

Major sources of computational complexity in complex decision-making

Jan Drugowitsch^{1,†} and Alexandre Pouget^{2,†}

¹Department of Neurobiology, Harvard Medical School, Boston, MA

²Department of Basic Neurosciences, University of Geneva, Switzerland

[†]Corresponding authors: jdrugow@gmail.com and alex.pouget@gmail.com

What makes decision-making hard, and what determines the time it takes to make a decision? For simple decisions, the standard answer implicates neuronal noise: its presence makes decisions hard, and averaging it out comes at the cost of long reaction times. We argue that this explanation is unlikely to hold for complex decisions. Instead, complex decisions are constrained by two main factors: memory retrieval and value computation. Indeed, most decisions require retrieving relevant information from memory, and use it to compute the choice options' values. For the large memories of vertebrate brains, both operations can be extremely complex even if the neural circuits implementing them are perfectly noiseless. Yet, the importance of these factors has not yet been fully recognized in systems neuroscience, which tends to focus on tasks in which values are retrieved from simple noisy look-up tables. The interrogation of more complex and realistic tasks, similar to the ones used in human research, might help bridge this gap.

Introduction

We consider here the general problem of value-based decision-making; that is, decisions that require a choice between two or more options whose subjective values can only be determined from information stored in long-term memories (Shadlen & Shohamy, 2016). Such value-based decisions span a wide range – from relatively simple, such as deciding which of two fruits to eat, to highly complex, such as choosing a destination for your vacation or the next move in a game of chess. This contrasts with perceptual decision-making, which is based primarily on sensory information that is readily available, such as deciding whether a set of dots move to the right or left side of the screen (Hanks & Summerfield, 2017; Shadlen & Newsome, 1996).

Like any other type of decisions, value-based decisions are subject to speed accuracy trade-offs in the sense that the probability of picking the best option can only increase at the cost of spending more time deliberating (Chittka et al., 2009; Luce, 1991; Schouten & Bekker, 1967; Wickelgren, 1977). In simple perceptual decision-making tasks, the increase in accuracy with time is thought to reflect a very basic mechanism: neurons are noisy (Shadlen & Newsome, 1998) but the noise can be averaged out over time, thus decreasing uncertainty and, consequently, increasing the accuracy of the decision (Shadlen et al., 1996). According to this view, internal neuronal noise is the main factor determining decision time (Gold & Shadlen, 2007; Palmer et al., 2005). We argue here that, in complex value-based decision-making, neuronal noise plays only a minor role in determining decision time and more generally the speed accuracy trade-off.

Speed accuracy trade-offs in decision-making apply to all systems, even computers. The more time devoted to a decision, the more likely it is correct. However, in computer science, and in contrast to neuroscience, internal noise is a non-issue since digital computers are effectively noise-free. The only noisy part in such systems are the sensors, such as cameras or microphones. For these sensors, more time can indeed result in better signal-to-noise ratio via averaging, but once inside the computer, further averaging is of no use since any additional noise is carefully corrected for. Yet, computers nonetheless feature a speed accuracy trade-off

for complex problems (e.g., Huang et al., 2017), for two reasons. The first has to do with retrieving memories, an extremely complex – and potentially slow – process for large databases. The second has to do with computing the value associated with an option given a retrieved memory. In both cases, the more time devoted to these computations, the more accurate the decision. In the case of memory, the probability of retrieving the relevant memory for a given decision grows with search time, while for value computation, the precision with which the value is being computed scales with the time devoted to the computation. This is exemplified in traditional computer chess programs, that evaluate the quality of a move by considering its consequences for future board positions (Russell et al., 2010). The further they look into the future, the more accurate this evaluation, but the more time they take doing so (Ferreira, 2013). Furthermore, this evaluation requires recalling the value of future board positions, which is more precise when using more complex value heuristics that might take longer to compute.

We argue here that the same must be true in the brain: for complex, value-based tasks, decision-making is primarily constrained by two distinct computations: memory retrieval and computational complexity, as opposed to neuronal noise. In addition, we discuss another critical factor, namely, how to turn these computations into a policy – a way of deciding whether or not to take more time and access more memories, or make an immediate decision.

Simple versus complex decisions

Let's consider a simple choice such as whether you want to eat a peach or a mango for dessert. To evaluate the respective values of these choices, you could recall from memory the values you experienced every time you ate either of these fruits (Shadlen & Shohamy, 2016). Assuming each experience led to a slightly different value, sampling these memories would generate two time-series of noisy momentary values, which you can then use to estimate the value of each fruit (**Fig. 1a**). Collecting more memory samples implies that you can estimate the expected value of the two options with higher accuracy. However, whatever time is spent on forming this choice is not spent on other activities. It is therefore in your interest to decide as quickly as possible. The typical approach to resolving this speed-accuracy trade-off consists of seeking a policy, which we will refer to as the *optimal policy*, that maximizes the reward (the pleasure you get from eating the peach or mango) relative to the cost of time and mental operations.

To derive the optimal policy, the decision maker first needs to compute the probability distribution over values (in Bayesian terms, the *belief*) for each of the options, given all the evidence (or samples, we use the two terms interchangeably) that has been collected up to the current time. We will use $p(v_k | e(1:T))$ to denote the posterior distribution over the value, v_k , of option k (for this example either 'peach' or 'mango') given the time series of T memory samples, denoted $e(1:T)$. According to Bayes' rules, this posterior distribution is proportional to the product of two quantities. The first quantity is the prior over the value, $p(v_k)$, that is, the frequency, or probability, with which we are of being offered choices with value v_k in our life. For instance, we face small-valued choices every day, such as deciding on dessert at the end of a meal, while high-valued choices, like your next vacations, only happen a few times per year and have, therefore, a much smaller prior probability. The second quantity is the likelihood function, $p(e(1:T) | v_k)$ (**Fig. 1a**). This tells us how likely it is to observe the sequence of evidence, $e(1:T) = \{e(1), e(2), \dots, e(T)\}$, if option k has value v_k . For instance, for simple decisions, it is typically assumed that a single sample of evidence at time t , $e(t)$, is the option's true value, v_k , corrupted by Gaussian noise, arising for instance from neuronal noise in memory circuits (Ratcliff, 1978; Shadlen & Shohamy, 2016; Tajima et al., 2016). Under this assumption, $p(e(t) | v_k)$ is a Gaussian function of v_k . If evidence is drawn over time from memory in a statistically independent way, $p(e(1:T) | v_k)$ can be factorized into a product of marginal likelihoods, $p(e(t) | v_k)$, which significantly simplifies incrementally updating the posterior with each recalled memory (Gold & Shadlen, 2007). If, in addition to assuming that

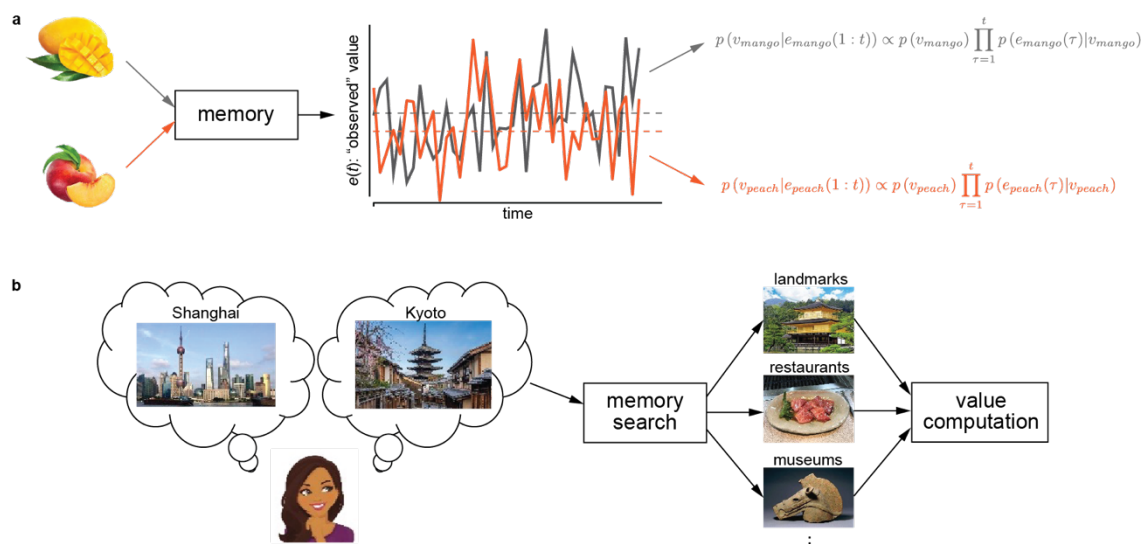


Figure 1: Simple versus complex value-based decisions. **a)** Simple choices, like deciding whether to eat a peach or a mango. For such choices, models typically assume that memory provides a time series of noisy samples of the value of each object (the black and red traces), from which the agent can infer probability distributions over the values of the options, which in turn determine the choice. **b)** For complex decisions, such as choosing the destination for your next vacation, the process must involve at least two steps which are considerably more complicated. First, relevant facts for the choices at hand must be retrieved from memory, such as the list of museums or restaurants, in the case of cities. Second, the value of the options, or rather a probability distribution over these values must be inferred from the retrieved facts. Both steps are subject to the speed-accuracy trade-off. The more time is devoted to a memory search, the more likely it is to be exhaustive. Likewise, the accuracy of the value computation is dependent on the time allocated to the computation.

the likelihood function is Gaussian, we further assume that the prior is Gaussian, the posterior distribution--the product of the prior and likelihood-- is also Gaussian and easy to compute.

Next, the decision maker needs to determine an optimal stopping policy, i.e., a rule that determines when to stop collecting samples before making a choice (Bogacz et al., 2006; Drugowitsch et al., 2012; Drugowitsch, Deangelis, et al., 2014; Furl & Averbeck, 2011; Gold & Shadlen, 2007; Tajima et al., 2016; Wald, 1945). In the Gaussian case, both the optimal stopping policy and computing the posterior can be implemented exactly with a very simple model, known as drift diffusion model (DDM), which we review in a next section. A DDM accumulates over time the difference in momentary evidences across the two choices and stops whenever this accumulated evidence exceeds a threshold, which acts as a stopping bound (Chernoff, 1961; Tajima et al., 2016).

This general strategy – computing the posterior distribution over the available options and then finding a stopping policy, usually in the form of boundaries on some function of the posterior – is applicable to all decisions, whether value-based or not (Tajima et al., 2016, 2019). However, this seemingly simple strategy can hide a great degree of complexity. This complexity arises from at least two interacting components: computing the posterior $p(v_k|e(1:T))$, and using this posterior to decide when to make a choice. In our simple example, in which all the evidences are drawn from a Gaussian distribution, the optimal policy reduces to a DDM. In general, however, computing the likelihood, the posterior and the optimal stopping policy is much harder.

For instance, imagine you’ve narrowed down your next vacation to either Shanghai or Kyoto. Unless you have visited these cities multiple times, you don’t have the option of averaging over samples of past value evaluations. An alternative strategy would be to collect a variety of facts from your memory that would help constrain the values of these two cities (**Fig. 1b**). For instance, their values will necessarily depend on urban activities that you would enjoy, such as museums, restaurants, concerts, architectural landmarks, etc. But which memories should you recall to inform your decision? For example, recalling that both cities are located

in Asia does not differentiate between the two cities and therefore does not help to decide, whereas recalling that they are in China and Japan might. Moreover, how do you infer the values of the two cities given the memories you have recovered? And how do you determine that you have searched your memories long enough; that is, when do you stop collecting more evidence? Clearly, at this point we are no longer dealing with just a time series of values drawn from Gaussian distributions. As a result, computing the posterior distribution over the values of each options, and the associated stopping bound, will be considerably more complicated.

As mentioned in the introduction, we argue below that these are indeed the three main difficulties in complex decision making: how to efficiently search a very large database for relevant facts, how to compute the posterior over option values given these facts, and coming up with a policy for stopping either of these operations. It is the combination of these three factors that determine the speed accuracy trade off.

Drift diffusion models

Before discussing complex decisions, let us return to the simple peach vs mango example above. In this example we assumed evidence to be drawn independently and identically distributed from Gaussian distributions with means equal to the actual values of the option. What is the best way to handle such choices? We could treat each choice as a single, isolated choice, but in the course of our life we have many such choices to consider. For instance, within a meal we may choose between several appetizers, entrees, dessert (peach vs mango) and after-dinner drinks. In such a situation, it makes more sense to ask how to optimize a sequence of such binary choices rather than a single one (Drugowitsch et al., 2012; Tajima et al., 2016). The faster we decide each choice, the quicker we can move to the next one. This suggests a policy in which, at each time step, say every 100ms, we collect two samples of value (one per choice option) from our memory, compute the difference, and accumulate these differences over time until their sum exceeds some stopping bound (**Fig. 2b**). In most models, the height of the stopping bound either remains the same at all times, or collapses over time. If it collapses progressively, less accumulated evidence is required to make a choice as time elapses (**Fig. 2b**). Implementations of such a strategy are known as drift diffusion models, or DDMs (Ratcliff, 1978; Ratcliff & McKoon, 2008; Tajima et al., 2019). If the prior over values in the world is Gaussian, the values accessed by memory are corrupted by Gaussian noise with identical variance, samples are drawn in pairs, and the bound is chosen correctly (and in particular collapses over time), DDMs maximize the average reward rate (Chernoff, 1961; Tajima et al., 2016).

In DDMs, the decision time depends on the drift rate, which is the difference in the true mean values of the two options (**Fig. 2a**). If the two options have nearly identical means, the drift rate is small, and reaction times are long. This predicts long reaction times *even when both values are very large*. This seems counterintuitive: why would a decision maker agonize over a choice when both options are good? It would seem to make more sense to pick an option quickly, since a high reward is guaranteed. This is true indeed but only if the decision maker knows in advance that both options are very likely to have similar high values. However, in real life, there is no reason to believe that most choices will have similar values. It is more realistic to assume that values are drawn independently from the same Gaussian distribution, in which case the options will frequently have very dissimilar values. If dissimilar values are common, it is advantageous to use a decision-making strategy that leads to quick decisions when the two options have dissimilar values, at least when the values follow Gaussian distributions (Tajima et al., 2016). Such a strategy is precisely what a DDM provides, since it is sensitive to the difference in values: the larger the difference, the faster the decision (**Fig. 2b** left vs. right). This comes as the cost of slow decisions for choices with similar values but that's fine as long as such occurrences are rare.

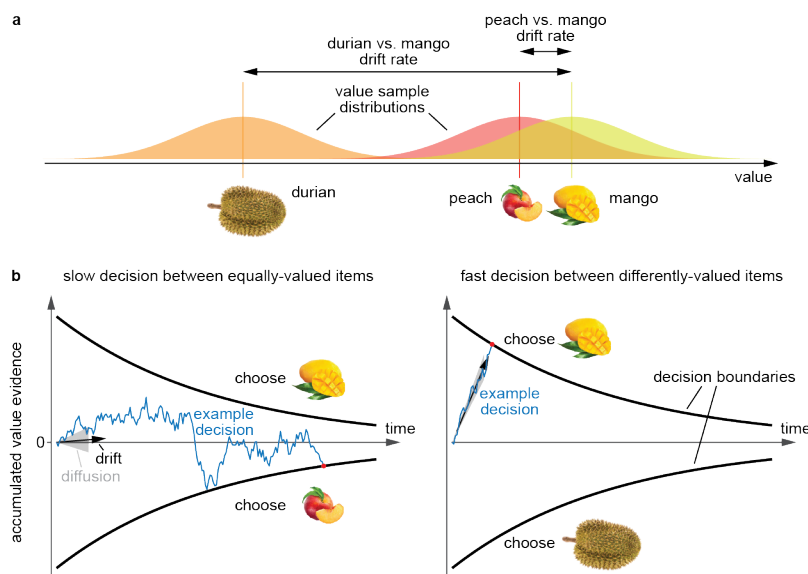


Figure 2: Drift diffusion models (DDMs) of simple decision. a) For simple choices, it is typically assumed that value samples are drawn from memorized past experiences for either choice item (Fig. 1a). DDMs assume these samples to be independently drawn from a Gaussian distribution, centered on the item's actual value (vertical lines), here illustrated for two subjectively similarly valued items (peach & mango) and one subjectively lower-valued item (durian; generally considered unappealing). b) In DDMs, the difference between choice items of sampled item values is accumulated over time until one of two decision boundaries (solid black) is reached, triggering the corresponding choice. The accumulated evidence resembles a random walk (blue example trajectory) with deterministic drift (black arrow) equal to the actual item value difference, and the stochastic diffusion (grey shaded area showing diffusion SD) reflecting the noisy value samples. This results in a counterintuitive behavior: choice between two similarly high value items, such as the mango and the peach in our example, are slow because the difference in actual values is small, resulting in a low drifting rate. This is counterintuitive because one would imagine that choosing should be fast when all options are highly valuable.

Interestingly, humans appear to follow the classic DDM strategy, that is, deciding when the accumulated difference in value between two options exceeds some bound. Indeed, DDMs have been shown to provide exquisite fits to psychometric and chronometric curves in perceptual and value-based decision-making (Krajbich et al., 2010; Palmer et al., 2005; Ratcliff & McKoon, 2008). Despite their success, however, DDMs are unlikely to provide a general framework for decision-making beyond the most basic tasks. The DDM strategy is optimal only under very restrictive conditions (Bogacz et al., 2006; Drugowitsch et al., 2012; Drugowitsch, Deangelis, et al., 2014; Tajima et al., 2016). For example, the information each evidence sample provides needs to be independent of all others, and their collective information must be summarizable by a single variable, such as their cumulated sum, that evolves towards the stopping bounds. As soon as these conditions are violated, DDMs are suboptimal (Drugowitsch, Moreno-Bote, et al., 2014; Jang et al., 2021; Tajima et al., 2016). These restrictive conditions are especially likely to be violated in complex decision-making tasks. In such tasks, a set of memorized facts or features is retrieved from memory, from which value must be inferred. Optimal decision making reduces to DDMs only if the memories can be turned into noisy samples of the true value, the variance of the noise is the same for all options, and the prior on value is Gaussian. As we discuss below, none of these conditions are satisfied in complex decision-making tasks.

Memory retrieval

In computer science, computing time is of paramount importance. Enormous efforts are devoted to developing algorithms that scale as well as possible with the size of the problem. In fact, the computational complexity of a problem is defined precisely with respect to this scaling (Moore & Mertens, 2011). Likewise, databases must be carefully organized to minimize retrieval time and maintenance. A poorly organized database can considerably slow down retrieval time, or can lower the probability of recovering relevant information, which

in turn can dramatically alter the performance of an algorithm. For instance, when asked to answer a query such as ‘name animals similar to a horse’, computing time and performance would be largely improved if one were to organize the data according to a taxonomic tree rather than in simple list of animals along with their properties (Kemp & Tenenbaum, 2008) (Fig. 3a,b).

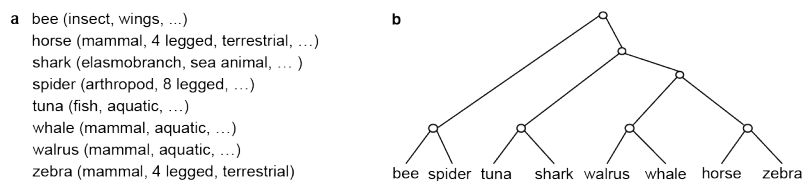


Figure 3: Two different representational structure for animals. a) Alphabetical list along with the properties of each animal. Search time is efficient when looking for a specific animal but answering questions like “find an animal similar to zebra” would scale with the number of items in the list. **b)** A tree-like representation. In this case, finding a specific animal would take more time but finding similar animals is very efficient since they are close-by leaves of the tree.

The same constraints must apply to the nervous system (Austerweil et al., 2012). The human brain presumably stores several million memories. How do we search through them? How are they organized? How do we know that we have recovered all the relevant information supporting a given decision? These are extremely complex issues. Ultimately, we can’t be sure we have retrieved all the relevant, stored information and whether it is indeed possible to retrieve all the stored information in a reasonable amount of time. We have all experienced situations in which we spent hours if not days to reach a decision, such as choosing between different careers, or picking a school for our kids, simply because we want to make sure we have left no stone unturned, that we have indeed considered all the relevant facts.

Indeed, given the inherent uncertainty in this process, memory retrieval faces the standard speed accuracy trade-off: the more time is spent on retrieving relevant memories, the more accurate the retrieval will be (Osth et al., 2018; Ratcliff, 1978). To find the optimal compromise, the decision maker should have knowledge of the statistics of memory retrieval, which correspond to the likelihood function associated with each recalled memory e . For simple decisions, we have assumed this likelihood $p(e | v_k)$ to depend on each option’s value v_k , in line with the idea that the retrieved memories are simply samples of previous value evaluations. For complex decisions, we either have no memories of previous value evaluations (e.g., we are considering going to Kyoto even though we have never been there), or memories that only apply in very restrictive contexts which limits their use. Thus, we instead use likelihood $p(e | c_k)$, where c_k is the identity of option k (e.g., c_1 =“Shanghai”, c_2 =“Kyoto”).

Conditioning on the choice option’s identity c_k rather than its value v_k makes explicit that memories are recalled when cued with the option’s identity rather than being noisy samples of the option’s value. This also reflects that, as already discussed, not all memories are necessarily informative about the option’s value. This likelihood $p(e | c_k)$ is what is commonly known as a generative model, that is, a probabilistic model of how the decision maker recalls relevant information from memory. For instance, if you’re pondering a trip to Paris, your memory circuits are very likely to reactivate the memory of the Eiffel tower, which is to say that $p(e=\text{Eiffel tower} | c_k=\text{Paris})$ is very high, while the probability that you will remember the Guimet museum, which hosts a unique collection of art from South-East Asia, is much smaller. Combining this likelihood with a model for how memories inform the choice option’s value ($p(v_k | e(1:T), c_k)$) results in a full model that in turn results in value estimates from which we can determine the optimal stopping time.

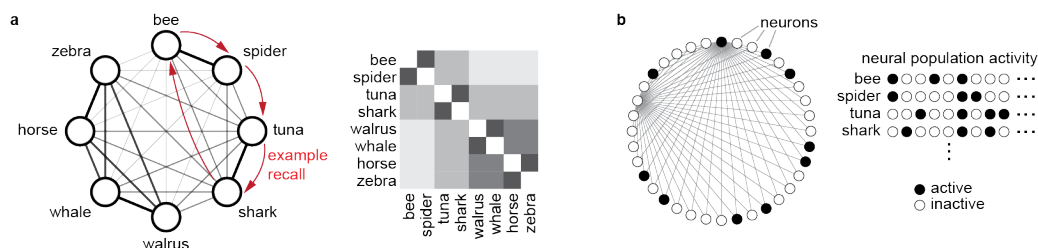


Figure 4: A network model for memory retrieval. **a)** Memory retrieval by traversing a semantic network. The similarity between memories (right, darker = more similar, from Fig. 3b) determines the connectivity strength between nodes in a semantic network (left). Memory recall follows the strongest connections while excluding just visited memories. The red arrows show one example recall sequence, starting with “tuna”, followed by “shark”, “bee”, and “spider”. Recall terminates upon repetition of the same memory sequence. **b)** A Hopfield network (fully connected, here only illustrated for two neurons), in which memories are encoded by sparse population pattern, called, attractors, can implement the recall strategy from **a**. Here, more similar activity patterns reflect more similar memories (shown as neuronal population activity patterns). Memory transitions are enforced by fluctuating inhibitory currents. Figure based on principles discussed in (Katkov et al., 2017; Romani et al., 2013).

The problem is that characterizing $p(e|c_k)$ requires a thorough understanding of how memories are organized in the nervous system, as well as how they can be queried by the rest of the brain, of which we have only a scant understanding. Several studies suggest that memory retrieval is consistent with random diffusion in semantic networks (Austerweil et al., 2012).

More recently, Tsodyks and colleagues demonstrated that memory recall can be modelled with Hopfield networks, a class of recurrent networks in which memories correspond to stable patterns of activity, also known as point attractors. For instance, the memory for a bee in Fig. 4b correspond to a pattern in which neurons 1, 4, and 6 are active while the other neurons are silent. Memory recall in this model starts by converging onto an attractor, the first recalled memory, which subsequently fades away due to adaptation, leading the network to fall into another point attractor, corresponding to another memory, and so forth (Katkov et al., 2017; Romani et al., 2013) (Fig. 4). To the extent that the topology of the Hopfield network, that is, the spatial proximity of the attractors, reflects the representational proximity of the stored items in a semantic network, this mechanism is similar to a random diffusion in a semantic network.

The topology of the semantic network remains to be determined but, in some cases, it seems clear that it is learned from experience. For instance, Behrens and colleagues recently suggested that conceptual spaces, which are very much like semantic networks, are organized in 2D maps in the hippocampus and entorhinal cortex (Constantinescu et al., 2016). Furthermore, Knudsen and Wallis proposed this map to be directly over the space of values (Knudsen & Wallis, 2021). Searching through these spaces is then equivalent to a navigation problem in a physical environment. Having such conceptual maps raises the possibility that searches could be goal-directed by learning paths within a semantic network to retrieve specific memories, similar to how we direct our saccades to actively gather information about our visual environment (Yang et al., 2016).

In fact, one could imagine that memory retrieval involves simulations of complex models of the environment, and does not merely rely on diffusion through 2D conceptual spaces. For instance, when choosing between Shanghai and Kyoto, one could mentally simulate what a typical weekend would be like in each city in order to reactivate the proper memories, such as museums, restaurants or architectural landmarks. This strategy of focusing on memories that best help deciding between the available choice options would be vastly more effective than a random search but would require a sophisticated model of the environment. This is precisely what cortical circuits, particularly in prefrontal and frontal areas, might provide (Benoit et al., 2014). The vast expansion of these areas in the human brain is quite likely related

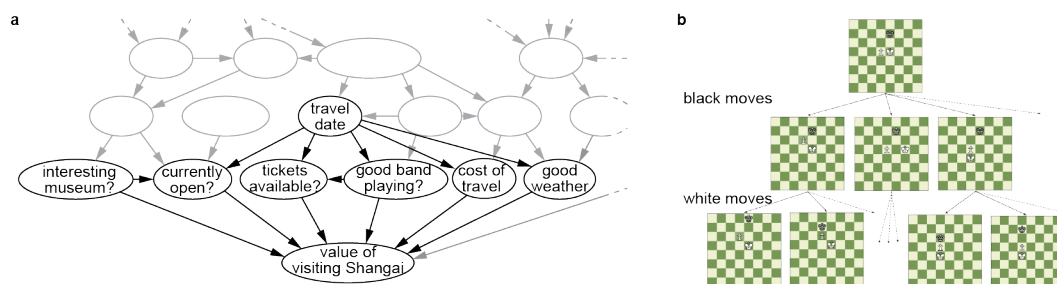


Figure 5: Bayesian networks and internal simulations for decision making. a) Excerpt of an example Bayesian Network that might be used to compute the value of visiting Shanghai. Ellipses show random variables (examples in black) whose dependency structure is specified by connecting arrows. “tickets available?” for example, depends on “travel date” and “good band playing?”. More formally, its probability is the conditional probability distribution $p(\text{tickets available?} \mid \text{travel date, good band playing?})$. b) Similar to human chess players, game-playing algorithms consider future board positions and their associated value when planning the next move.

to humans’ remarkable abilities to develop and simulated complex models of the environment.

Value computation

As we have seen, the statistic of memory retrieval is a likely key determinant of optimal decision-making policies but an equally important factor is the computation of the value themselves.

In the simplest case, a value might just be a weighted sum of the perceived value of each feature of each of the recalled memories (e.g., Iigaya et al., 2020), but, more generally, complex cases require nonlinear combinations. For instance, the fact that your favorite band is playing in Shanghai shouldn’t have much of an impact on how much you value Shanghai, unless you also know you can get tickets. In this case, the value is related to the product of these two features, which is to say, it is nonlinear. In general, the more complex and nonlinear the computation, the more computation time it will require. There are exceptions to this rule, such as deep networks that can compute remarkably complex nonlinear functions in about the same time as linear functions (for instance, switching the activation of neurons in a deep network from nonlinear to linear, barely affects simulation time). However, there are many situations in which simple feedforward architectures may not be sufficient to compute value. As we have pointed out above, the decision maker should try to infer a posterior distribution over value given the information retrieved in memory. In general, this would involve performing inference on a Bayesian network, which can be quite involved when the statistical relationship between the variables are complex (Fig. 5a). For instance, for complex inference, the method of choice is known as Monte Carlo Markov Chain (Gilks et al., 1998), which involves drawing numerous samples from the posterior distribution, a time-consuming step. If the brain were to use a similar approach (Orbán et al., 2016), a large fraction of the decision time would be spent on this step, leading to yet another speed accuracy trade-off: how long should the brain sample this network before stopping and deciding? Indeed, with any sampling technique, accuracy of inference scales with the number of samples and therefore time.

In addition, for some problems, it might make sense to rely on further internal simulations, not to retrieve facts but to generate states that can help evaluate the value of the current options. Going back to our example of chess, the current options are all the possible legal moves given the current board configuration. The decision faced by the player at every turn consists of picking one move among the legal ones. In computer programs, this involves simulating several moves forward, and evaluating the resulting values of the board configuration, a process known as tree search (Fig. 5b; (Russell et al., 2010)). Therefore, in this case, the value of the current moves is not obtained by recovering stored board positions from memory, along with their associated values. Instead, simulations are used to generate future

board positions several moves ahead and use a function evaluator (e.g., a deep neural network; (Silver et al., 2018)) to evaluate the value of these future board configurations, which are then used to estimate the value of the current moves. Although no memory search is required here, such simulations are time consuming and would involve a speed accuracy trade-off.

The optimal stopping policy

So far, we have argued that what makes complex decision making hard is memory retrieval and value computation, as opposed to noise averaging in standard models of perceptual decision-making. Both computations are subject to speed accuracy trade-off in the sense that the more time is devoted to these operations, the more accurate the decision will be. Which raises a third problem: when should the agent make a choice between the options? How should the agent determine that enough memories have been searched and that the computation of the values is accurate enough?

The goal of the memory search and value computation is to infer a posterior distribution over the value of all available options. However, knowing the posteriors is not enough to determine a stopping policy – that is, it's not enough to decide whether to choose an option or to gather more evidence (in the form of more memories and additional computation). For that, two more quantities are needed: an estimate of how much the two distributions over value will change if more evidence is received, and the cost of spending time accumulating more evidence versus doing something else. The stopping policy then involves a tradeoff: is additional information worth the cost of time?

Estimating how much information is gained by gathering more evidence is a hard problem – at least as hard as computing the posterior. Except for very simple generative models, doing this exactly is, at the very least, computationally costly, and in the typical case it's intractable (Griffiths et al., 2019). Qualitatively, however, simple heuristics may work well: observe how much the posteriors change with new memories, and when they stop changing very much, make a decision (Drugowitsch et al., 2012; Vul et al., 2014). But how much is “very much”? This is hard to answer for two reasons. First, as discussed above, single memories can have a massive effect on value. So, the fact that the posterior hasn't changed much with the last few memories doesn't mean it won't when a new memory arrives, and that must be taken into account. The second reason has to do with the cost of time. In a laboratory setting, where participants perform a single task, the cost of time is simply the average reward rate (Tajima et al., 2016). In realistic situations, one almost always has the option of doing many things. For instance, if there's a critical deadline at work, it's likely that little time will be spent planning a vacation; whereas if work is especially easy, a great deal of time will be spent planning. In addition, simple things like vacation rate matters: if you take one vacation in your life you're going to spend a lot more time planning it than if you take a vacation every six months.

In summary, the time spent on hard decisions – in this case, how many memories to recall before making a decision – depends strongly on the reward structure in the rest of a person's life. This can vary wildly, both from one person to the next, and over time for each individual; concomitantly, the decision time will vary wildly. How people determine a stopping policy that is even remotely good under these circumstances is a bit of a mystery.

Buridan's ass revisited

Buridan's ass is a well-known dilemma in philosophy in which an equally thirsty and hungry donkey happens to be at an equal distance of equally appetizing sources of water and food. How could the animal break the tie? Clearly, such a conundrum only exists if both choice options are exactly equally appetitive. The tie is easily broken as soon as their associated value signals are noisy, such as assumed by even the simple diffusion models. Nonetheless, studies have indeed shown that it takes people a long time to decide between two options with

Text box 1: Averaging out the noise. The typical recipe to average out the noise in a series of samples all drawn independently from the same distribution is simple. It requires adding the sample and dividing by the total number of samples, yielding the empirical mean of the distribution. Note, however, that the division by the total number of samples is not required to average out the noise. Indeed, consider the variable corresponding to the sum of the samples. The mean of this variable is proportional to the number of samples. Its variance is also proportional to the number of samples, which implies, that its standard deviation scales with the square root of the number of samples. This implies in turn that the signal to noise ratio, the ratio of the mean over the standard deviation, grows with the square root of the number of samples. Therefore, simple accumulating samples without dividing by their number, as it done in the DDM, already increase the signal to noise ratio, effectively averaging out the noise.

comparable high values (Krajbich et al., 2010). Why agonize over such decisions? Wouldn't the decision maker be better off just picking either option right away and then move on to the next decision? The standard explanation for this counterintuitive behavior is the one we discussed previously, and which relies on DDMs: human decision making is well modeled by drift diffusion models, which predicts long reaction times for making decisions between choices of equally high value (**Fig. 2b**).

We saw that for the simple case in which diffusion models implement the normative policy, this strategy does make sense, and in fact can even be optimal. However, this conclusion only applies to the simplest type of choices in which participants can be assumed to receive noisy value samples over time, making DDMs optimal. For complex decisions, as we have argued, noise may not be the bottleneck. Instead, retrieving the relevant information in memory and computing value are the computationally demanding steps, which we believe offer a new perspective on why people agonize over decisions involving high valued items. Indeed, retrieving the relevant facts to decide between a weekend in Shanghai or Kyoto takes time and is subject to uncertainty. Even after recalling facts about, say, museums and landmarks, we can't be sure that further contemplation doesn't bring up essential, decision-changing information. For example, we might recall cherry blossom season in Kyoto just after having booked a non-refundable flight to Shanghai. In other words, the decision options might have very different values once all relevant information is considered. Retrieving all the relevant information, however, takes time, because memory retrieval is a complex and stochastic process (Austerweil et al., 2012; Katkov et al., 2017; Romani et al., 2013). Likewise, inferring probability distribution over values can be time consuming, particularly if it is based on sampling strategies like MCMC. Then, being conservative and taking time to make good decisions might indeed be the best strategy, as it prevents us from committing seemingly good choices which could turn out to be a lot worse than the alternative once all the evidence has been recovered and assessed. Therefore, the reasons we are slow at deciding between two equally good options for complex decisions may not be so much due to the use of DDMs. Instead, they may be related to increased complexity of memory retrieval and value computation that require time and cannot be captured by simple models such as DDMs.

Discussion

Over the past few decades, neuroscientists and cognitive scientists have started to uncover the neural basis of decision making, guided by the development of precise quantitative behavioral and neural models (Beck et al., 2008; Gold & Shadlen, 2007; O'Connell et al., 2018). This remarkable line of work has been made possible by focusing on relatively simple forms of decision-making such as perceptual and simple value-based choices. As we have seen, for these types of decisions, the decision-related evidence determines the choices via a very simple process. For instance, in models of simple value-based decision making, the pieces of decision-related evidence are drawn from Gaussian distributions with means equal to the true

values. Then, the noise corrupting the evidence is the only factor limiting behavioral performance. Its deleterious impact can be mitigated by integrating increasingly more evidence, which averages out the noise. In fact, in simple models, averaging out noise is the main factor to the speed-accuracy trade off because decisions can be arbitrarily accurate, but only at the cost of spending an inordinate amount of time accumulating evidence, and thus reducing noise.

In contrast, noise is unlikely to be the main limiting factor for complex decision making. In fact, even for simple decision, whether neuronal noise is the main factor determining performance is far from clear (Beck et al., 2012; Drugowitsch et al., 2016). However, more importantly, other factors will necessarily contribute to the speed-accuracy trade-off, including in particular memory searches and value computation. As we discussed, both of these processes are major and unavoidable time sinks when making complex decisions, and are vastly more complex than simply averaging noise. Averaging noise is trivial: it simply requires adding samples over time (see Text box 1). In contrast, efficient database search or value inference in large Bayesian networks requires considerably more complex algorithms. This is indeed well known in the field of artificial intelligence, where computation and efficient search of large databases are central problems (Moore & Mertens, 2011; Murphy, 2012), while noise is a non-issue given that computers are effectively noise free. With regard to memory, the speed of searching through a database depends on its organization and its number of entries. However, regardless of the organization, in a fixed amount of time, only a subset of entries can be considered, leading to reduced precision that is prone to a speed accuracy trade-off. Similar principles must apply to neural memories. However, while multiple neural structures are thought to be involved in memory storage and memory retrieval, we still know very little about how the brain retrieve memories efficiently. It is quite likely that memory search not only relies on the decision's general context (Bornstein & Norman, 2017), but furthermore involves simulating complex models of the task at hand (e.g., simulating a day in a city to recover relevant facts about a city you intend to visit) but we are far from understanding how this is implemented in neural circuits. The rich literature of replay in the hippocampus (Foster & Wilson, 2006; Johnson & Redish, 2007; Pfeiffer & Foster, 2013) provides some intriguing cues, suggesting in particular that the brain may rely on internal simulations of the outside world to re-activate memories. However, much of the hippocampal literature is limited to the spatial domain or extremely simple abstract environments (Constantinescu et al., 2016) in which the use of internal simulations seems almost unwarranted. Hopefully, future studies will be able to consider how humans retrieve memories in much more complex environments in which memory retrieval becomes the computational bottleneck. It is difficult to predict which parts of the brain implement these computations but given the implication of the thalamus and cortex in storing long-term memory, it is likely that memory retrieval will involve more than just the hippocampal circuits. Recent studies suggest that, indeed, replays are observed in human cortex and hippocampus during decision-making and play a role in memory retrieval and consolidation (Eldar et al., 2018, 2020; Kurth-Nelson et al., 2016; Liu et al., 2019). This demonstrates the feasibility of studying the neural basis of memory retrieval in humans and highlight the similarities to research in rodents. Likewise, speed accuracy trade-offs are at play when computing values. If this process relies on simulations — whether deterministic (e.g. imagining future chess board positions) or stochastic (e.g. Monte-Carlo sampling-based methods) — the accuracy of the computation will typically scale with the time spent on the computation. Just like for memory search, the neural basis of value computation is still very much unknown. Part of the bottleneck in this research has been the paucity of detailed neuronal models that would help inspire future experiments. One promising avenue of future research is to combine such network models with symbolic systems (sometimes called systems 1 and 2; (Kahneman, 2013)) whose algorithms traditionally directly manipulate the abstract symbols that are believed to govern human conscious thoughts, irrespective of how those symbols are represented by our brain's neurons. Hybrid models rely on a mixture of

such abstract representations (e.g., a tree of possible sequences of future moves in chess) along with ones learned by neural networks. These models (e.g., Schrittwieser et al., 2020), but also offer a promising path towards understanding how our brains use past information to make efficient decisions.

Author contributions

The authors contributed equally to all aspects of the article.

Competing interests

The authors declare no competing interests.

References

- Austerweil, J., Abbott, J. T., & Griffiths, T. (2012). Human memory search as a random walk in a semantic network. *Advances in Neural Information Processing Systems*, 25. <https://papers.nips.cc/paper/2012/hash/14d9e8007c9b41f57891c48e07c23f57-Abstract.html>
- Beck, J. M., Ma, W. J., Kiani, R., Hanks, T., Churchland, A. K., Roitman, J., Shadlen, M. N., Latham, P. E., & Pouget, A. (2008). Probabilistic Population Codes for Bayesian Decision Making. *Neuron*, 60(6), 1142–1152. <https://doi.org/10.1016/j.neuron.2008.09.021>
- Beck, J. M., Ma, W. J., Pitkow, X., Latham, P. E., & Pouget, A. (2012). Not Noisy, Just Wrong: The Role of Suboptimal Inference in Behavioral Variability. *Neuron*, 74(1), 30–39. <https://doi.org/10.1016/j.neuron.2012.03.016>
- Benoit, R. G., Szpunar, K. K., & Schacter, D. L. (2014). Ventromedial prefrontal cortex supports affective future simulation by integrating distributed knowledge. *Proceedings of the National Academy of Sciences*, 111(46), 16550–16555. <https://doi.org/10.1073/pnas.1419274111>
- Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced-choice tasks. *Psychological Review*, 113(4), 700–765. <https://doi.org/10.1037/0033-295X.113.4.700>
- Bornstein, A. M., & Norman, K. A. (2017). Reinstated episodic context guides sampling-based decisions for reward. *Nature Neuroscience*, 20(7), 997–1003. <https://doi.org/10.1038/nn.4573>
- Chernoff, H. (1961). Sequential Tests for the Mean of a Normal Distribution. *Berkeley Symposium on Mathematical Statistics and Probability, 1961*, 79–91. <https://projecteuclid.org/ebooks/berkeley-symposium-on-mathematical-statistics-and-probability/Proceedings-of-the-Fourth-Berkeley-Symposium-on-Mathematical-Statistics-and-probability/Chapter/Sequential-Tests-for-the-Mean-of-a-Normal-Distribution/bsmsp/1200512160>
- Chittka, L., Skorupski, P., & Raine, N. E. (2009). Speed–accuracy tradeoffs in animal decision making. *Trends in Ecology & Evolution*, 24(7), 400–407. <https://doi.org/10.1016/j.tree.2009.02.010>
- Constantinescu, A. O., O'Reilly, J. X., & Behrens, T. E. J. (2016). Organizing conceptual knowledge in humans with a gridlike code. *Science*, 352(6292), 1464–1468. <https://doi.org/10.1126/science.aaf0941>
- Drugowitsch, J., Deangelis, G. C., Klier, E. M., Angelaki, D. E., & Pouget, A. (2014). Optimal multisensory decision-making in a reaction-time task. *eLife*, 2014(3), 1–19. <https://doi.org/10.7554/eLife.03005.001>
- Drugowitsch, J., Moreno-Bote, R., Churchland, A. K., Shadlen, M. N., & Pouget, A. (2012). The Cost of Accumulating Evidence in Perceptual Decision Making. *Journal of Neuroscience*, 32(11), 3612–3628. <https://doi.org/10.1523/JNEUROSCI.4010-11.2012>
- Drugowitsch, J., Moreno-Bote, R., & Pouget, A. (2014). Optimal decision-making with time-varying evidence reliability. *Advances in Neural Information Processing Systems*, 748–756. <https://doi.org/10.7554/eLife.03005>
- Drugowitsch, J., Wyart, V., Devauchelle, A.-D., & Koechlin, E. (2016). Computational Precision of Mental Inference as Critical Source of Human Choice Suboptimality. *Neuron*, 92(6), 1–14. <https://doi.org/10.1016/j.neuron.2016.11.005>
- Eldar, E., Bae, G. J., Kurth-Nelson, Z., Dayan, P., & Dolan, R. J. (2018). Magnetoencephalography decoding reveals structural differences within integrative decision processes. *Nature Human Behaviour*, 2(9), 670–681. <https://doi.org/10.1038/s41562-018-0423-3>
- Eldar, E., Lièvre, G., Dayan, P., & Dolan, R. J. (2020). The roles of online and offline replay in planning. *eLife*, 9, e56911. <https://doi.org/10.7554/eLife.56911>
- Ferreira, D. R. (2013). The Impact of the Search Depth on Chess Playing Strength. *ICGA Journal*, 36(2), 67–80. <https://doi.org/10.3233/ICG-2013-36202>
- Foster, D. J., & Wilson, M. A. (2006). Reverse replay of behavioural sequences in hippocampal place cells during the awake state. *Nature*, 440(7084), 680–683. <https://doi.org/10.1038/nature04587>
- Furl, N., & Averbeck, B. B. (2011). Parietal cortex and insula relate to evidence seeking relevant to reward-related decisions. *Journal of Neuroscience*, 31(48), 17572–17582. <https://doi.org/10.1523/JNEUROSCI.4236-11.2011>
- Gilks, W. R., Richardson, S., & Spiegelhalter, D. J. (Eds.). (1998). *Markov chain Monte Carlo in practice*. Chapman & Hall.
- Gold, J., & Shadlen, M. (2007). The neural basis of decision making. *Annu. Rev. Neurosci.*, 30, 535–574. <https://doi.org/10.1146/annurev.neuro.29.051605.113038>
- Griffiths, T. L., Callaway, F., Chang, M. B., Grant, E., Krueger, P. M., & Lieder, F. (2019). Doing more with less: Meta-reasoning and meta-learning in humans and machines. *Current Opinion in Behavioral Sciences*, 29, 24–30. <https://doi.org/10.1016/j.cobeha.2019.01.005>
- Hanks, T. D., & Summerfield, C. (2017). Perceptual Decision Making in Rodents, Monkeys, and Humans. *Neuron*, 93(1), 15–31. <https://doi.org/10.1016/j.neuron.2016.12.003>

- Huang, J., Rathod, V., Sun, C., Zhu, M., Korattikara, A., Fathi, A., Fischer, I., Wojna, Z., Song, Y., Guadarrama, S., & Murphy, K. (2017). Speed/Accuracy Trade-Offs for Modern Convolutional Object Detectors. *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 3296–3297. <https://doi.org/10.1109/CVPR.2017.351>
- Iigaya, K., Yi, S., Wahle, I. A., Tanwisuth, K., & O'Doherty, J. P. (2020). *Aesthetic preference for art emerges from a weighted integration over hierarchically structured visual features in the brain* [Preprint]. *Neuroscience*. <https://doi.org/10.1101/2020.02.09.940353>
- Jang, A. I., Sharma, R., & Drugowitsch, J. (2021). Optimal policy for attention-modulated decisions explains human fixation behavior. *eLife*, 10, e63436. <https://doi.org/10.7554/eLife.63436>
- Johnson, A., & Redish, A. D. (2007). Neural Ensembles in CA3 Transiently Encode Paths Forward of the Animal at a Decision Point. *Journal of Neuroscience*, 27(45), 12176–12189. <https://doi.org/10.1523/JNEUROSCI.3761-07.2007>
- Kahneman, D. (2013). *Thinking, fast and slow* (1st pbk. ed). Farrar, Straus and Giroux.
- Katkov, M., Romani, S., & Tsodyks, M. (2017). Memory Retrieval from First Principles. *Neuron*, 94(5), 1027–1032. <https://doi.org/10.1016/j.neuron.2017.03.048>
- Kemp, C., & Tenenbaum, J. B. (2008). The discovery of structural form. *Proceedings of the National Academy of Sciences*, 105(31), 10687–10692. <https://doi.org/10.1073/pnas.0802631105>
- Knudsen, E. B., & Wallis, J. D. (2021). Hippocampal neurons construct a map of an abstract value space. *Cell*, 184(18), 4640–4650.e10. <https://doi.org/10.1016/j.cell.2021.07.010>
- Krajbich, I., Armel, C., & Rangel, A. (2010). Visual fixations and the computation and comparison of value in simple choice. *Nature Neuroscience*, 13(10), 44. <https://doi.org/10.1038/nn.2635>
- Kurth-Nelson, Z., Economides, M., Dolan, R. J., & Dayan, P. (2016). Fast Sequences of Non-spatial State Representations in Humans. *Neuron*, 91(1), 194–204. <https://doi.org/10.1016/j.neuron.2016.05.028>
- Liu, Y., Dolan, R. J., Kurth-Nelson, Z., & Behrens, T. E. J. (2019). Human Replay Spontaneously Reorganizes Experience. *Cell*, 178(3), 640–652.e14. <https://doi.org/10.1016/j.cell.2019.06.012>
- Luce, R. D. (1991). *Response Times: Their Role in Inferring Elementary Mental Organization*. Oxford University Press, Incorporated. <https://public.ebookcentral.proquest.com/choice/publicfullrecord.aspx?p=3051923>
- Moore, C., & Mertens, S. (2011). *The nature of computation*. Oxford University Press.
- Murphy, K. P. (2012). *Machine Learning: A Probabilistic Perspective*. MIT Press.
- O'Connell, R. G., Shadlen, M. N., Wong-Lin, K., & Kelly, S. P. (2018). Bridging Neural and Computational Viewpoints on Perceptual Decision-Making. *Trends in Neurosciences*, 41(11), 838–852. <https://doi.org/10.1016/j.tins.2018.06.005>
- Orbán, G., Berkes, P., Fiser, J., & Lengyel, M. (2016). Neural Variability and Sampling-Based Probabilistic Representations in the Visual Cortex. *Neuron*, 92(2), 530–543. <https://doi.org/10.1016/j.neuron.2016.09.038>
- Osth, A. F., Jansson, A., Dennis, S., & Heathcote, A. (2018). Modeling the dynamics of recognition memory testing with an integrated model of retrieval and decision making. *Cognitive Psychology*, 104, 106–142. <https://doi.org/10.1016/j.cogpsych.2018.04.002>
- Palmer, J., Huk, A. C., & Shadlen, M. N. (2005). The effect of stimulus strength on the speed and accuracy of a perceptual decision. *Journal of Vision*, 5(5), 376–404. <https://doi.org/10.1167/5.5.1>
- Pfeiffer, B. E., & Foster, D. J. (2013). Hippocampal place-cell sequences depict future paths to remembered goals. *Nature*, 497(7447), 1–8. <https://doi.org/10.1038/nature12112>
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, 85(2), 59–108. <https://doi.org/10.1037/0033-295X.85.2.59>
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, 20(4), 873–922. <https://doi.org/10.1162/neco.2008.12.06.420>
- Romani, S., Pinkoviezky, I., Rubin, A., & Tsodyks, M. (2013). Scaling Laws of Associative Memory Retrieval. *Neural Computation*, 25(10), 2523–2544. https://doi.org/10.1162/NECO_a_00499
- Russell, S. J., Norvig, P., & Davis, E. (2010). *Artificial intelligence: A modern approach* (3rd ed). Prentice Hall.
- Schouten, J. F., & Bekker, J. A. M. (1967). Reaction time and accuracy. *Acta Psychologica*, 27, 143–153. [https://doi.org/10.1016/0001-6918\(67\)90054-6](https://doi.org/10.1016/0001-6918(67)90054-6)
- Schrittwieser, J., Antonoglou, I., Hubert, T., Simonyan, K., Sifre, L., Schmitt, S., Guez, A., Lockhart, E., Hassabis, D., Graepel, T., Lillicrap, T., & Silver, D. (2020). Mastering Atari, Go, chess and shogi by planning with a learned model. *Nature*, 588(7839), 604–609. <https://doi.org/10.1038/s41586-020-03051-4>
- Shadlen, M. N., Britten, K. H., Newsome, W. T., & Movshon, J. A. (1996). A computational analysis of the relationship between neuronal and behavioral responses to visual motion. *Journal of Neuroscience*, 16(4), 1486–1510. <https://doi.org/10.1523/JNEUROSCI.16-04-01486.1996>
- Shadlen, M. N., & Newsome, W. T. (1996). Motion perception: Seeing and deciding. *Proceedings of the National Academy of Sciences*, 93(2), 628–633. <https://doi.org/10.1073/pnas.93.2.628>
- Shadlen, M. N., & Newsome, W. T. (1998). The variable discharge of cortical neurons: Implications for connectivity, computation, and information coding. *The Journal of Neuroscience: The Official Journal of the Society for Neuroscience*, 18(10), 3870–3896. <https://doi.org/10.1523/JNEUROSCI.18-10-1998.3870>
- Shadlen, M. N., & Shohamy, D. (2016). Decision Making and Sequential Sampling from Memory. *Neuron*, 90(5), 927–939. <https://doi.org/10.1016/j.neuron.2016.04.036>
- Silver, D., Hubert, T., Schrittwieser, J., Antonoglou, I., Lai, M., Guez, A., Lanctot, M., Sifre, L., Kumaran, D., Graepel, T., Lillicrap, T., Simonyan, K., & Hassabis, D. (2018). A general reinforcement learning algorithm that masters chess, shogi, and Go through self-play. *Science*, 362(6419), 1140–1144. <https://doi.org/10.1126/science.aar6404>
- Tajima, S., Drugowitsch, J., Patel, N., & Pouget, A. (2019). Optimal policy for multi-alternative decisions. *Nature Neuroscience*, 22(9), 1503–1511. <https://doi.org/10.1038/s41593-019-0453-9>

- Tajima, S., Drugowitsch, J., & Pouget, A. (2016). Optimal policy for value-based decision-making. *Nature Communications*, 7, 12400. <https://doi.org/10.1038/ncomms12400>
- Vul, E., Goodman, N., Griffiths, T. L., & Tenenbaum, J. B. (2014). One and done? Optimal decisions from very few samples. *Cognitive Science*, 38(4), 599–637. <https://doi.org/10.1111/cogs.12101>
- Wald, A. (1945). Sequential Tests of Statistical Hypotheses. *The Annals of Mathematical Statistics*, 16(2), 117–186. <https://doi.org/10.1214/aoms/1177731118>
- Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica*, 41(1), 67–85. [https://doi.org/10.1016/0001-6918\(77\)90012-9](https://doi.org/10.1016/0001-6918(77)90012-9)
- Yang, S. C.-H., Lengyel, M., & Wolpert, D. M. (2016). Active sensing in the categorization of visual patterns. *eLife*, 5(FEBRUARY2016), 1–22. <https://doi.org/10.7554/eLife.12215>