencedirect.com/science/article/pii/S0149763413000833

[60] D. Lee, H. Seo, M. W. Jung, Neural Basis of Reinforcement Learning and Decision Making, Annual review of neuroscience 35 (2012) 287–308. doi:10.1146/annurev-neuro-062111-150512.

[61] T. Yarkoni, The generalizability crisis, The Behavioral and brain sciencesPublisher: Behav Brain Sci (Dec. 2020). doi:10.1017/S0140525X20001685.
URL https://pubmed.ncbi.nlm.nih.gov/33342451/

[62] X. Liu, J. Hairston, M. Schrier, J. Fan, Common and distinct networks underlying reward valence and processing stages: A meta-analysis of functional neuroimaging studies, Neuroscience and Biobehavioral Reviews 35 (5) (2011) 1219–1236. doi:10.1016/j.neubiorev.2010.12.012.
URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC3395003/

[63] J. Davidow, K. Foerde, A. Galvan, D. Shohamy, An Upside to Reward Sensitivity: The Hippocampus Supports Enhanced Reinforcement Learning in Adolescence, Neuron 92 (1) (2016) 93–99. doi:10.1016/j.neuron.2016.08.031.
URL http://linkinghub.elsevier.com/retrieve/pii/S0896627316305244

[64] T. E. J. Behrens, M. W. Woolrich, M. E. Walton, M. F. S. Rushworth, Learning the value of information in an uncertain world, Nature Neuroscience 10 (9) (2007) 1214–1221. doi:10.1038/nn1954.
URL https://www.nature.com/articles/nn1954

[65] G. Lefebvre, M. Lebreton, F. Meyniel, S. Bourgeois-Gironde, S. Palminteri, Behavioural and neural characterization of optimistic reinforcement learning, Nature Human Behaviour 1 (4) (2017) 0067. doi:10.1038/s41562-017-0067.
URL http://www.nature.com/articles/s41562-017-0067

[66] N. Daw, S. Gershman, B. Seymour, P. Dayan, R. Dolan, Model-Based Influences on Humans' Choices and Striatal Prediction Errors, Neuron 69 (6) (2011) 1204–1215. doi:10.1016/j.neuron.2011.02.027.

[67] K. Katahira, The statistical structures of reinforcement learning with asymmetric value updates, Journal of Mathematical Psychology 87 (2018) 31–45. doi:10.1016/j.jmp.2018.09.002.
URL http://www.sciencedirect.com/science/article/pii/S0022249617302407

[68] M. Sugawara, K. Katahira, Dissociation between asymmetric value updating and perseverance in human reinforcement learning, Scientific Reports 11 (1) (2021) 3574, number: 1 Publisher: Nature Publishing Group. doi:10.1038/s41598-020-80593-7.
URL https://www.nature.com/articles/s41598-020-80593-7

[69] L. Xia, S. Master, M. Eckstein, L. Wilbrecht, A. G. E. Collins, Learning under uncertainty changes during adolescence, in: Proceedings of the Cognitive Science Society, 2020.

[70] J. H. Decker, F. S. Lourenco, B. B. Doll, C. A. Hartley, Experiential reward learning outweighs instruction prior to adulthood, Cognitive, Affective & Behavioral Neuroscience 15 (2) (2015) 310–320. doi:10.3758/s13415-014-0332-5.

[71] S. Palminteri, E. J. Kilford, G. Coricelli, S.-J. Blakemore, The Computational Development of Reinforcement Learning during Adolescence, PLoS Compu-

tational Biology 12 (6) (Jun. 2016). doi:10.1371/journal.pcbi.1004953.
URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4920542/

[72] L. Deserno, R. Boehme, A. Heinz, F. Schlagenhauf, Reinforcement Learning and Dopamine in Schizophrenia: Dimensions of Symptoms or Specific Features of a Disease Group?, Frontiers in Psychiatry 4, publisher: Frontiers (2013). doi:10.3389/fpsyt.2013.00172.
URL https://www.frontiersin.org/articles/10.3389/fpsyt.2013.00172/full

[73] W.-Y. Ahn, J. R. Busemeyer, Challenges and promises for translating computational tools into clinical practice, Current Opinion in Behavioral Sciences 11 (2016) 1–7. doi:10.1016/j.cobeha.2016.02.001.
URL https://www.sciencedirect.com/science/article/pii/S2352154616300237

[74] S.-J. Blakemore, T. W. Robbins, Decision-making in the adolescent brain, Nature Neuroscience 15 (9) (2012) 1184–1191, number: 9 Publisher: Nature Publishing Group. doi:10.1038/nn.3177.
URL http://www.nature.com/articles/nn.3177

[75] S. DePasque, A. Galván, Frontostriatal development and probabilistic reinforcement learning during adolescence, Neurobiology of Learning and Memory 143 (2017) 1–7. doi:10.1016/j.nlm.2017.04.009.
URL http://www.sciencedirect.com/science/article/pii/S107474271730062X

[76] M. K. Eckstein, S. L. Master, L. Xia, R. E. Dahl, L. Wilbrecht, A. G. E. Collins, Learning Rates Are Not All the Same: The Interpretation of Computational Model Parameters Depends on the Context, bioRxiv (2021) 2021.05.28.446162Publisher: Cold Spring Harbor Laboratory Section: New Results. doi:10.1101/2021.05.28.446162.
URL https://www.biorxiv.org/content/10.1101/2021.05.28.446162v1

[77] S. M. Groman, C. Keistler, A. J. Keip, E. Hammarlund, R. J. DiLeone, C. Pittenger, D. Lee, J. R. Taylor, Orbitofrontal circuits control multiple reinforcement-learning processes, Neuron 103 (4) (2019) 734–746.e3. doi:10.1016/j.neuron.2019.05.042.
URL https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6893860/

[78] C. K. Starkweather, S. J. Gershman, N. Uchida, The Medial Prefrontal Cortex Shapes Dopamine Reward Prediction Errors under State Uncertainty,

Neuron 98 (3) (2018) 616–629.e6. doi:10.1016/j.neuron.2018.03.036.
URL http://www.sciencedirect.com/science/article/pii/S0896627318302423

[79] S. J. Gershman, N. Uchida, Believing in dopamine, Nature Reviews Neuroscience 20 (11) (2019) 703–714, number: 11 Publisher: Nature Publishing Group. doi:10.1038/s41583-019-0220-7.
URL https://www.nature.com/articles/s41583-019-0220-7

[80] R. Frömer, C. K. Dean Wolf, A. Shenhav, Goal congruency dominates reward value in accounting for behavioral and neural correlates of value-based decision-making, Nature Communications 10 (1) (2019) 4926, number: 1 Publisher: Nature Publishing Group. doi:10.1038/s41467-019-12931-x.
URL https://www.nature.com/articles/s41467-019-12931-x

[81] W. van den Bos, R. Hertwig, Adolescents display distinctive tolerance to ambiguity and to uncertainty during risky decision making, Scientific Reports 7 (1) (2017) 40962, number: 1 Publisher: Nature Publishing Group. doi:10.1038/srep40962.
URL https://www.nature.com/articles/srep40962

[82] N. Sendhilnathan, M. Semework, M. E. Goldberg, A. E. Ipata, Neural Correlates of Reinforcement Learning in Mid-lateral Cerebellum, Neuron 106 (1) (2020) 188–198.e5. doi:10.1016/j.neuron.2019.12.032.

[83] S. D. McDougle, A. G. E. Collins, Modeling the influence of working memory, reinforcement, and action uncertainty on reaction time and choice during instrumental learning, Psychonomic Bulletin & Review 28 (1) (2021) 20–39. doi:10.3758/s13423-020-01774-z.
URL https://doi.org/10.3758/s13423-020-01774-z

[84] A. Konovalov, I. Krajbich, Neurocomputational Dynamics of Sequence Learning, Neuron 98 (6) (2018) 1282–1293.e4. doi:10.1016/j.neuron.2018.05.013.
URL http://www.sciencedirect.com/science/article/pii/S0896627318303854

[85] D. Kalashnikov, A. Irpan, P. Pastor, J. Ibarz, A. Herzog, E. Jang, D. Quillen, E. Holly, M. Kalakrishnan, V. Vanhoucke, S. Levine, QT-Opt: Scalable Deep Reinforcement Learning for Vision-Based Robotic Manipulation, arXiv:1806.10293 [cs, stat]ArXiv: 1806.10293 (Nov. 2018).
URL http://arxiv.org/abs/1806.10293

[86] A. Bakkour, D. J. Palombo, A. Zylberberg, Y. H. Kang, A. Reid, M. Verfaellie, M. N. Shadlen, D. Shohamy, The hippocampus supports deliberation during value-based decisions, eLife 8 (2019) e46080, publisher: eLife Sciences Publications, Ltd. doi:10.7554/eLife.46080.
URL https://doi.org/10.7554/eLife.46080

[87] I. Momennejad, E. M. Russek, J. H. Cheong, M. Botvinick, N. D. Daw, S. J. Gershman, The successor representation in human reinforcement learning, Nature Human Behaviour 1 (9) (2017) 680–692. doi:10.1038/s41562-017-0180-8.