An Introduction to Psychologically Plausible Sampling Schemes for Approximating Bayesian Inference

Jian-Qiao Zhu[1], Nick Chater[2], Pablo León-Villagrá[1], Jake Spicer[1], Joakim Sundh[1], and Adam Sanborn[1]

[1]Department of Psychology, University of Warwick, UK, CV4 7AL

[2]Warwick Business School, University of Warwick, UK, CV4 7AL

Author Note

Full name	Email Address	Small bio
Jian-Qiao Zhu	jianqiao.zhu@warwick.ac.uk	Zhu is a Research Fellow of Psychology at University of Warwick. He is interested in the rational principles of cognition.
Nick Chater	nick.chater@wbs.ac.uk	Nick is a Professor of Behavioural Science at Warwick Business School. He is interested in the cognitive and social foundations of rationality.
Pablo León-Villagrá	pablo.leon.villagra@gmail.com	Pablo is a Research Fellow of Psychology at University of Warwick. His interests concern how learning and development shape human cognition.
Jake Spicer	jake.spicer@warwick.ac.uk	Jake is a Research Fellow of Psychology at University of Warwick. His research looks at sampling approaches to human cognition.
Joakim Sundh	joakim.sundh@psyk.uu.se	Joakim is a Research Fellow of Psychology at University of Warwick. His research concerns exploring sampling-based cognitive models of Bayesian inference.
Adam Sanborn	A.N.Sanborn@warwick.ac.uk	Adam is a Professor of Psychology. He is interested in the rationality of human behaviour and studies with Bayesian models of cognition.

Keywords: sampling algorithms, Bayesian models of cognition, approximation, resource rationality, autocorrelation

Subject index: Cognitive Science, Behavioural Science, Behavioural Economics, Computational Modeling, Brain Science

Abstract

The brain must make inferences about, and decisions concerning, a highly complex and unpredictable world, based on sparse evidence. An "ideal" normative approach to such challenges is often modeled in terms of Bayesian probabilistic inference. But for real-world problems of perception, motor control, categorization, language understanding or commonsense reasoning, exact probabilistic calculations are computationally intractable. Instead, we suggest that the brain solves these hard probability problems approximately, by considering one, or a few, samples from the relevant distributions. Here we provide a gentle introduction to the various sampling algorithms that have been considered as the approximation used by the brain. We broadly summarise these algorithms according to their level of knowledge and their assumptions regarding the target distribution, noting their strengths and weaknesses, their previous applications to behavioural phenomena, as well as their psychological plausibility.

An Introduction to Psychologically Plausible Sampling Schemes for Approximating Bayesian Inference

Both natural and built environments are complex, and people often have to operate under great uncertainty about the true state of the world. One major cause of this uncertainty, as discussed in several of the other chapters in this book, is that we often have access to only a small number of samples of information from the environment. These investigations of information sampling often study situations in which this information is either biased or unreliable. Though even when it is both unbiased and reliable, the number of relevant experiences we have available in most tasks is far fewer than the 25,500 that Jacob Bernoulli estimated were required to have "moral certainty" about the probability of even a simple binary event (Stigler, 1986).

Fortunately, Bayesian models of cognition provide a principled way for the mind to deal with this uncertainty: stating all hypotheses h, defining a prior probability for these hypotheses $p(h)$, and then updating these beliefs according to the rules of probability theory as information about the environment d becomes available. These updated beliefs are the posterior probabilities $p(h|d)$, and are proportional to the prior multiplied by the likelihood that the data are produced from each hypothesis $p(d|h)$. Interestingly in many complex domains, from language production and common-sense reasoning to vision and motor control and intuitive physics, human behaviour corresponds well to this probability calculus (Chater & Manning, 2006; Battaglia, Hamrick, & Tenenbaum, 2013; Sanborn, Mansinghka, & Griffiths, 2013; Wolpert, 2007; Sanborn & Chater, 2016). If the assumptions of the Bayesian model are correct, there are straightforward ways to incorporate the costs and benefits of each action to decide the best action to take under uncertainty.

To enjoy many of the benefits of the Bayesian approach, the brain, and, in fact, any biological and physical machinery, has to calculate the posterior in realistic amounts of time.

Unfortunately, computing exact posterior degrees of belief has been proven computationally difficult and is typically not possible in most real-life applications (Roth, 1996; van Rooij, 2008). To give an intuition as to why, we consider the following running example: suppose that I plan to meet up with a friend in London for a bite to eat at a place of their choosing, though they have not yet told me where. I call my friend as I jump into the cab, but they do not answer. Where should I tell the driver to go? I don't have the time (or money) to ride around London visiting possible locations, so I must make a single decision based on my internal knowledge. Ideally, in order to minimise how far I need to walk once my friend finally answers their phone, I would like to be dropped off at the average location (as opposed to the most likely location) that I expect my friend to be[1]. To calculate this average location exactly, I would have to assign a probability to each establishment that serves food in London, based on my knowledge of my friend. Then, I would multiply the longitude of every establishment by the probability that my friend is at that establishment, sum up the results, and then repeat the calculation for latitude. However, because London has more than 44,000 establishments that serve food, calculating average locations in this way is implausibly difficult.

There is no doubt that a compromise must be made: for example, one might settle for an approximation of Bayesian inference rather than an exact calculation. One class of algorithms for approximating the posterior, developed in computer science and statistics, is based on Monte Carlo methods. The idea is simple: instead of using the entire posterior distribution, $p(h|d)$, generate a sequence of samples from the posterior distribution, $h^i \sim p(h|d)$, where h^i is the i-th sample of the hypotheses, and then use them to guide future behaviour. These samples are then internal mental samples, rather than samples of

[1] For simplicity, we assume that I want to minimize only the squared walking distance

information from the environment. While these samples provide a less accurate characterization of the posterior distribution than the probabilities do, they help with the complexity of calculating with that distribution. For example, to calculate the location I should travel to meet my friend, if I'm given a set of sampled locations where my friend might be, then I can simply take the average latitude and longitude of the samples. The accuracy of this estimate increases with the number of samples: while in the limit of an infinite number of samples the accuracy of my calculation would be perfect, an imperfect but still useful answer can still be found using a psychologically realistic sample size, one that is far smaller than the total number of eating establishments in London (Vul et al., 2014). In this chapter, we review a psychological literature that pursued this line of research and examine the potential link between sampling algorithms and cognitive psychology (summarized in Table 1 and Figure 1), with a more in-depth discussion of how sampling algorithms can explain human behavior given in Chapter X.

Table 1.

Sampling algorithms and their statistical and psychological implications.

Knowledge Required of Probability Distribution	Algorithms	Deviations From Ideal Inference	Example applications
Global knowledge	Direct sampling	Stochastic behavior	Probability matching behaviours (Vul et al., 2014); Exploration-exploitation tradeoff (Speekenbrink & Konstantinidis, 2015; Gershman, 2018)

Approximate global knowledge	Importance sampling; Particle filters	Stochastic behavior; Overweighting extreme events; Order effects in updating	Overweighting of extreme events and the four-fold pattern of risky preference (Lieder et al., 2018; Nobandegani et al., 2018); Reproductions from memory of perceptual stimuli, and predictions about the duration of real-life events (Shi et al., 2010); Serial dependence and working memory capacity in category learning (Sanborn et al., 2010; Lloyd et al., 2019); Order effects in human causal learning (Abbott & Griffiths, 2011); Classical conditioning in animal behaviours (Daw & Courville, 2008; Gershman et al., 2010); Decision making in changing environments (Yi et al., 2009; Brown & Steyvers, 2009).
Local knowledge	MCMC algorithms: Random walk Metropolis; Gibbs sampling; Metropolis-coupled	Stochastic behavior; Framing effects; Autocorrelated behavior	Bistable perception (Gershman et al., 2012); Anchoring bias (Lieder et al., 2018); Biases in probability judgments (Sanborn & Chater, 2016; Dasgupta et

	Markov chain Monte Carlo	al., 2017); Autocorrelation in human causal learning (Bramley et al., 2017); *1/f* noise and Lévy flight in repeated estimation and memory retrieval (Zhu et al., 2018);

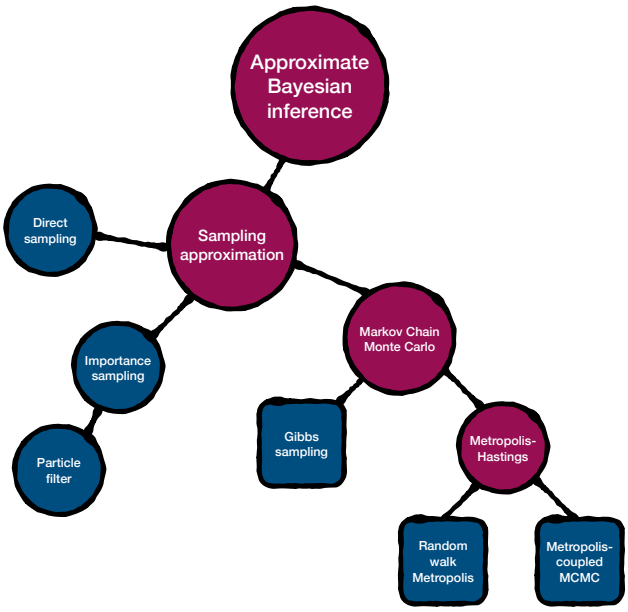

Figure 1. A 'family tree' of sampling-based approximate inference, where higher parent nodes (in red) represent more generalized concepts and lower leaf nodes (in dark blue) denote specific sampling algorithms. The algorithms requiring global or approximate global knowledge are circles, whereas those requiring local knowledge are squares.

Sampling with Global Knowledge

Direct Sampling

The simplest idea in sampling is to draw a random sample from the probability distribution of interest, termed the *target distribution*, or in our example, the distribution of where my friend might be, directly: $h^i \sim p(h|d)$. Each of the samples are independent of one another and the sample average is also unbiased. This requires *global knowledge* of the distribution—the knowledge of all states and their probabilities. For the meeting-a-friend example, by analogy, direct sampling simply means drawing samples of many possible locations and taking the empirical average of these samples to approximate my friend's true average location. We do not need to integrate over all possible places, but can instead rely on a few samples drawn from the distribution of the possible meetup places to get a good estimate of where to travel.

Psychological Applications of Direct Sampling.

Given a specific input, Bayesian models of cognition are deterministic: each input typically leads to a single response for that input. In two-alternative forced choice where the reward history of both options has been observed, a reward-maximising agent should always choose the option with the higher chance of reward. This contrasts with the observation that human behavior is almost inevitably noisy, even in tasks in which the stimuli are clear and so there is unlikely to be any sensory noise (Mosteller & Nogee, 1951; Spicer, Mullett, & Sanborn, submitted). Instead, an extensive empirical literature shows that people "probability match" and choose an option with a frequency proportional to the probability of reward (Vulkan, 2000).

Adding the assumption of direct sampling makes Bayesian models of cognition stochastic, and so offers an explanation for the noise in human behavior without losing the benefits of the normative framework of Bayesian reasoning. When decisions are carried out based on one or a few samples, stochastic behaviour is expected to occur because samples are

randomly generated. The matching law fits nicely with the sampling view of decision making, especially when making decisions based on a single sample (Vul et al., 2014). Further assuming people draw one sample from past trials and act optimally towards the sample, the predicted choice pattern of one option should match the probability of reward of the option: direct sampling adds noise but does not add any bias (at least beyond any bias in the distribution it draws samples from).

Indeed, direct sampling is also widely used in non-Bayesian psychological models; most stochastic models of human behaviour tend to use direct sampling by default. For example, the drift diffusion model of choice and response time supposes that decision making is a process of evidence accumulation until a threshold is reached. Each piece of evidence is typically assumed to be independent of the last and is directly sampled from memory or the environment (Ratcliff & Rouder, 1998; Usher & McClelland, 2001; Pleskac & Busemeyer, 2010; Nosofsky, 2011; Blurton et al., 2020; Shadlen & Shohamy, 2016). Global memory matching models assume that recalled items are directly sampled from a distribution over items in memory (Raaijmakers & Shiffrin, 1981; Brown, Neath, & Chater, 2007), while influential models of categorization and many other tasks assume that responses are directly sampled from a distribution of responses (Nosofsky, 1984).

Sampling with Approximate Global Knowledge

There remain, however, some psychologically implausible aspects of direct sampling from Bayesian models. Aside from the issue of generating samples that are independent and unbiased draws from a distribution (which people are notoriously poor at anyway; Wagenaar, 1972), there is the issue of obtaining and representing global knowledge: knowledge of all of the states and their probabilities. Without these probabilities, direct sampling is not possible to implement.

Returning to our example, let's say that while trying to decide what destination to tell my cab driver in London, I receive a somewhat ambiguous text message from my friend saying that they are hungry for a curry. This data (i.e., new piece of information from my friend), d, allows me to update my probabilities as to where I should tell the cab driver to go. But there is a tractability issue that arises from Bayes rule itself. When updating a probability distribution with new information, the posterior probability of each hypothesis is equal to its likelihood, $p(d|h)$, multiplied by its prior, $p(h)$, divided by a proportionality constant or *partition function* Z that also needs to be calculated:

$$p(h|d) = \frac{p(d|h)p(h)}{Z} = \frac{p^*(h|d)}{Z}$$

This partition function Z is the sum of each and every hypothesis multiplied by its likelihood:

$$Z = \sum_h p(h|d)p(h)$$

This means that while it is relatively easy to calculate a value $p^*(h|d) = p(d|h)p(h)$ that is proportional to the posterior probability of a single hypothesis, it is much more computationally demanding to determine the constant Z needed to find the exact value of the posterior probability. And, because the average longitude and average latitude require multiplying each longitude and latitude by its probability (whether or not sampling is used), knowing only $p^*(h|d)$ instead of $p(h|d)$ is not enough – it means that my answer will be off by Z, which is an unknown constant[2,3]. If, as before, I wish to avoid summing across the more

[2] For discrete hypotheses, Z lies between zero and one, while for continuous hypotheses Z can take any value greater than zero.

[3] Another intuition about the importance of the partition function is to think about calculating the chances of winning a neighborhood lottery in which the ticket sellers have gone door to door. You will know how many tickets you yourself have purchased, but your chances of winning will depend on an unknown constant: the total number of tickets sold.

than 44,000 eating establishments in London, a new approximation approach is much needed[4].

One way to do so is to utilize *approximate global knowledge* of the posterior probability distribution: starting with a rough idea of the posterior that is refined by calculating values $p^*(h|d)$ proportional to the posterior distribution for only a sample of the hypotheses, rather than for every hypothesis as calculating Z requires. However, unlike direct sampling, for small sample sizes this method can introduce biases which depend on the particulars of the sampling algorithm. We next introduce a key example of a method which draws samples with such approximate global knowledge of the posterior, importance sampling.

Importance Sampling

If directly sampling from $p(h|d)$ is computationally daunting except for the simplest toy examples, one may consider an alternative sampling strategy that first draws samples from conventional and simple distributions (e.g., a Gaussian distribution) and later adjusts these samples to align with the target distribution. Indeed, this is the key idea behind the importance sampling algorithm. This method draws samples from a simpler distribution $q(h)$ (also known as the proposal distribution), and then *reweights* them in reference to the target posterior distribution $p(h|d)$, correcting for the difference between the two distributions. As we shall soon see, the reweighting scheme (i.e., the correction for the difference) can be done

[4] Another intuition to understand why it could be hopeless to do direct sampling for complex problems can be drawn from the history and evolution of scientific theories. In science, deciding how probable a theory is when new data arrives requires integration over all possible theories. It is, however, an almost impossible job to imagine all possible theories. Furthermore, if we chose theories at random, the likelihood of the theory predicting the new data would almost always be zero. Instead, scientists make local adjustments to theories with only occasionally breakthroughs (most dramatically in Kuhn [1962] called paradigm shifts)). On explaining how light bends around heavy objects, for example, it is impossible to find a high likelihood account before Einstein's theory of relativity or by integrating over all possible Newtonian theories of physics. Normalization is hard because it requires a comprehensive survey of all possible theories.

with unnormalized probabilities. Therefore, the partition function, which requires summation over the all possible hypotheses, is no longer needed for importance sampling. In general, to make samples from q more representative of p, samples from more probable states in p and less probable states in q will receive more weight.

Following on the same meeting-a-friend example as above, importance sampling does not require the global knowledge of all possible meeting places and their associated probabilities used by direct sampling. Instead, the extent of knowledge needed in importance sampling can be understood by analogy to rough knowledge of which areas in London have the highest chance of my friend being there (e.g., their home or workplace) and how quickly the chance drops off from those streets. I then could sample establishments guided by this rough knowledge and correct for the estimated probabilities of meetup based on the probability of that place. Formally, the correction step that reweights samples is defined as follows:

$$w^i = \frac{p(h^i|d)/q(h^i)}{\sum_{j=1}^{N} p(h^j|d)/q(h^j)} \text{ with } \sum_{i=1}^{N} w^i = 1$$

where N is total number of samples. The weights of samples can be seen as a measure of how well the proposal distribution q (i.e., the rough knowledge) fits the target distribution p. Consider an extreme scenario when q matches p exactly (i.e., $q = p$), the samples will be equally weighted (i.e., $w^i = \frac{1}{N}, i = 1,2, \dots, N$) because there is no need of reweighting to correct the differences between two distributions. To this end, the method may appear to be redundant: why bother to draw samples from another distribution if we can readily evaluate $p(h|d)$? The usefulness of the importance sampling algorithm is realised since the equation above can be further simplified as follows:

$$w^i = \frac{\frac{p^*(h^i|d)}{Zq(h^i)}}{\sum_{j=1}^{N} \frac{p^*(h^j|d)}{Zq(h^j)}} = \frac{p^*(h^i|d)/q(h^i)}{\sum_{j=1}^{N} p^*(h^j|d)/q(h^j)} \text{ where } h^i \sim q(h)$$

As the equation above suggests, the normalisation constant of $p(h|d)$, Z, can be ignored in calculating weights, meaning that one can use the un-normalised values $p^*(h|d)$, significantly reducing the amount of computation required to run the algorithm. One common choice for the proposal distribution is to use the prior distribution: $q(h) = p(h)$. That is, samples are initially drawn from the prior distribution, and then reweighted to act as samples from the posterior distribution. Note that the importance sampler needs to know the unnormalized posterior at the sampled hypotheses, but not the normalized posterior which requires a global knowledge of the entire hypothesis space. The weights for this scheme, called *likelihood-weighted importance sampling*, are particularly simple because the prior probabilities cancel

$$w^i = \frac{p^*(h^i|d)/p(h^i)}{\sum_{j=1}^{N} p^*(h^j|d)/p(h^j)} = \frac{p(d|h^i)}{\sum_{j=1}^{N} p(d|h^j)} \text{ where } h^i \sim p(h)$$

so the weights are just samples proposed from the prior weighted by their likelihoods. As with any proposal distribution, this scheme works well when the proposal distribution (i.e., now the prior) is similar to the target distribution (i.e., the posterior), but poorly when they are very different.

While a speed-up in computation can be achieved through using the unnormalised distribution (i.e., $p^*(h|d)$), the proposal distribution still requires knowledge at least the space of states in p that are non-zero, and covers the same space in q with non-zero probabilities as well. This is simply because the importance sampling algorithm will never propose states from the spaces where the proposal distribution, q, has zero probabilities. In this sense, the importance sampling algorithm requires approximate global knowledge of the target distribution: a rough knowledge of where the probable states in the target distribution

might be located. Without this knowledge the algorithm will be very inefficient. If q is too broad compared to p (e.g., q is a uniform distribution across all possible meetup locations, meaning that I have no prior knowledge of believing one restaurant is more probable for a meetup than another), many of the proposed states will be low probability states in the target distribution and a very large number of samples will be needed to produce a good approximation. And alternatively, when q is too narrow compared to p, high probability states in p are very unlikely to be proposed based on q, and so again a very large number of samples will be needed to produce a good approximation.

In practice, the importance sampling algorithm has been widely used to approximate averages of some function $f(h)$ with state probabilities distributed according to $p(h|d)$. For example, when the function $f(h)$ is a utility function, the algorithm can be used to approximate expected utility. In our example of meeting a friend for a bite to eat, $f(h)$, is either the longitude or latitude of the eating establishment h (i.e., a hypothesis of where my friend might be). The sample-based approximation to function average is simply the weighted average of the function values at sampled hypotheses:

$$\mathbb{E}_p[f] \approx \frac{1}{N}\sum_{i=1}^{N} w^i f(h^i)$$

And in fact, this functional form provides a generalization of the well-known exemplar model of categorization (Nosofsky, 1986). In an exemplar model, exemplars of each category are remembered and the category label of a new exemplar is inferred from their similarity of the new exemplar to each of the old exemplars (Nosofsky, 1986). If the remembered exemplars are considered to be samples from a prior distribution (where the sampling is done by the environment rather than internally), and the similarity of each new exemplar to the old exemplars is encoded by the likelihood function, then the exemplar model can be written exactly as a likelihood-weighted importance sampler, providing a mathematical link between

a sampling scheme and an empirically supported cognitive model (Shi, Griffiths, Feldman, & Sanborn, 2010).

While the link of the importance sampler to the exemplar model is an unbiased use of samples, in other situations unbiased samples are not the most useful samples. When function averages are approximated by importance sampling with a limited amount of samples, certain states should matter more in their impact on the average and thus should be prioritised. Intuitively, those states that are highly probable in p and are more extremely valued in the function f should influence the average stronger than the other states. In fact, the optimal proposal distribution for calculating function average has been proved to capture these intuitions (Geweke, 1989; Murphy, 2012):

$$q^{\text{optimal}}(h) = \frac{|f(h)|p(h|d)}{\sum_{h'}|f(h')|p(h'|d)}$$

Paradoxically, in order to optimally use the importance sampling to approximate function average, the optimal proposal distribution requires global knowledge of both f and p.

Psychological Applications of the Importance Sampler

The optimality of preferentially sampling more extreme states in importance sampling has inspired a rational reinterpretation of people's overestimation in judging the frequency of extreme events (Lieder, Griffiths, & Hsu, 2018; Nobandegani et al., 2018). The overrepresentation of extreme events is often measured by observing people's risky choices in experiments in which the extremities and probabilities of outcomes were systematically manipulated (Tversky & Kahneman, 1973; Lieder et al., 2018). While the rational model of risky decision making, expected utility theory (von Neumann & Morgenstern, 2007), prescribes that people choose the option with the highest expected utility (with some axiomatic assumptions on the utility function), the computation of expected utility as a weighted average of the utility function is generally intractable.

This, however, can be approximated by importance sampling using a utility-weighted sampler which prioritizes eventualities according to their extremity and frequency: $|u(h)|p(h|d)$ (Lieder et al., 2018; Nobandegani et al., 2018). If people are assumed to calculate expected utility with importance sampling, the optimal proposal distribution to do so is determined by the extremity of the utility, meaning people should oversample (relative to their probabilities) hypotheses with large positive or negative utilities: in essence, it is important to consider rare events if their consequences could be life-changing (in either a positive or negative sense). Such overrepresentation can then persist in judgements even when samples are reweighted as described above. This model shows how a rational agent, who is bounded by computational resources and time in calculating exact expected utility (i.e., only small sample sizes are affordable), should overestimate extreme eventualities because they are drawn from an importance distribution designed to efficiently estimate expected utility.

Particle Filters

The importance sampling algorithm is designed for situations in which there are no new observations about the environment (i.e., sampling from a fixed target distribution), but in many realistic scenarios, especially those highlighted in other chapters in this book, it is important to deal with incoming information (i.e., a constantly evolving target distribution). To draw an analogy with the example of meeting my friend for a bite to eat, we can now imagine that my friend is sending me a series of text messages: first telling me that they want a curry, then telling me that they don't want to spend too much money, and then telling me that they don't want to travel too far. Each of these texts conveys additional information that I will want to use to update the best location to tell my cab driver, but the other algorithms that

we review in this chapter are ill-suited for inference of this type because they need to start sampling from scratch whenever the posterior distribution changes.

Luckily, importance sampling can be generalized to sample from changing posterior distributions by performing importance sampling sequentially. At its simplest, this involves updating the weights of each sampled hypothesis by multiplying each weight by the likelihood of that hypothesis producing the new incoming data, iteratively applying likelihood-weighted importance sampling. Sequential importance sampling is part of a family of algorithms known as *particle filters,* which often for efficiency reasons involve resampling the set of hypotheses from the weight distribution so that each resampled hypothesis has equal weight (Doucet et al., 2001; Arulampalam et al., 2002; Speekenbrink, 2016). In cognitive terms, the particle filter acts like an evolving set of hypotheses in the mind: the hypotheses that better explain the observed data have a higher chance of surviving and multiplying while hypotheses that poorly explain observed data are likely to go extinct.

Psychological Applications of Particle Filters

Particle filters have been used to explain primacy effects found in human cognition, in which early data have an outsized influence. These primacy effects are often difficult to explain with Bayesian models because the temporal structure of data typically should not matter for the tasks people are tested on, but particle filters can produce them as a result of the evolving set of hypothesis samples concentrating on initially-promising hypotheses, and so failing to have any samples of hypotheses that should later dominate. For example, particle filters have been used to explain garden-path sentences such as *'The women brought the sandwich from the kitchen tripped'* which first may lead the listener into first thinking that the sandwiches were brought by the women, until the word *'tripped'* makes it clear that the sandwiches were brought to the women (Levy, Reali, & Griffiths, 2009). Primacy effects have also explained serial dependence and working memory capacity in category learning

(Sanborn et al., 2010; Lloyd et al., 2019), order effects in human causal learning (Abbott & Griffiths, 2011), classical conditioning in animal behaviours (Daw & Courville, 2008; Gershman et al., 2010), and decision making in changing environments (Prat-Carrabin et al., 2021; Yi et al., 2009; Brown & Steyvers, 2009).

Sampling with Local Knowledge

While approximate global knowledge significantly reduces computational costs, is it possible to sample with a lesser form of knowledge? What if I don't know where my friend is likely to be, even approximately? I might only be able to easily calculate the unnormalized probabilities $p^*(h|d)$, and must somehow sample from the posterior distribution without access to the partition function Z. Here we introduce a general family of sampling algorithms that operate on this principle, Markov chain Monte Carlo (MCMC). The core idea in MCMC is to simulate transitions of a Markov chain whose stationary distribution is proportional to $p(h|d)$. That is, we can sample from a conditional probability $q(h'|h)$ (i.e., the transition probability of a Markov chain)[5] recursively long enough, such that the stationary distribution of the Markov chain is in proportion to $p(h|d)$. In our meetup example, this is similar to comparing the likelihood of two possible restaurants: we do not know the exact probability that my friend selected either location, but we do know which has the better chance, and, with enough pairwise comparisons, can create an approximation of the underlying distribution given certain assumptions (Tierney, 1994).

While a Markov chain converges to a unique stationary distribution, there are many ways to design the Markov chain's transitions, $q(h'|h)$, to do so while using only *local*

[5] Following on our example of meeting a friend, the transition probability of a Markov chain can be thought as the probability of thinking about next restaurant based on the current restaurant. It is also possible that the next restaurant turns out to be the same as the current one.

knowledge of the target distribution. We define local knowledge of the distribution to be the ability to determine the unnormalized probability of any state $p^*(h|d)$, but there is no requirement to store all of the $p^*(h|d)$ or even to store a list of all possible states h. It is only necessary to consider a very small number of states and unnormalized probabilities at any one time, and these can be forgotten following evaluation, which greatly increases the psychological plausibility of these sampling algorithms. Different $q(h'|h)$ suit different target distribution geometries. As a result, the sampler dynamics vary among algorithms; so do their psychological implications. In this section, we will focus on discussing the Random Walk Metropolis (RWM) that has achieved empirical successes in explaining aspects of human data, while also briefly mentioning related algorithms including Gibbs sampling and Metropolis-coupled MCMC (MC^3).

Random Walk Metropolis

A commonly used proposal distribution of MCMC is a Gaussian distribution centered on the current state: $q(h'|h) = \mathcal{N}(h'; h, \sigma^2)$. That is, new potential states are likely to fall close to the previous state. However, not every proposed state h' will be automatically accepted; if rejected, the new state should remain h. There exist numerous rules for accepting or rejecting proposed states, whose ultimate objectives are to approximate a probability distribution $p(h|d)$ by staying at a state for an amount of times that is also proportional to its probability (Dellaportas & Roberts, 2003). The combination of a Gaussian proposal distribution and the Metropolis-Hastings acceptance rules is also known as a Random Walk Metropolis (RWM) algorithm (Metropolis et al., 1953).

To solve the meeting-a-friend example as a RWM sampler, I should plan my direction of travel by first mentally comparing restaurants in London in a sequential fashion. I start with one restaurant and then think about the unnormalized probability that my friend picked

it. Then I think of a nearby restaurant and the unnormalized probability that my friend picked it instead. If the new restaurant has higher chance of being picked by my friend than the last, then I focus on the new restaurant. Conversely, if the new restaurant has lower chance than the last, then I stochastically decide whether to focus on this restaurant or focus on the previous one. Whichever restaurant I focus on, I again then randomly think of one of its neighbours next. Surprisingly, the proportion of times I focus on a restaurant gives an estimate of that establishment's probability of meetup.

How does the Metropolis-Hastings acceptance rule guarantee that the sampler spends time in each state in proportion to its probabilities in the target distribution $p(h|d)$? Because we want a proposal distribution that does not relate to the target distribution (i.e., does not appeal to the global knowledge), the sampler cannot *a priori* know how probable the proposed state h' is in the target distribution. The sampler does know, however, that the proposed state h' is either more probable or less probable than the current state h; that is, whether my friend is more likely to be at the new establishment. If the sampler greedily moves only to more probable states, it then should mimic optimisation algorithms that search for local optima. Our goal is not to find the maximum probability, however, but instead to draw representative samples from the target distribution; hence unlike most optimisation algorithms, the sampler has to move to less probable states occasionally. The quantity that governs the stochastic movement between proposed and current states is the relative probability between the two states $\frac{p(h'|d)}{p(h|d)}$. The formal condition underlying the convergence of RWM is known as the detailed balance (O'Hagan & Forster, 2004). In practice, the probability of accepting the proposed state follows:

$$p^{\text{accept}} = \min\left(1, \frac{p(h'|d)q(h|h')}{p(h|d)q(h'|h)}\right)$$

In RWM, a symmetric Gaussian distribution is used as the proposal distribution, meaning that $q(h|h') = q(h'|h)$. Furthermore, the calculation of relative probability between h and h' can be further simplified using unnormalised distribution because the normalisation constant cancels out:

$$\frac{p(h'|d)}{p(h|d)} = \frac{p^*(h'|d)/Z}{p^*(h|d)/Z} = \frac{p^*(h'|d)}{p^*(h|d)}$$

Metropolis-Hastings algorithms are useful in practice because we now can sample from $p(h|d)$ even if the normalisation constant Z is unknown: we do not need to know the probability of every eatery in London, but only the relative chance of one location over another. For the algorithms running symmetric proposal distribution such as the random walk Metropolis, the acceptance probability can be further shortened as follows:

$$p^{\text{accept}} = \min\left(1, \frac{p^*(h'|d)}{p^*(h|d)}\right)$$

Psychological Applications of the random walk Metropolis

RWM typically generates autocorrelated samples even when the target distribution remains unchanged. This is because the proposal distribution often suggests new states within a local neighborhood of the current state (i.e., local jumps). As a result, RWM is sensitive to its starting point, which allows it to explain framing effects. For example, RWM has been used to explain the anchoring bias where people's estimates are biased toward a previously-presented irrelevant quantity: though irrelevant to the subsequent task, people may use the irrelevant quantity as the starting point for RWM. If only a few samples are generated, there will only be a few local jumps from the starting point; hence estimates are biased toward the starting point (Lieder et al., 2018).

The unpacking effect also can be explained by a biased starting point. Depending on the examples that are unpacked, people judge the probability of an implicitly unpacked event (e.g., "*death from heart attack, cancer, and other natural courses*") differently than the

probability of an equivalent packed one (e.g., "*death from natural causes*") despite probability theory requiring the two to be the same (Sloman et al., 2014; Fox & Tversky, 1998; Dasgupta et al., 2017). Empirically, unpacking to typical examples increases probability judgments (i.e., subadditivity), while unpacking to atypical examples decreases probability judgments (i.e., superadditivity). In RWM, due to local search and autcorrelation, starting in high probability region will likely miss out atypical examples, while starting in low probability region will likely overestimate atypical examples, therefore explaining these unpacking effects as a starting point bias (Dasgupta et al., 2017). In addition to these two examples, the starting point has been used to explain a range of other biases, including the observed autocorrelation in human hypothesis generation (Sanborn & Chater, 2016; Dasgupta et al., 2017; Bonawitz et al., 2014; Gershman et al., 2012; Vul & Pashler, 2008), which we discuss further in Chapter X.

Gibbs Sampling

When the probability distribution of one variable conditioned on the values of all the other variables is relatively simple to express, it can be more efficient to use the conditional probability as the proposal distribution. Gibbs sampling implements this very idea that the transition probability $q(h'|h)$ is defined as the probability of one variable given fixed values of all of the other variables (Geman & Geman, 1984). The conditional probability is often easy to calculate in probabilistic graphic models where conditional dependencies are explicitly expressed in a graph (Koller & Friedman, 2009).

If we consider the meeting-a-friend example again, instead of thinking of new eating establishments based on their straight-line distance from my current focus, I instead think of possible meetup places along the longitudes and latitudes of London iteratively. Making the unrealistic assumption that London is laid out on a grid this would be like sampling among all

the eating establishments on a single (either north-south or east-west) road. For instance, I fix a latitude first, and then randomly sample an eatery along it (i.e., a longitude). Then, fixing that sampled longitude, I randomly sample an eatery along it (i.e., a new latitude). By iteratively swapping back and forth, I can gradually build up my estimate of the probability of each establishment being the actual meetup place. In this example, one iteration of the Gibbs sampling could be as follows:

$$h_x^i \sim p(h_x | h_y^{i-1}, d)$$

$$h_y^i \sim p(h_y | h_x^i, d)$$

where h_x and h_y are longitude and latitude of the restaurant h. It is straightforward to see that we are effectively sampling from the joint distribution, $p(h|d) = p(h_x|h_y, d)p(h_y|d)$, when the latitude that have been conditioned on (i.e., h_y) are sufficiently explored[6].

Psychological Applications of Gibbs Sampling

Causal relationships are often represented as a set of boxes and arrows with the boxes representing variables and the arrows indicating how variables causally influence one another. Gibbs sampling is often used for performing approximate inference because it is relatively easy to sample the value of a variable given fixed values of all of the other variables. Interestingly, the dynamics of Gibbs sampling match the sequential belief changes of human participants in online causal learning experiments, in which participants are asked to report their causal model after each observation (Bramley et al., 2017). More specifically, when asked to identify true causal models, there is a strong tendency for people to identify a model that only differs in a single aspect from the causal model they identified on the previous trial. That is, people appear to maintain a global hypothesis about the causal model

[6] Here, h_x and h_y are interchangable.

that is updated by making local changes. This behavior matches Gibbs sampling which focuses on one specific aspect of the causal models at a time and updates conditional on the rest.

Metropolis-coupled MCMC (MC^3)

While the random walk Metropolis and Gibbs samplers can be successful in exploring simple distributions, these algorithms can have issues with more complex spaces. A key example of this is provided by multimodal distributions: new states are unlikely to fall far from the current state, meaning the RWM sampler will have difficulty finding distant modes, particularly if those modes are separated by low value regions. Returning to our example of meeting my friend, if I start out sampling a restaurant that is surrounded by only steak houses then, knowing that my friend wants a curry, I might never sample beyond that initial restaurant and incorrectly conclude that this is where my friend must end up.

To allow a sampler to escape from a single mode, one strategy is to "melt" the target distribution to make the differences between probabilistic peaks and valleys less extreme. In doing so, the sampler can quickly traverse through valleys between modes (low probability regions); hence, making the remote modes easier to sample from. In statistics, the process of smoothing out a distribution can be done through controlling the computational temperature of the distribution. Increasing the temperature of a distribution leads to flatter probabilistic landscape; the distribution converges to a flat, uniform distribution when the temperature approaches infinity.

The distorted distribution, though easier to explore, is no longer the original distribution. The computational temperature, which was initially increased in order to smooth out the distribution, should be decreased back to 1 in order to ensure sampling from the target distribution. This process of heating and cooling has been implemented both serially and

parallelly in algorithms. The serial implementation is known as simulated annealing in statistics (Kirkpatrick et al., 1983); this algorithm intuitively resembles the way that metals are forged by repeatedly heating and cooling in order to align the molecules. The parallel implementation, which will be our focus, is known as Metropolis-coupled MCMC (MC^3), or parallel tempering in statistics (Geyer, 1991; Earl & Deem, 2005). This algorithm maintains multiple parallel Markov chains at different temperatures, and allows communication among chains. The higher temperature chains, which preferentially make long distance moves in the probabilistic landscape, assist lower temperature chains to make non-local jumps.

Psychological Applications of MC^3

The sampler dynamic of MC^3, due to swapping between chains, differs from that of RWM: a swap often induces a long-distance, non-local jump. As a result, the cold chain will perform local jumps on most occasions, but may intermittently swap to better locations with non-local jumps. The overall distribution of jump distances thus can resemble a power-law distribution, which is associated with Lévy flights in the animal and cognitive foraging literature, and has also been observed in people "internally foraging" for the names of animals (Todd & Hills, 2020; Rhodes & Turvey, 2007). For the same reason, MC^3 also emits long-range, non-Markovian dependencies for every individual chain. This serial dependence can be characterised as *1/f* noise, a type of long-range dependence which is ubiquitous in cognitive time series (Zhu et al., 2018; Gilden et al., 1995; Gilden, 2001).

Discussion and Conclusions

Sampling algorithms are methods for generating examples to approximate a probability distribution about which little might be known in advance. We have reviewed a number of algorithms here and discussed their psychological plausibility and how they have been deployed by researchers to explain how internal samples are generated. While many

algorithms are effective only when sampling from an unchanging distribution, there are some, such as particle filters, that are able to operate as information is obtained from the outside environment, as in information sampling.

The juxtaposition of information and internal sampling in this book suggests an interesting prospect for investigation. Perhaps the sampling algorithms that we have outlined here are not just used internally, as in sampling from memory, but perhaps also used externally for information sampling. For instance, direct sampling, in combination with a greedy policy that selects the best action given the sample, has been widely applied in multi-arm bandit problems. The combined algorithm, famously known as *Thompson sampling*, balances exploitation of the current best alternative with exploration of potentially better alternatives in a way that is optimal for the simple bandit task and many other sequential decision-making problems (Russo et al., 2017; Thompson, 1933). Human behaviour in these sequential tasks, to some extent, also corresponds to Thompson sampling (Speekenbrink & Konstantinidis, 2015; Gershman, 2019).

Finally, moving beyond the behaviour of individuals, there is a long tradition in economics of seeing trading in markets, and especially in betting and financial markets, as implementing a form of distributed computation aggregating the knowledge of distinct rational (and sometimes boundedly rational) agents (e.g., Bowles, Kirman & Sethi, 2017). It is also an intriguing possibility that individual agents viewed as *samplers* may, when able to interact with each other (whether through copying, communication, or trade), in aggregate be viewed as carrying out distributed Bayesian computation by sampling (see, e.g., Krafft et al, 2021).

References

Abbott, J. T., & Griffiths, T. L. (2011). Exploring the influence of particle filter parameters on order effects in causal learning. In *Proceedings of the annual meeting of the cognitive science society* (Vol. 33, No. 33).

Arulampalam, M. S., Maskell, S., Gordon, N., & Clapp, T. (2002). A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking. *IEEE Transactions on signal processing*, *50*(2), 174-188.

Battaglia, P. W., Hamrick, J. B., & Tenenbaum, J. B. (2013). Simulation as an engine of physical scene understanding. *Proceedings of the National Academy of Sciences*, *110*(45), 18327-18332.

Blurton, S. P., Kyllingsbæk, S., Nielsen, C. S., & Bundesen, C. (2020). A Poisson random walk model of response times. *Psychological Review*, *127*(3), 362.

Bonawitz, E., Denison, S., Griffiths, T. L., & Gopnik, A. (2014). Probabilistic models, learning algorithms, and response variability: sampling in cognitive development. *Trends in Cognitive Sciences*, *18*(10), 497-500.

Bowles, S., Kirman, A., & Sethi, R. (2017). Retrospectives: Friedrich Hayek and the market algorithm. *Journal of Economic Perspectives*, *31*(3), 215-30.

Bramley, N. R., Dayan, P., Griffiths, T. L., & Lagnado, D. A. (2017). Formalizing Neurath's ship: Approximate algorithms for online causal learning. *Psychological Review*, *124*(3), 301.

Brown, G. D., Neath, I., & Chater, N. (2007). A temporal ratio model of memory. *Psychological Review*, *114*(3), 539-576.

Brown, S. D., & Steyvers, M. (2009). Detecting and predicting changes. *Cognitive Psychology*, *58*(1), 49-67.

Chater, N., & Manning, C. D. (2006). Probabilistic models of language processing and acquisition. *Trends in Cognitive Sciences*, *10*(7), 335-344.

Dasgupta, I., Schulz, E., & Gershman, S. J. (2017). Where do hypotheses come from? *Cognitive Psychology*, *96*, 1-25.

Daw, N., & Courville, A. (2008). The pigeon as particle filter. *Advances in Neural Information Processing Systems*, *20*, 369-376.

Dellaportas, P., & Roberts, G. O. (2003). An introduction to MCMC. In *Spatial statistics and computational methods* (pp. 1-41). Springer, New York, NY.

Doucet, A., De Freitas, N., & Gordon, N. (2001). An introduction to sequential Monte Carlo methods. In *Sequential Monte Carlo methods in practice* (pp. 3-14). Springer, New York, NY.

Earl, D. J., & Deem, M. W. (2005). Parallel tempering: Theory, applications, and new perspectives. *Physical Chemistry Chemical Physics*, *7*(23), 3910-3916.

Fox, C. R., & Tversky, A. (1998). A belief-based account of decision under uncertainty. *Management Science, 44,* 879–895.

Geman, S., & Geman, D. (1984). Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. *IEEE Transactions on pattern analysis and machine intelligence*, (6), 721-741.

Gershman, S. J. (2019). Uncertainty and exploration. *Decision*, *6*(3), 277.

Gershman, S. J., Blei, D. M., & Niv, Y. (2010). Context, learning, and extinction. *Psychological Review*, *117*(1), 197.

Gershman, S. J., Vul, E., & Tenenbaum, J. B. (2012). Multistability and perceptual inference. *Neural Computation*, *24*(1), 1-24.

Geweke, J. (1989). Bayesian inference in econometric models using Monte Carlo integration. *Econometrica: Journal of the Econometric Society*, 1317-1339.

Geyer, C. J. (1991). Markov chain Monte Carlo maximum likelihood.

Gilden, D. L. (2001). Cognitive emissions of *1/f* noise. *Psychological Review*, *108*(1), 33.

Gilden, D. L., Thornton, T., & Mallon, M. W. (1995). *1/f* noise in human cognition. *Science*, *267*(5205), 1837-1839.

Gittins, J. and D. Jones. 1979. “A dynamic allocation index for the discounted multiarmed bandit problem”. *Biometrika.* 66(3): 561– 565.

Griffiths, T. L., & Tenenbaum, J. B. (2006). Optimal predictions in everyday cognition. *Psychological Science*, *17*(9), 767-773.

Griffiths, T. L., Vul, E., & Sanborn, A. N. (2012). Bridging levels of analysis for probabilistic models of cognition. *Current Directions in Psychological Science*, *21*(4), 263-268.

Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, *3*(3), 430-454.

Koller, D., & Friedman, N. (2009). *Probabilistic graphical models: principles and techniques*. MIT press.

Krafft, P. M., Shmueli, E., Griffiths, T. L., & Tenenbaum, J. B. (2021). Bayesian collective learning emerges from heuristic social learning. *Cognition, 212,* 104469.

Kuhn, T. (1962). *The Structure of Scientific Revolutions*, Chicago, IL: University of Chicago Press.

Levy, R. P., Reali, F., & Griffiths, T. L. (2009). Modeling the effects of memory on human online sentence processing with particle filters. In *Advances in Neural Information Processing Systems* (pp. 937-944).

Lieder, F., Griffiths, T. L., & Hsu, M. (2018). Overrepresentation of extreme events in decision making reflects rational use of cognitive resources. *Psychological Review*, *125*(1), 1.

Lloyd, K., Sanborn, A., Leslie, D., & Lewandowsky, S. (2019). Why higher working memory capacity may help you learn: Sampling, search, and degrees of approximation. *Cognitive Science*, *43*(12), e12805.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equation of state calculations by fast computing machines. *The Journal of Chemical Physics*, *21*(6), 1087-1092.

Mosteller, F., & Nogee, P. (1951). An experimental measurement of utility. *Journal of Political Economy*, *59*(5), 371-404.

Murphy, K. P. (2012). *Machine learning: a probabilistic perspective*. MIT press.

Nobandegani, A. S., Castanheira, K. D. S., Otto, A. R., & Shultz, T. R. (2018). Over-representation of extreme events in decision-making: A rational metacognitive account. *arXiv preprint arXiv:1801.09848*.

Nosofsky, R. M. (1984). Choice, similarity, and the context theory of classification. *Journal of Experimental Psychology: Learning, memory, and cognition*, *10*(1), 104-114.

Nosofsky, R. M. (2011). The generalized context model: An exemplar model of classification. *Formal approaches in categorization*, 18-39.

O'Hagan, A., & Forster, J. J. (2004). *Kendall's advanced theory of statistics, volume 2B: Bayesian inference* (Vol. 2). Arnold.

Pleskac, T. J., & Busemeyer, J. R. (2010). Two-stage dynamic signal detection: a theory of choice, decision time, and confidence. *Psychological Review*, *117*(3), 864.

Prat-Carrabin, A., Wilson, R. C., Cohen, J. D., & Azeredo da Silveira, R. (2021). Human inference in changing environments with temporal structure. *Psychological Review, 128*(5), 879–912.

Raaijmakers, J. G., & Shiffrin, R. M. (1981). Search of associative memory. *Psychological Review*, *88*(2), 93-134.

Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*(5), 347-356.

Rhodes, T., & Turvey, M. T. (2007). Human memory retrieval as Lévy foraging. *Physica A: Statistical Mechanics and its Applications*, *385*(1), 255-260.

Roth, D. (1996). On the hardness of approximate reasoning. *Artificial Intelligence*, *82*(1-2), 273-302.

Russo, D., Van Roy, B., Kazerouni, A., Osband, I., & Wen, Z. (2017). A tutorial on Thompson sampling. *arXiv preprint arXiv:1707.02038*.

Sanborn, A. N., & Chater, N. (2016). Bayesian brains without probabilities. *Trends in Cognitive Sciences*, *20*(12), 883-893.

Sanborn, A. N., Griffiths, T. L., & Navarro, D. J. (2010). Rational approximations to rational models: alternative algorithms for category learning. *Psychological Review*, *117*(4), 1144.

Sanborn, A. N., Mansinghka, V. K., & Griffiths, T. L. (2013). Reconciling intuitive physics and Newtonian mechanics for colliding objects. *Psychological Review*, *120*(2), 411.

Shadlen, M. N., & Shohamy, D. (2016). Decision making and sequential sampling from memory. *Neuron*, *90*(5), 927-939.

Shi, L., Griffiths, T. L., Feldman, N. H., & Sanborn, A. N. (2010). Exemplar models as a mechanism for performing Bayesian inference. *Psychonomic Bulletin & Review*, *17*(4), 443-464.

Sloman, S., Rottenstreich, Y., Wisniewski, E., Hadjichristidis, C., & Fox, C. R. (2004). Typical versus atypical unpacking and superadditive probability judgment. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 30,* 573–582.

Speekenbrink, M. (2016). A tutorial on particle filters. *Journal of Mathematical Psychology*, *73*, 140-152.

Speekenbrink, M., & Konstantinidis, E. (2015). Uncertainty and exploration in a restless bandit problem. *Topics in cognitive science*, *7*(2), 351-367.

Spicer, J., Mullett, T., & Sanborn, A.N. (submitted). Repeated risky choices become more consistent with themselves but not expected value, with no effect of trial order.

Stigler, S. M. (1986). *The history of statistics: The measurement of uncertainty before 1900*. Harvard University Press.

Thompson, W. R. (1933). On the likelihood that one unknown probability exceeds another in view of the evidence of two samples. *Biometrika*, *25*(3/4), 285-294.

Tierney, L. (1994). Markov chains for exploring posterior distributions. *the Annals of Statistics*, 1701-1728.

Todd, P. M., & Hills, T. T. (2020). Foraging in mind. *Current Directions in Psychological Science*, *29*(3), 309-315.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. *Cognitive Psychology*, *5*(2), 207-232.

Usher, M., & McClelland, J. L. (2001). The time course of perceptual choice: the leaky, competing accumulator model. *Psychological Review*, *108*(3), 550.

Van Rooij, I. (2008). The tractable cognition thesis. *Cognitive Science*, *32*(6), 939-984.

Von Neumann, J., & Morgenstern, O. (2007). *Theory of games and economic behavior (commemorative edition)*. Princeton university press.

Vul, E., & Pashler, H. (2008). Measuring the crowd within: Probabilistic representations within individuals. *Psychological Science*, *19*(7), 645-647.

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.

Vulkan, N. (2000). An economist's perspective on probability matching. *Journal of Economic Surveys*, *14*(1), 101-118.

Wagenaar, W. A. (1972). Generation of random sequences by human subjects: A critical survey of literature. *Psychological Bulletin*, *77*(1), 65-72.

Wolpert, D. M. (2007). Probabilistic models in human sensorimotor control. *Human Movement Science*, *26*(4), 511-524.

Yi, M. S., Steyvers, M., & Lee, M. (2009). Modeling human performance in restless bandits with particle filters. *The Journal of Problem Solving*, *2*(2), 5.

Zhu, J.-Q., Sanborn, A., & Chater, N. (2018). Mental Sampling in Multimodal Representations. *Advances in Neural Information Processing Systems*, *31*, 5748-5759.