Learning to Selectively Attend

Samuel J. Gershman (sjgershm@princeton.edu)
Jonathan D. Cohen (jdc@princeton.edu)
Yael Niv (yael@princeton.edu)
Department of Psychology and Neuroscience Institute, Princeton University
Princeton, NJ 08540 USA

Abstract

How is reinforcement learning possible in a high-dimensional world? Without making any assumptions about the structure of the state space, the amount of data required to effectively learn a value function grows exponentially with the state space's dimensionality. However, humans learn to solve high-dimensional problems much more rapidly than would be expected under this scenario. This suggests that humans employ inductive biases to guide (and accelerate) their learning. Here we propose one particular bias—sparsity—that ameliorates the computational challenges posed by high-dimensional state spaces, and present experimental evidence that humans can exploit sparsity information when it is available.

Keywords: reinforcement learning; attention; Bayes.

Introduction

Reinforcement learning (RL) in high-dimensional state spaces is a notoriously difficult problem in machine learning (Sutton & Barto, 1998), primarily because of the *curse of dimensionality*: The number of states grows exponentially with dimensionality (Bellman, 1957), and thus if one were naively to represent a separate value (expected reward) for each state, one would require astronomical amounts of data to effectively learn the value function (and thereby behave adaptively). Nonetheless, humans appear to learn rapidly from small amounts of data. Thus, while substantial evidence has accumulated that human behavior follows the predictions of RL models (Dayan & Niv, 2008), these models may fundamentally underestimate the learning capabilities of humans.

Following work in other areas of cognition (Braun, Mehring, & Wolpert, 2009; Kemp & Tenenbaum, 2009), we suggest that rapid learning arises from the exploitation of structured knowledge in the form of inductive biases. In particular, our proposal is that humans employ a sparsity assumption: that only one (or a small number) of dimensions is relevant at any given time for predicting reward. For example, when you are at a stoplight, only the color of the light matters, not its shape, size, etc. In other domains (such as ordering food in a restaurant), you may know that dimensional relevance is sparse, but not which particular dimensions are relevant (does it matter which restaurant it is? which table I am sitting at? who the chef is? who the waiter is?); for this purpose, one requires a learning algorithm that can exploit sparsity. We formalize this idea in terms of rational statistical inference, and present new experimental evidence that human behavior is consistent with such a model.

Central to our analysis is the idea that selective attention is a direct consequence of Bayesian inference under the sparsity assumption: Restricting attention to only a few dimensions is akin to the belief that only those dimensions are relevant for earning reward. This has the effect of reducing the space of possible value functions to a much smaller subspace.

While Bayesian probability theory stipulates the ideal observer model, in general it may not be computationally tractable to perform the necessary calculations exactly (Kruschke, 2006; Daw & Courville, 2008). We therefore consider a "hybrid" model that combines the computational efficiency of model-free RL with the statistical efficiency of Bayesian inference. We compare the ideal observer and hybrid models to a naive RL model and show that models that exploit structured knowledge better capture choice behavior in our experiment.

The Computational Problem

For concreteness, we consider one particular example of the general class of reinforcement learning problems for which the sparsity assumption holds. This example is meant to capture the abstract structure of many problems facing humans in the real world, where they must make choices between several multidimensional stimuli under conditions where most dimensions are unpredictive of reward. This example will also serve as a formal description of the task that we asked human subjects to perform, the results of which we report in a later section.

The subject plays N trials, and observes M stimuli simultaneously on each trial. The ith stimulus on trial n is denoted by a D-dimensional vector $\mathbf{x}_{ni}$, where each integer-valued component x_{nij} indicates the property of the jth stimulus dimension (for instance, [color = green, shape = triangle, texture = dots]). Each set of trials has a target dimension d (e.g., 'shape') and target property f on that dimension (e.g., 'circle'). The subject chooses a stimulus c_n on each trial and observes a binary reward r_n. The probability of reward given choice and target is

$$P(r_n = 1 | c_n, d, f, \mathbf{X}_n) = \begin{cases} \theta_1 & \text{if } x_{nc_nd} = f \\ \theta_0 & \text{otherwise,} \end{cases} \quad (1)$$

In other words, the subject receives a reward with probability θ_1 if the chosen stimulus posseses the target property on the target dimension, and with probability θ_0 otherwise.

Ideal Observer Model

Given uncertainty about the target dimension and property, an ideal observer would use Bayes' rule to infer the posterior over the target dimensions and property and then calculate the

value of the stimulus by taking the expectation of reward with respect to the posterior:

$$V_n(c) = \sum_d \sum_f P(r_n = 1 | c_n = c, d, f, \mathbf{X}_n) P(d, f | \mathcal{D}_{n-1}, \mathbf{r}_{1:n-1}), \tag{2}$$

where $\mathcal{D}_{n-1} = \{\mathbf{X}_{1:n-1}, \mathbf{c}_{1:n-1}\}$. The intuition behind the ideal observer model is that the observer weights the expected reward in each possible state of the world (i.e., target dimension and property) by the probability of the world being in that state given past observations. A key characteristic of this model is that a complete posterior distribution is maintained over states of the world, rather than a point estimate. The posterior distribution used by the ideal observer is given by Bayes' rule:

$$P(d, f | \mathcal{D}_{n-1}, \mathbf{r}_{1:n-1}) \propto P(\mathbf{r}_{1:n-1} | \mathcal{D}_{n-1}, d, f) P(d, f), \tag{3}$$

where the prior is assumed to be uniform and the likelihood is given by:

$$P(\mathbf{r}_{1:n-1} | \mathcal{D}_{n-1}, d, f) = \prod_{t=1}^{n-1} P(r_t | c_t, d, f, \mathbf{X}_t). \tag{4}$$

Note that this model describes an ideal *observer*, not an ideal *actor*: we assume that subjects are "weakly" rational in their decision rule (see the softmax choice function described below), even if they update their value function optimally.

Reinforcement Learning Models

We now consider several alternative models based on RL. The intuition behind these models is that what ultimately matters for the choice value is the *expectation* under the posterior; so incrementally updating an estimate of this expectation from experience will eventually converge to the optimal choice values, even though these updates do not make optimal use of information on each trial. The various RL models differ principally in their construction of the value function.

Naive RL Model

The naive RL model represents a separate value for every possible stimulus-dimension-property combination. Specifically, the choice value estimate is given by:

$$V_n(c) = v_n(\mathbf{x}_{nc}). \tag{5}$$

This estimate is updated according to the learning rule:

$$v_{n+1}(\mathbf{x}_{nc_n}) = v_n(\mathbf{x}_{nc_n}) + \alpha \Delta_n, \tag{6}$$

where α is a learning rate and Δ is the *prediction error*:

$$\Delta_n = r_n - V_n(c). \tag{7}$$

Although the optimal solution is learnable by this model, its highly unconstrained structure means that learning will be very slow.

Function Approximation Models

One reason why the naive RL model may be ineffective in this task is that it lacks the ability to generalize across different combinations of features. Intuitively, if you knew the target dimension and property, then the value of a stimulus should be independent of the properties on the non-target dimension. However, the naive RL model yokes these together, such that learning operates on configurations of properties and hence fails to exploit this invariance. For example, the naive RL model learns a different value for green triangles with dots and for green triangles with waves, although the texture dimension may be completely incidental and not predictive of reward.

A more structured RL model that generalizes across configurations can be obtained by constructing the value function as a linear combination of D basis functions ϕ:

$$V_n(c) = \sum_{d=1}^{D} w_n(d, x_{ncd}) \phi_d, \tag{8}$$

where the weight matrix $\mathbf{W}_n$ determines how the basis functions are combined, with one weight for each dimension-property pair. Each basis function ϕ_d is a scalar. This type of model is known as a *function approximation architecture* (Sutton & Barto, 1998). RL learning is used to update the weights according to:

$$w_{n+1}(d, x_{ncd}) = w_n(d, x_{ncd}) + \alpha \Delta_n \phi_d, \tag{9}$$

where the prediction error Δ_n is computed the same way as in the naive RL model (Eq. 7). This update can be understood as performing gradient ascent on the value function by optimizing the weight parameters (Williams, 1992).

We will consider a family of basis functions parameterized by η:

$$\phi_d = \frac{P(d | \mathcal{D}_{n-1}, \mathbf{r}_{1:n-1})^\eta}{\sum_d P(d | \mathcal{D}_{n-1}, \mathbf{r}_{1:n-1})^\eta}. \tag{10}$$

The basis function can be thought of as an "attentional focus" that encodes the subject's beliefs about what dimension is currently relevant. Thus, rather than maintaining the full posterior over target dimension and property (which may be quite computationally expensive), with the function approximation model the subject maintains the *marginal* posterior over target dimension (i.e., the probability of a dimension being the target, averaging over different target properties), which is then used to weight separate value functions, one for each dimension. When reward feedback is received, credit (or blame) is assigned to each value function in proportion to its posterior probability.[1] We refer to this model as the *hybrid* model, because it combines properties of RL learning and the ideal observer.

Different settings of η lead to several special cases of interest:

[1] Note that the subject need not maintain and update the full posterior; any procedure that estimates the marginal posterior directly is consistent with this formulation.

- $\eta = 0$: uniform weighting of dimensions (diffuse focus).
- $\eta = 1$: exact posterior weighting of dimensions (optimal focus).
- $\eta \to \infty$: maximum a posterior (MAP) weighting (myopic focus).

Other intermediate scenarios are also possible. Thus, the value of η tells us how much information about the posterior distribution the subject is using to focus attention, with $\eta = 1$ being optimal focus and $\eta = 0$ completely ignoring information from the posterior and attending equally to all dimensions.[2] When η is larger than 1, the subject discards posterior uncertainty by focusing on the mode of the distribution, and is therefore overconfident in her beliefs about the relevant dimension.

One way to interpret the function approximation model is as a neural network in which the basis functions represent attentional units focusing on different sensory channels, and the weights represent synaptic connections between the attentional units and a reward prediction unit (Figure 1). The synaptic weights are updated using RL learning (Eq. 9). This interpretation resonates with ideas in computational neuroscience that view the dorsolateral prefrontal cortex as encoding attentional or task-related biases that interact with a striatal reward prediction system (Braver, Barch, & Cohen, 1999; Rougier, Noelle, Braver, Cohen, & O'Reilly, 2005; Todd, Niv, & Cohen, 2009). The prediction error Δ driving the weight updates is thought to be signaled by midbrain dopaminergic afferents to the striatum (Schultz, Dayan, & Montague, 1997)

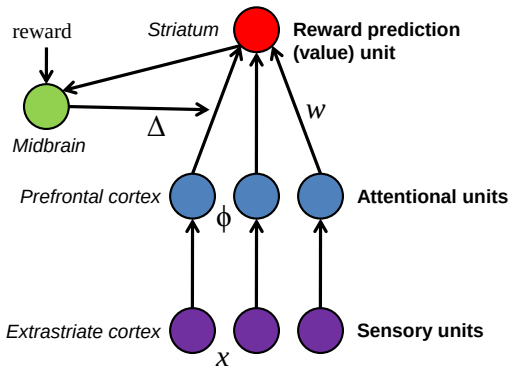

Figure 1: **Neural network interpretation of the hybrid model**.

Method

We now describe a behavioral experiment designed to quantitatively evaluate these models. Our experiment was inspired by the intra-dimensional/extra-dimensional set-shifting task (Dias, Robbins, & Roberts, 1996; Owen, Roberts, Polkey, Sahakian, & Robbins, 1991), in which subjects are asked to discriminate between visual stimuli on the basis of a particular (but unknown) dimension which they must learn from feedback, as well as the Wisconsin card-sorting task (Milner, 1963; Stuss et al., 2000). We have adapted this task to a multi-armed bandit setting, such as has been used in many previous studies of reinforcement learning (e.g., Daw, O'Doherty, Dayan, Seymour, & Dolan, 2006; Schonberg, Daw, Joel, & O'Doherty, 2007).

Participants and Stimuli

Sixteen Princeton University undergraduate students participated in the experiment and received 12 dollars reimbursement. The stimuli were triplets of stimuli varying along three dimensions: color (red, yellow, green), shape (circle, triangle, square), and texture (waves, dots, lattice). An example triplet is shown in Figure 2.

Figure 2: **Example experimental stimuli**.

Procedure

For each game, the target dimension and property are chosen randomly and with equal probability. On each trial, the subject was presented with a random triplet and asked to choose one of the stimuli. The stimuli on each trial were generated by a random permutation of the property assignments. After making the choice, the subject received feedback about whether or not her choice resulted in a reward. If the subject chose the stimulus with the target dimension/property pair, she received a reward with probability 0.75. Otherwise, reward was delivered with probability 0.25. The targets changed on each game (lasting 10-30 trials), and subjects were informed when a new game was beginning.

Choice Probabilities

To map from values to choices, we define a *policy* π_n that specifies the probability $\pi_n(c)$ of making choice c on trial n. Here we use the "softmax" policy defined by

$$\pi_n(c) = \frac{e^{\beta V_n(c)}}{\sum_a e^{\beta V_n(a)}}, \tag{11}$$

where β is an *inverse temperature* parameter, which allows us to model stochasticity in the subject's responses.

Parameter Estimation and Model Comparison

We fit the parameters of each model with MAP estimation using gradient descent and calculated the Laplace approximation (Kass & Raftery, 1995) to the log marginal likelihood

[2] It is important to note that diffuse focus is not the same as the naive RL model. For all values of η, the function approximation model is still able to generalize across different configurations, unlike the naive RL model.

(evidence) for each model m according to:

$$E_m = \ln \int_{\omega_m} P(\omega_m) \prod_{n=1}^{N} \pi_n(c_n|\omega_m) d\omega_m$$
$$\approx \ln P(\hat{\omega}_m) + \sum_{n=1}^{N} \ln \pi_n(c_n|\hat{\omega}_m) + \frac{G_m \ln 2\pi - \ln |\mathbf{H}_m|}{2}, \quad (12)$$

where ω_m is the set of parameters for model m, $\hat{\omega}_m$ is the MAP estimate of the parameters, G_m is the number of parameters (length of ω_m), and $\mathbf{H}_m$ is the Hessian matrix of second derivatives of the negative log-posterior evaluated at the MAP estimate. We then calculated the log Bayes Factor relative to chance (where all choices are equiprobable) according to $E_m - N\ln(1/3)$. A larger Bayes Factor indicates greater support for a model. Note that the chance (null) model has no parameters. In addition to comparing models based on Bayes Factors, we also calculated predictive log-likelihood on a held-out game using a leave-one-out cross-validation procedure.

For all the models, we fit an inverse temperature β, placing on it a Gamma(2,2) prior. This served to ameliorate a well-known degeneracy in models with both a temperature and learning rate, such that these parameters tend to trade-off against each other (inverse temperature becoming very large and learning rate very small). For the RL models, we fit a learning rate α, placing on it a Beta(1.2,1.2) prior, which slightly biases the fits away from the parameter boundaries. For the ideal observer model, we also allowed θ_1 and θ_0 to vary across subjects, since we only told subjects that the target would be rewarding more often than non-targets, placing on θ_1 a Beta(12,4) prior and on θ_0 a Beta(4,12) prior; these priors were chosen to have as their expected value the true—but unknown—values of θ_1 and θ_0. Finally we placed a Uniform(0,10) prior on η.

Results

Figure 3 presents the log Bayes Factors for each model, summed across subjects, along with the cross-validation resutls. Zero represents the null (chance) model in both cases. Clearly all the models do better than chance, but the naive RL model appears to perform substantially worse than the others. Overall, the hybrid model appears to best match behavior on this task. Figure 4 displays a the of log Bayes Factors for the ideal and hybrid models, showing that there are also individual differences in which model is favored for each subject.

Additional insight into these models can be gained by inspecting aggregate learning curves (the probability of choosing the optimal stimulus as a function of trials within a game). As shown in Figure 5, the naive RL model appears to consistently underestimate the speed of learning exhibited by subjects, whereas both the ideal observer and hybrid models hew closely to the empirical learning curve. One peculiarity of the learning curve is that subjects appear to learn faster than the ideal observer. We believe that this is an artifact of the softmax choice probability function: the inverse temperature parameter appears to be too low early in a game and slightly too high later in a game. No single value of the inverse temperature would be able to capture this pattern. We have also fit a model with a non-stationary inverse-temperature, but in the interest of parsimony we only report the simpler stationary model.

	Log Bayes Factor	*Held-out log-likelihood*
Ideal	5425	5620
Hybrid	**5892**	**6208**
Naive	3307	3312

Figure 3: **Model comparison results**. Highest scores are shown in bold.

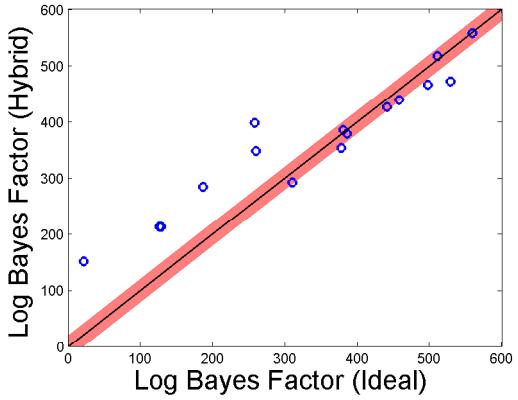

Figure 4: **Comparison of Log Bayes Factors for ideal and hybrid models**. Points above the diagonal are favored by the hybrid model. The red shaded region indicates the confidence interval outside of which one model is more likely than the other with $p < 0.05$.

Another question we can ask is whether subjects who behave more in accordance with the ideal observer or hybrid model earn more reward overall. Figure 6 does indeed show this relationship (measured as the correlation between reward earned and the log Bayes Factors for ideal relative to the hybrid model), suggesting that subjects who more effectively exploit the structure of the task tend to perform better. The correlation is significant at $p < 0.01$. There is also a correlation with total reward ($p < 0.02$) for the hybrid model log Bayes Factor relative to the null model.

Figure 7 shows the parameter estimates for η on a log-scale, demonstrating that subjects cluster around 0, corresponding to exact posterior weighting (optimal focus). This was supported by a sign test which failed to reject (p=0.45) the null hypothesis that $\ln(\eta)$ was drawn from a distribution with 0 median. Thus, within the family of possible basis functions, Bayesian attentional weighting best describes human behavior on this task.

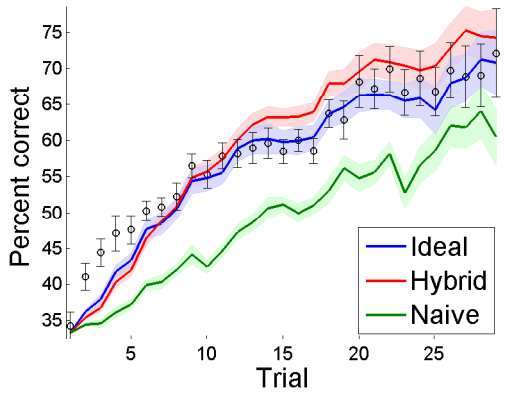

Figure 5: **Learning curves**. Probability of choosing the optimal stimulus as a function of trial within a game. The circles represent the empirical choice probability. The model-based curves are computed from the softmax equation with values fitted to choice behavior. Error bars are standard errors.

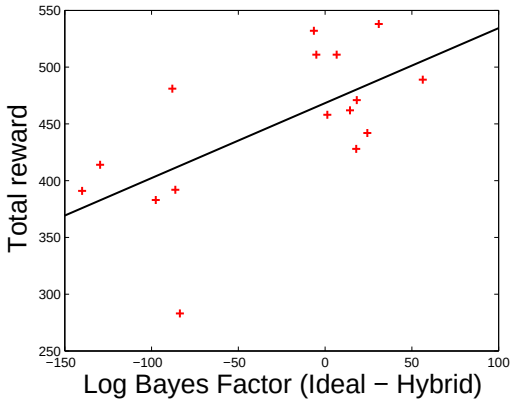

Figure 6: **Individual differences in earned reward**. On the x-axis is plotted the log Bayes Factors of the ideal model relative to the hybrid model, and on the y-axis is plotted the total reward earned. A least-squares line is superimposed on the scatter plot.

Discussion

In this paper we have posed a problem that humans face in everyday life: how to learn value functions in high-dimensional state spaces. The crucial assumption that makes this possible is that only one or a few dimensions is relevant at any given time. By employing this sparsity assumption in the machinery of Bayesian inference, the effective dimensionality of the problem is reduced. This can be understood as a kind of selective attention that is learned through experience.

Our experimental results demonstrate that humans can exploit sparsity information when it is available. We compared an ideal observer and a family of sophisticated RL algorithms against a naive RL model that ignores sparsity information. In essence, this ignorance prevents the agent from generalizing across stimulus configurations, the key ingredient to efficient learning. Our computational analysis of behavior on this task suggests that humans combine reinforcement learning with Bayesian inference, rather than using a purely Bayesian strategy. This makes sense if the brain's learning algorithms are designed to deal with high-dimensional problems for which exact Bayesian inference is intractable. The hybrid model represents a tractable compromise between the statistical inefficiency of naive RL and the computational inefficiency of the ideal observer.

The idea that selective attention can be framed as the outcome of Bayesian inference has been explored by several authors (Dayan, 2009; Rao, 2005; Yu, Dayan, & Cohen, 2009). Most relevant to our work is the competitive combination model of Dayan, Kakade, and Montague (2000), in which stimuli are assumed to vary in how reliably they predict reward. Dayan et al. (2000) showed that selective attention to particular stimuli falls naturally out of inference over the causal relationships between stimuli and reward in such a model. Our work is conceptually similar, with the extension that we model inference over dimensions, rather than

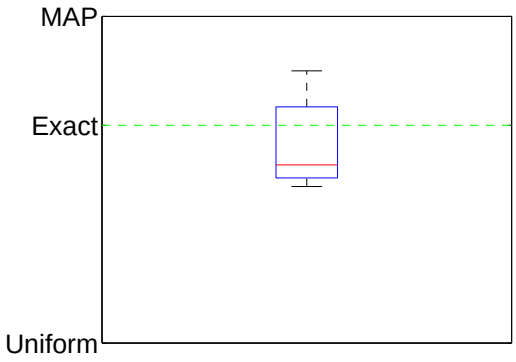

Figure 7: **Boxplot of** $\ln(\eta)$ **estimates**.

just stimuli. As emphasized by Dayan et al. (2000), the selectivity of attention in our model is based on proceses of statistical inference, rather than resource constraints. This point is particularly important to explaining how attention is *learned*; resource-limitation models, without further elaboration, do not speak to this issue.

The central role of selective attention has been extensively explored in the category learning literature, notably by Nosofsky (1986) and Kruschke (1992). The basic idea is that learned attentional weights amplify or attenuate specific stimulus dimensions in a way that facilitates category discrimination. Recently, Kruschke (Kruschke, 2006) has attempted to connect these ideas to a form of approximate Bayesian inference he dubs "locally Bayesian learning" (LBL). Much as in our work, attention arises in LBL as a consequence of weighting different hypotheses about the currently relevant stimulus dimension in response to new evidence. Our hybrid model embodies a similar idea, but by using parameterized family of

basis functions to implement attentional weighting it covers a spectrum of possible inductive influences on reinforcement learning.

While our work was partly inspired by earlier neural network models (Braver et al., 1999; Rougier et al., 2005), our goal in this paper was to step away from implementational details and interrogate computational- and algorithmic-level concerns. Future work will need to examine more systematically how the algorithmic ideas presented here map onto neural mechanisms. We are currently investigating this question with functional magnetic resonance imaging.

In conclusion, the main theoretical and experimental contribution of this paper is showing that the human RL system is more sophisticated than previous computational models have given it credit for. This may not, after all, be that surprising; many years of machine learning research have shown that the naive assumptions of previous models simply do not scale up to high-dimensional real world problems. It remains to be seen what other hidden sophistications in the RL system await discovery.

Acknowledgments

We thank Michael Todd for invaluable discussion. SJG was supported by a Quantitative Computational Neuroscience training grant from the National Institute of Mental Health.

References

Bellman, R. (1957). *Dynamic Programming*. Princeton, NJ: Princeton University Press.

Braun, D., Mehring, C., & Wolpert, D. (2009). Structure learning in action. *Behavioural Brain Research*.

Braver, T., Barch, D., & Cohen, J. (1999). Cognition and control in schizophrenia: A computational model of dopamine and prefrontal function. *Biological Psychiatry*, *46*(3), 312–328.

Daw, N., & Courville, A. (2008). The pigeon as particle filter. *Advances in neural information processing systems*, *20*, 369–376.

Daw, N., O'Doherty, J., Dayan, P., Seymour, B., & Dolan, R. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876.

Dayan, P. (2009). Load and attentional bayes. In D. Koller, D. Schuurmans, Y. Bengio, & L. Bottou (Eds.), *Advances in neural information processing systems 21* (pp. 369–376).

Dayan, P., Kakade, S., & Montague, P. (2000). Learning and selective attention. *Nature Neuroscience*, *3*, 1218–1223.

Dayan, P., & Niv, Y. (2008). Reinforcement learning: the good, the bad and the ugly. *Current opinion in neurobiology*, *18*(2), 185–196.

Dias, R., Robbins, T., & Roberts, A. (1996). Dissociation in prefrontal cortex of affective and attentional shifts.

Kass, R., & Raftery, A. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*(430).

Kemp, C., & Tenenbaum, J. (2009). Structured statistical models of inductive reasoning. *Psychological review*, *116*(1), 20–58.

Kruschke, J. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, *99*(1), 22–44.

Kruschke, J. (2006). Locally Bayesian learning with applications to retrospective revaluation and highlighting. *Psychological Review*, *113*(4), 677–698.

Milner, B. (1963). Effects of different brain lesions on card sorting: The role of the frontal lobes. *Archives of Neurology*, *9*(1), 90.

Nosofsky, R. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General*, *115*(1), 39–57.

Owen, A., Roberts, A., Polkey, C., Sahakian, B., & Robbins, T. (1991). Extra-dimensional versus intra-dimensional set shifting performance following frontal lobe excisions, temporal lobe excisions or amygdalo-hippocampectomy in man. *Neuropsychologia*, *29*(10), 993–1006.

Rao, R. (2005). Bayesian inference and attentional modulation in the visual cortex. *Neuroreport*, *16*(16), 1843.

Rougier, N., Noelle, D., Braver, T., Cohen, J., & O'Reilly, R. (2005). Prefrontal cortex and flexible cognitive control: Rules without symbols. *Proceedings of the National Academy of Sciences of the United States of America*, *102*(20), 7338.

Schonberg, T., Daw, N., Joel, D., & O'Doherty, J. (2007). Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. *Journal of Neuroscience*, *27*(47), 12860.

Schultz, W., Dayan, P., & Montague, P. (1997). A neural substrate of prediction and reward. *Science*, *275*(5306), 1593.

Stuss, D., Levine, B., Alexander, M., Hong, J., Palumbo, C., Hamer, L., et al. (2000). Wisconsin Card Sorting Test performance in patients with focal frontal and posterior brain damage: effects of lesion location and test structure on separable cognitive processes. *Neuropsychologia*, *38*(4), 388–402.

Sutton, R., & Barto, A. (1998). *Reinforcement Learning: An Introduction*. Cambridge, MA: MIT Press.

Todd, M., Niv, Y., & Cohen, J. (2009). Learning to use Working Memory in Partially Observable Environments through Dopaminergic Reinforcement. In *Neural information processing systems* (pp. 1689–1696).

Williams, R. (1992). Simple statistical gradient-following algorithms for connectionist reinforcement learning. *Machine Learning*, *8*(3), 229–256.

Yu, A., Dayan, P., & Cohen, J. (2009). Dynamics of attentional selection under conflict: Toward a rational Bayesian account. *Journal of Experimental Psychology: Human Perception and Performance*, *35*(3), 700–717.