Sampling Data, Beliefs, and Actions

Erik Brockbank, Cameron Holdaway, Daniel Acosta-Kane & Edward Vul

University of California, San Diego

Author Note

Correspondence concerning this article should be addressed to Erik Brockbank, Department of Psychology, University of California San Diego, 9500 Gilman Drive, La Jolla, CA 92037. E-mail: ebrockbank@ucsd.edu

Abstract

Sampling—using a stochastically drawn subset of possibilities—has been at the core of many influential modeling frameworks of human decision making for the last half century. Although these frameworks all refer to their core operation as "sampling," they differ dramatically in the behaviors and inferences they aim to account for. Here we review this landscape of sampling models under a unified expected utility framework which treats diverse sampling accounts as approximating different terms in the expected utility calculation. We show that a broad range of sample-based models in psychology are built around sampled *data*, *beliefs*, or *actions* and can therefore support downstream expected utility maximization. To compare these models on even footing, our review focuses on how the *number of samples* and the *sample distribution* differ within each element of the expected utility calculation. This integrated summary allows us to identify opportunities for fruitful cross-pollination across sampling domains, and to highlight outstanding challenges for accounts that might aim to integrate these disparate models.

Keywords: sampling, data, beliefs, actions, expected utility

Sampling Data, Beliefs, and Actions

Introduction

This volume tackles the ambitious challenge of reviewing the heterogeneous set of results, phenomena, accounts, and models that fall under the banner of sampling in judgment and decision making. These accounts address different domains of behavior, in different contexts, with different constraints, yet the reliance on "sampling" suggests that they share some central properties. In this chapter we aim to provide an organizing framework for a sample of this literature, to highlight the conceptual differences and the core similarities across domains.

What do sampling accounts have in common? Indeed, what constitutes a "sampling" account of human behavior? In all cases, sampling accounts posit that people perform some calculation using a small *subset* of all the values that are relevant for the calculation. This subset is often generated through stochastic processes over which the person may not have full control. In recent years, the use of sampling as a mechanism to explain various forms of inference and decision making has spread across broad domains of psychology. The most familiar type of sampling account is concerned with *observations* of the world (Fiedler & Juslin, 2006). These accounts are motivated by the similarity between what a statistician must do to draw inferences from limited data, and what individual humans must do to act based on noisy, sparse observations of their environment. These accounts formalize the notion that the world is far larger, more complicated, and more dynamic than any one person can apprehend. Consequently, people must act based on a small subset of observations—a sample—rather than a complete snapshot of the world state. Though initially developed to explain inferences about perception and the physical state of the world (Braddick, 1974), these models have been applied to a range of more abstract inferences about social structure and the behavior of others (Fiedler, 2000). However, sampling has also been proposed as a modeling framework for the *computations* that people undertake to revise their beliefs, or to generate predictions, about the state of the world (Griffiths et al., 2012). On this view, human beliefs are probability distributions over possible states of the world, and prior beliefs are updated in light of new data via Bayes' rule to form posterior beliefs. This belief updating is usually analytically intractable and in applied settings is often approximated via Monte Carlo methods that draw samples from the posterior (Doucet et al., 2001). Work in this space has proposed that similar strategies are adopted by the brain; beliefs about world states are represented as finite sets of samples. More recently, sampling has also been postulated as a strategy for choosing actions (Phillips et al., 2019). Under these accounts, the set of possible actions is impossibly large, so choosing the best one cannot be accomplished by evaluating the prospective quality of *all* actions. Instead, we sample a subset of actions to consider and evaluate, then pick the best action from among those sampled. Far from being a comprehensive overview, these examples illustrate the wide reach that sampling has attained in computational models of decision making; the additional role attributed to samples from memory for instance (Nosofsky & Palmeri, 1997) broadens the purview of sampling accounts even further.

Given this diversity of sampling models in the literature and the range of research goals they support, we might ask whether it's helpful to view them as members of a coherent class of "sampling accounts" at all. Would we be better off instead choosing distinct terms for each, to avoid confusion? The premise of this chapter is that a unified review of these accounts will not only help disentangle their differences, but can also provide structure for considering their commonalities and identifying opportunities for cross-pollinating research ideas across fields. We propose that these different sampling accounts are all approximating different components of expected utility maximization: *data*, *beliefs*, and *actions*. We review a broad swath of sampling literature within this framework and show how describing sampling accounts as approximating specific components of expected utility clarifies how they are related and poses novel questions and comparison points for seemingly distinct areas of research. We close by highlighting opportunities for synthesis across sampling domains and outlining challenges that arise in light of the fact that in real-world decisions, all of these components must be sampled together.

Expected utility framework

We propose that most sampling accounts can be best understood in the context of choosing an action to maximize expected utility. The computations necessary to pick an action that maximizes expected utility require combining *data*, marginalizing *beliefs*, and optimizing over *actions*. Within this framework, existing sampling accounts describe the consequences of using small subsets of alternatives to approximate each of these operations.

Sampling to approximate expected utility

Many of our everyday decisions revolve around choosing the best action in a given situation. *What will we wear to a friend's party in the evening and what should we bring? What will be the best route to drive there? How long should we stay?* All of these decisions involve selecting an ideal course of action from among many. How do people evaluate their choices and ultimately settle on one? We start with the expected utility hypothesis, the assumption that for an intelligent agent, the best choice (A^*) is the action that yields the largest expected utility ($U(\cdot)$).

$$A^* = \arg\max_A \mathbb{E}[U(A)] \tag{1}$$

.

This formulation, though intuitive, is ultimately a crude decision making policy because it disregards context-dependent variation in action outcomes. For intelligent behavior, we would expect the payoffs of actions to vary as a function of context, as indicated to the agent by observed data (x).

$$A^* = \arg\max_A \mathbb{E}[U(A \mid x)] \tag{2}$$

.

In this richer formulation, utilities associated with actions vary based on observed data; we *condition* on the data x to yield different predicted action outcomes. Expected utility in Equation 2 is equivalent to the Q-function in reinforcement learning (Watkins & Dayan, 1992), associating payoffs with specific action-context combinations. While equating context with observed data is sufficient in settings with simple, unambiguous observations, in more realistic settings where observations are complex and noisy, it is more parsimonious to consider payoffs to be contingent on a latent state (s), rather than the data directly. Under this formulation, the utility function abstracts away from data, considering utility to be a property of an action in a state. Uncertainty arises from ambiguity about what state the agent is in.

$$A^* = \arg\max_A \sum_s U(A; s)P(s \mid x) \tag{3}$$

.

Equation 3 emphasizes the different roles that data (x), beliefs ($P(s)$), and actions (A) play in decision making: We *optimize* over actions ($\arg\max_A$), we take an expectation, or *marginalize*, over beliefs ($\sum_s f(s)P(s)$), and we *condition* on data ($P(s \mid x)$).[1] Choosing an action that maximizes expected utility requires that an agent perform all of these calculations in order to decide what to do. Given an infinite number of available actions and possible states of the world, these calculations are computationally intractable in their general formulation. Thus, human behavior is typically seen as approximating the expected utility maximization described in Equation 3 (S. Gershman et al., 2015).

Broadly, sample-based accounts of behavior and decision making propose that when

[1] We rely on this simple formulation throughout the paper. However, a more thorough model-based view of utility would rewrite $U(A; s)$ based on state transitions arising from actions: $U(A; s) = \sum_{s'} U(s')P(s' \mid A, s)$. This ascribes utilities to particular states, rather than state-action combinations, and has the advantage of clarifying the role of *predictive* beliefs—beliefs not just about the current state (s), but also about future states (s'). This formulation may be further expanded to consider

solving problems which are intractable via brute force calculation, people rely on a limited set of samples to *approximate* the underlying calculation. This can be done for each of the operations in Equation 3 described above. Instead of considering all the available information, people make decisions based on a sample of observations, acting as "intuitive statisticians" (Fiedler & Juslin, 2006). Likewise, instead of considering the full probability distribution over states of the world, *Monte Carlo* methods (Robert & Casella, 2004) demonstrate that one can consider a sample of possible states to achieve the same end. Finally, numerical optimization techniques show that a global optimum such as the best action in a given context may be found with high probability by considering a small subset of alternatives (Karnopp, 1963). In this way, sample-based accounts offer an answer to the question of how people might maximize expected utility in their decisions; an approximate solution to Equation 3 can be estimated using *samples* from the relevant distributions over data, beliefs, and actions without relying on the full underlying distributions. This approximation of classical expected utility (Von Neumann & Morgenstern, 1947) contrasts somewhat with forms of *subjective* expected utility (Savage, 1954) which argue that the perceived utility or probability of an outcome may be fundamentally distorted from the true underlying probability or value (a prominent example is the argument in *prospect theory* (Tversky & Kahneman, 1979) that the utility curve for losses and gains is asymmetric). As we describe below, various distortions to the perceived probability or utility of an outcome can arise from the process of sampling data, beliefs, or actions, and some of these may lead to behavior that is consistent with prominent accounts like prospect theory. However, an approximation that distorts the true probability can be considered somewhat distinct from maintaining a high fidelity but asymmetric utility curve.

To illustrate this sample-based account of approximate expected utility

---

sequential decisions. The proposal in this chapter, that diverse sampling accounts support approximate expected utility maximization, applies equally to expanded versions of Equation 3, but for our purposes the version in Equation 3 is sufficient.

maximization, imagine the familiar and sometimes daunting task of deciding what to wear to a party. Here, selecting the action that maximizes expected utility amounts to choosing the best outfit. A brute force solution would require examining every item of clothing that may already be available in one's closet, or which might be purchased in advance of the party. One must then calculate the expected utility of each combination of garments. However, we often opt for a simpler solution: sample a set of *actions*, i.e., candidate outfits from our wardrobe, and evaluate those. Formally, the $\arg\max_A$ in Equation 3 is evaluated over a *sample* of actions in A. But here too, we face a challenge in calculating the expected utility of a given outfit because it will be impacted by incidental information about the state of the world, like what sort of party we are going to. To calculate the expected success of a candidate outfit, we would need to consider how well that outfit would be received in every conceivable party: a pool party or a drunken wake, a reception by an ambassador or a birthday party for a three-year old, etc. To calculate the expected utility of a given outfit, a rational agent must marginalize over all of these possible states weighted by their probability in light of all the information they have about the party in advance. Obviously, we do not do anything so thorough. In this case, we might begin by sampling *data* from the external world which supports our decision making: who else is going to this party? What does the invitation look like? Is it somebody's birthday? If it's outside, what will the weather be like? And how late will it go? These data samples are used to update *beliefs* about what kind of party we are attending, which impact the expected utility estimation. This means that the $P(s \mid x)$ in Equation 3 is evaluated based on only a sample of data points in x—we do not ask *all* of our friends whether they will be going, but just a few. Further, we do not then consider every possible belief we might have about the party. Instead, we consider just a few different prospective party environments, sampled with frequency proportional to their probability under $P(s \mid x)$. This set of samples is an approximate representation of $P(s \mid x)$, and allows us to estimate the expected value of an outfit without considering infinitely many party possibilities. This makes our expected utility calculation in (3) computationally much simpler. We choose the best of a subset of possible actions (outfits), based on its utility averaged over a subset of possible states (party environments), sampled according to their probability in light of a sample of relevant data (e.g., who is attending): $A^* = \arg\max_{a \in \mathbb{A}} \sum_{i=1}^{n} U(a; s_i)/n$, where $s_i \sim P(s \mid x)$. This process is illustrated in Figure 1. Critically, decisions such as this are sometimes difficult when we choose to evaluate our options carefully, yet they are fairly commonplace[2] and people, for the most part, do them well.

How do people regularly make such rich, sample-based inferences and what are the limitations in this ability? We propose that the many sampling accounts in the psychology literature offer insights into these questions when viewed within the context of sampling data, beliefs, and actions to maximize expected utility. This framework may not cleanly encompass *every sampling model* or example of sample-based inference in decision making, and some important classes of sampling models, such as sampling from memory, may be cross-cutting categories that play a role in multiple terms of the expected utility calculation. Nonetheless, we argue that framing a large swath of existing models and their contributions as supporting expected utility approximation provides both a unified perspective on sampling research, and, critically, a means of guiding future research in these areas: how can prior work on *data sampling* accounts inform *belief sampling* research? And what does the literature on belief sampling tell us about models of *action sampling*?

Comparing sample-based accounts

Though a diverse set of sample-based accounts might in theory be integrated under expected utility calculation as described above, comparing them with this lens is only useful if it provides novel insights or future research directions. However, evaluating the

[2] This process may arise not just in one-off decisions but in repeated decisions or those made over a longer time frame. We thank a reviewer for pointing out that the iterative process of behavioral research itself relies on sampling data which approximates an underlying distribution, sampling beliefs about the data to formulate or refine hypotheses, and choosing from among sampled actions to pursue further inquiry.

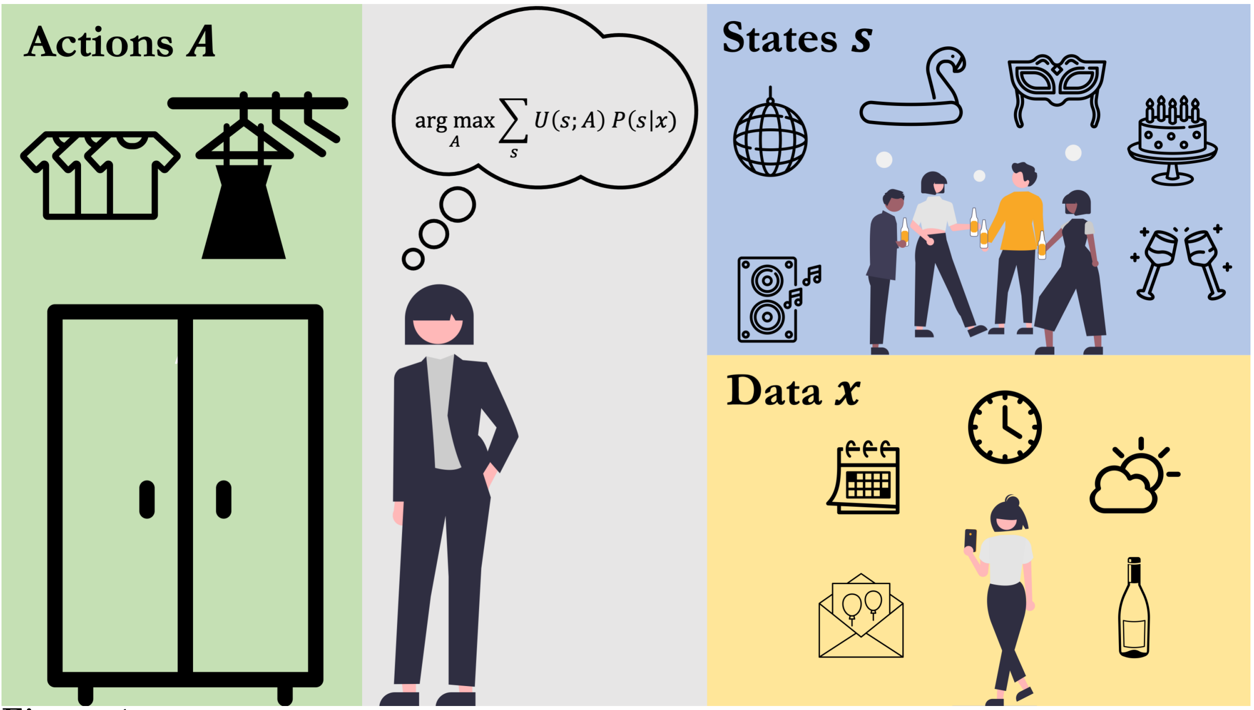

Figure 1

The expected utility framework for sample-based models in decision making, shown here for the decision about what to wear to a party. ***Left:*** *when making decisions, we typically sample a subset of options from the infinite space of possible actions (A), e.g., a handful of possible outfits.* ***Top right:*** *actions are evaluated based on their utility averaged over possible world states, sampled according to the agent's beliefs (probabilities of states s), such as what kind of party it will be.* ***Bottom right:*** *beliefs are informed by data (x) sampled from the external world, such as what the weather will be like and what time the party is.*

shared constraints and challenges across different sampling domains is itself non-trivial. For one, because of the unique questions motivating each of these research traditions, these different types of sampling have largely been considered in isolation. Further, out of practical considerations, researchers studying data, belief, or action sampling mostly ignore the other variables to provide greater experimental control. For example, researchers interested in data sampling typically design tasks where state estimates are determined by the data (i.e., there is little role for priors or uncertainty arising from internal models), and where the set of available actions is small and explicitly provided. Consequently, all the behavioral variability may be attributed to the data sampling process and calculations about beliefs or actions may be safely ignored. Research on belief and action sampling place similar restrictions on the other facets of expected utility maximization so as to effectively isolate the component under study. This raises the question of how to compare these distinct sampling paradigms with the goal of integrating them into a more unified account of sampling in decision making. Can we view data, belief, and action sampling as solving the same kinds of problems or operating within similar constraints? Or are they better understood as solving sufficiently different problems that their similarities stop at the use of samples? If the latter is the case, the mere inclusion of these models in a broader expected utility paradigm offers little additional ability to compare them or make progress on one by appeal to the others.

We believe that models of data, belief, and action sampling face a number of similar constraints which allow them to be usefully compared, and the successes of one potentially applied to the others. The constraints they share, and the point of departure for comparing them in this chapter, draws on what we consider to be essential aspects of any sampling model. Critically, a sampling-based account of behavior is only useful insofar as it makes divergent predictions from what would be expected if people were not sampling but instead considering the full set of possible alternatives. For instance, action sampling is only interesting insofar as it predicts something different than considering all possible actions. Where sampling accounts make identical predictions as inference based on complete information, we may have little reason to prefer the sampling models. What then differentiates a sampling model from one based on complete *analytical* inference? We propose that nearly all sampling models can be compared along two key dimensions that distinguish their predictions from alternative accounts: the *number of samples* and the *sampling distribution*. These features are central to what it *means* to be a sampling account and, critically, form the basis of a sampling account's unique predictions in decision making contexts.

Consider first the number of samples. Under all sampling accounts, if infinitely many samples are used, this is no different from relying on the full set of possibilities, so nothing is gained from postulating sampling. In contrast, when a decision maker relies on only a small number of samples, their inferences may be biased in various ways by the limits in their sampling. Therefore, the number of samples used, and the considerations that influence this, are central to the predictions that a sampling account offers. Here, we aim to show that how many samples are drawn and how this number is determined can be asked equally of data, belief, and action sampling models, and the answers provided by one can be useful for the others. Second, the *distribution generated* by sampling is similarly critical for a sampling account to make concrete predictions. Under optimal conditions, independent and identically distributed samples might lead to decisions consistent with use of the full distribution, in which case sampling models may not be identifiable. However, in many cases it is simply impossible to generate idealized samples, and whatever algorithms are used to obtain the samples will necessarily create some systematic deviation in the set of samples which will affect downstream behavior (Juni et al., 2016; Sanborn, 2017). Further, even samples obtained optimally for one purpose are likely to yield systematic biases with respect to another goal (Fiedler, 2008). For instance, when sampling in low-probability, high-stakes situations, one might either correctly estimate the probability of each outcome (by sampling according to probability), or correctly estimate the utility maximizing action payoffs (by sampling according to probability weighted by utility), but no adjustment strategy can allow both decisions to be unbiased using a small number of samples (Lieder, Griffiths, & Hsu, 2018; Vul et al., 2014). Thus, the sample distribution, like the number of samples taken, is a critical feature of sampling models that determines the particular ways in which sample-based decisions deviate from optimal behavior.

Given the fundamental role that the *number of samples* and the *sample distribution* play in sample-based models, we use these two characteristics as a basis for reviewing data, belief, and action sampling accounts of behavior under the umbrella of expected utility maximization. In what follows, we discuss how results in data, belief, and action sampling vary along these axes, and critically, how each one can inform future research in the others. In restricting the scope of our review to consider only the number of samples and the sample distribution, we do not intend to overlook other meaningful aspects of sampling models, such as the costs and benefits of samples, the algorithms that yield them, or how explicit the sampling process is. Instead, we propose that the consequences of these and other noteworthy sources of variation in sampling procedures are largely captured by virtue of their role in the number, and the distribution, of samples. As we show, comparison of sampling accounts along these dimensions alone offers fertile ground for identifying the major contributions of existing sampling models, as well as opportunities for future work which might improve our understanding of how people use samples to maximize the expected utility of their decisions.

Overview of sampling models in the literature

This chapter proposes that sampling accounts of behavior can be fruitfully examined within a unified framework of expected utility maximization. Here, we illustrate this process, reviewing a large range of sample-based models that are consistent with data, belief, and action sampling in sequence. For each of these sampling domains, we address considerations of the number, and the distribution, of samples, and how these results might inform or benefit from the other classes of sampling model.

Data sampling

At their core, models of *data sampling* are about gathering information which will reduce uncertainty about the present environment to support better (i.e., more informed) downstream decisions. In the example we provided at the outset (Figure 1), in which a person seeks to maximize the expected utility of possible clothes to wear to a party, data sampling amounts to seeking information which will refine their belief about the sort of party they are attending, e.g., what will the weather be like and who else will be there? From an expected utility standpoint, the role of each piece of data, x, is to improve posterior estimates over states. In most simple settings, each datum constrains the posterior distribution, $P(s \mid x)$, multiplicatively: $P(s \mid x) \propto P(s) \prod_x P(x \mid s)$. Given this broad formulation, we consider any process that obtains information from the outside world and supports subsequent decision making to be an instance of data sampling. Because acquiring information about the surrounding environment is a critical behavior for most if not all animals, some of the earliest sample-based models in psychology—and, as we'll discuss, some of the most well-formalized—have concerned data sampling.

The number of samples in data sampling

Data sampling models are rooted in Fechner (1860)'s two-alternative forced choice (2AFC) experimental paradigm, in which observers make repeated binary classifications of stimuli. In a canonical example, random dot kinematograms (Braddick, 1974) present participants with a field of dots each moving to the left or right. People are asked to judge the prevailing direction of the dots in each image as quickly and accurately as possible. Typical behavior in the task reflects a *speed-accuracy tradeoff*; intuitively, responding more quickly on any given trial decreases the probability of answering correctly, while increasing the number of subsequent trials that can be completed.

The 2AFC paradigm provides a precise and highly controlled environment for examining how people sample data from the external world. Computational accounts of behavior in the 2AFC task fall into the broad class of *drift diffusion models* (DDM) (Ratcliff & McKoon, 2008; Ratcliff & Smith, 2004) (see Voss et al. (2013) for a practical guide), which owe their origins to the *sequential probability ratio test* (SPRT) (Wald & Wolfowitz, 1948). Using speed and accuracy data for each participant, DDMs fit a decision threshold a which represents the level of certainty required to choose either option (essentially the desired level of accuracy), and a drift rate v which corresponds to the rate at which people accumulate evidence in the task (related to their task speed). The decision threshold indicates how much evidence people accumulate before making a decision, which imposes a distribution on the number of stimulus samples they consider in their choice (Vul et al., 2014). In this way, DDMs provide a descriptive account of how different contextual or environmental features, e.g., the proportion of left and right stimuli (Ratcliff & McKoon, 2008), impact the number of samples a subject uses to make a decision.

Critically, DDMs not only allow for precise characterization of the number of samples taken in data sampling settings, but also lend themselves to rational analysis: how many samples *should* one take in a given context? The decision policies adopted by the SPRT and DDM are *optimal* in the sense that they yield the fastest decision times for a given level of accuracy (Bogacz et al., 2006). This in turn reflects an optimality of data sampling; these models minimize the number of samples needed to achieve a desired level of accuracy. Critically, optimal thresholds in these models must reflect the objective or utility functions of the specific tasks: what are the relative benefits of speed and accuracy? Thus, decision thresholds connect an optimal agent's objectives and the number of samples they take. For instance, one often used objective function is maximizing reward rate, as determined by task specific parameters. Given a particular time cost of samples, fixed non-decision processing time, and inter-trial delays associated with both correct and incorrect responses, one can find a threshold function that optimizes the rate of reward. This function will in turn dictate an optimal number of samples. Other objective functions like minimizing Bayes risk can similarly be mapped onto specific threshold functions (Bogacz et al., 2006). Broadly, the SPRT and DDM therefore provide a normative approach to how many data samples to draw in 2AFC tasks.

Given the success of DDMs in characterizing both empirical and optimal data sampling behavior, these models provide a template for studying finite sampling across domains, but also highlight important directions for future development within data sampling. DDMs have been best characterized in 2AFC tasks, but many important human behaviors deviate from this simple case, for instance by facing choices with more than two alternatives, or data samples that are actively selected rather than passively received. In these scenarios, we must turn to more elaborate sampling algorithms that can grapple with these complications. However, even in these richer settings, the same core considerations clearly identified by the DDM paradigm will apply: what are the costs and benefits of samples, and how can a stopping rule be chosen to optimize an objective function for a given problem? In short, the formalization of number of samples offered by DDMs can guide efforts at similar precision in other related tasks.

The sample distribution in data sampling

In data sampling models, samples are stochastic observations of the world which people use to reduce uncertainty and inform action choices according to Equation 3. Notably, the drift diffusion paradigm typically places people in the role of passive observer, presented with *natural* samples (Gigerenzer & Hoffrage, 1995) that correspond to the likelihood function observers use (Ratcliff & Smith, 2004). In this canonical formulation, departures from rational behavior cannot be explained by the sample distribution as it is reflected in the likelihood function $P(x \mid s)$, since the sample-generating process and the observer's model thereof, are presumed to match.

However, a large body of research has examined what happens when samples from the decision maker's environment *do not* represent IID samples from the data distribution presumed by observers. Research examining decisions based on biased samples has mostly done so in the context of social inferences about the people around us. While this represents a far more abstract domain than, say, assessing the prevailing direction of dot kinematograms, the paradigm of sampled data from one's environment supporting downstream approximation is largely the same. For example, when people were asked to provide estimates for a range of health and well-being measures in the general population (e.g., income and education levels, work stress and health problems), their responses showed systematic biases relative to their own self-appraisal on these metrics; the biases are well predicted by a model in which people's inferences were based on samples from their immediate environment, which may have differed from the population distribution substantially (Galesic et al., 2012, 2018). Similar accounts based on biased samples from one's surroundings have been proposed in other domains of social reasoning, such as the role of peers in determining social attitudes and the robust tendency to judge one's in-group as more heterogeneous than one's out-group (Konovalova & Le Mens, 2020). These findings fall under the broad umbrella of *wicked* learning environments (Hogarth et al., 2015), in which the sampled data from which people learn and generalize deviates systematically from the population or "test" data (this is contrasted with *kind* learning environments in which there is a closer relation and any divergence is primarily a result of noise). Hogarth et al. (2015) show that a broad range of robust biases such as survivorship bias and the "hot stove" effect can be accounted for by small samples drawn in various kinds of wicked environments.

The previous examples suggest that biases in our social judgments and attitudes may be explained by *external* processes that systematically distort the distribution of data we sample in the course of everyday experience. However, there are a number of ways in which our *own* behavior can further bias the sample distribution of data on which we base decisions and actions. Learning and decision making scenarios in which people actively query the environment may require them to make critical decisions about what kind of data to sample and when to stop sampling. These *active learning* paradigms are often designed to reflect more naturalistic scenarios; in the example in Figure 1, deciding what to wear to a party has this character, since a person can exercise some control over how many friends to ask about party attendance, and how long to persist. Work on data sampling in more active settings has shown that decisions about when to stop sampling and what to sample can lead to biased sample distributions which account for further idiosyncrasies in decision making based on these samples.

First, consider the decision about when to stop sampling. In the *multi-armed bandit* paradigm (N. D. Daw et al., 2006; Gittins, 1979) participants sample outcomes from two "bandits" or levers, each of which has a unique underlying reward probability (or a distribution, e.g., 90% chance of one dollar and a 10% chance of five dollars). At the end of the trial period, participants make a (usually binary) decision about which bandit they will select for their final reward (Hertwig & Erev, 2009). A large body of work exploring binary choices in these settings finds that people tend to take very few samples, even when the cost of samples is negligible (Hau et al., 2010; Hertwig et al., 2004; Hertwig & Erev, 2009; Hertwig & Pleskac, 2010). Such sparse sampling tends to under-represent low-probability outcomes, and this bias in sample distributions can impact downstream behavior. While the biased distribution may make certain actions more efficient (e.g., choosing from among two gambles), people are unlikely to correct for this bias when their goals change or they are confronted with a new task, e.g., estimating the underlying distributional characteristics (Coenen & Gureckis, 2021; Jazayeri & Movshon, 2007).[3]

In addition to decisions about when to stop sampling, decision makers in active learning contexts can make goal-directed decisions about *what data* to sample. For example, when deciding where to go for dinner, we might sample reviews for a particular cuisine and look for positive ones or seek out positive samples and then choose a cuisine from among them. Such *selective sampling* can greatly increase the efficiency of decision making relative to natural sampling (Fiedler, 2008),[4] but selective sampling of a particular variable will necessarily produce biased samples with respect to that variable's base rate or to other correlated variables (Dawes, 1993). In problems of information search, people show little ability to correct for these biases in subsequent judgments; broadly, where data sampling reflects distortions from the underlying distribution brought about by people's

[3] In active learning settings where samples are generated through behavior, the line between data sampling and action sampling may seem somewhat blurred. For our purposes, we aim to distinguish between sampling actions for consideration, and trying actions to learn from the data they generate.

sampling choices, their downstream judgments will be similarly biased (Fiedler, 2000, 2008).

In some cases, decisions about what kind of data to sample can reflect broad biases in information search, as in the case of positive test strategies and confirmation bias (Klayman & Ha, 1987). Recent work has attempted to characterize the rational principles and goals that might guide such biased sampling decisions, for example maximizing expected information gain (Rothe et al., 2018) or testing *sparse hypotheses* (Navarro & Perfors, 2011; Oaksford & Chater, 1994). However, identifying such guiding principles in data sampling can be challenging (Rothe et al., 2018) and, more importantly, it remains the case that when such principles lead to biased samples, people typically fail to correct for these distortions in the sample distribution, leading to a range of familiar behavioral biases (Coenen & Gureckis, 2021; Fiedler, 2000).

Belief sampling

In the last few decades, prominent models of higher-level cognition, reasoning, and decision making have relied on probabilistic inference over rich internal knowledge structures (Chater et al., 2006; Knill & Richards, 1996; Oaksford, Chater, et al., 2007; Tenenbaum et al., 2011). These probabilistic reasoning accounts postulate that human beliefs can be characterized as probability distributions like the ones supporting expected utility maximization in Equation 3: the posterior distribution over states conditioned on observations, $P(s \mid x)$. In the example discussed at the beginning, the sampled set of beliefs follow this pattern; what kind of party we are attending and whether there will be dancing allow us to approximate a more complicated distribution over states given the available data (see Figure 1). This illustrates one of the central challenges of belief

[4] It is worth noting that samples from memory may also be considered along these lines; decision making can be aided by reaching back in memory for relevant experiences that will inform the current decision. However, due to the breadth and complexity of such models, we save this for the discussion.

sampling models, namely that estimating the underlying probability distributions may be arbitrarily difficult depending on what constitutes a "state," as well as the complexity of the models postulated for the prior ($P(s)$) and likelihood ($P(x \mid s)$) distributions.

The sophistication of these models has raised concerns about the plausibility of the human brain carrying out such fundamentally intractable inference (Gigerenzer & Goldstein, 1996; Jones & Love, 2011; Kwisthout et al., 2011). This tension has created an active area of research on biologically and psychologically plausible inference algorithms that might approximate probabilistic inference over large knowledge structures (Tenenbaum et al., 2011), with the most attention paid to variations of sampling algorithms (E. Bonawitz, Denison, Gopnik, et al., 2014; Sanborn et al., 2010; Shi et al., 2010). These *belief sampling* accounts propose that the probability distributions over knowledge structures—i.e., beliefs, under the probabilistic reasoning framework—are approximated by sets of samples (Lieder, Griffiths, & Hsu, 2018; Sanborn et al., 2010; Vul et al., 2014). Such belief sampling accounts arise in models of physical reasoning (Battaglia et al., 2013; Ullman et al., 2017), category learning (E. B. Bonawitz & Griffiths, 2010; Goodman et al., 2008; Shi et al., 2010), sentence parsing (Levy et al., 2009), theory of mind (Baker et al., 2009), creative thinking (Smith et al., 2013), multiple object-tracking (Vul et al., 2009), and many more. In all these domains, inference is supported by a sampled set of beliefs about what sort of world we might be in given the available data: what rule governs category membership, what a particular sentence's meaning evaluates to, or what sort of physical outcome is likely to occur from a particular starting state.

The number of samples in belief sampling

Belief sampling accounts start with the assumption that sampling is done by an algorithm that approximates the relevant probability distribution without bias, at least in the limit of infinitely many samples (S. Gershman et al., 2015). This assumption is important for addressing the motivating challenge of biological and psychological plausibility of probabilistic inference noted previously. However, while the sampling algorithms underlying belief sampling models may make certain guarantees in the limit, when these algorithms are used to model human behavior, they often rely on only a few samples. After all, considering infinitely many samples seems just as implausible as working with the complex probability distribution directly. Further, relying on few samples ensures that a sample-based approximation to a probabilistic model generates novel predictions (otherwise, it perfectly mimics the probabilistic model with no sampling). The net result is that belief sampling accounts start with the assumption that few samples are used and aim to characterize just how small that number is.

One approach to characterizing the number of samples relies on the role of sampling variability in explaining differences between individual and group-level behavior (Dasgupta et al., 2017; Vul et al., 2014). In a range of settings, decision making across a group of participants or aggregated over many trials closely resembles complete probabilistic reasoning (i.e., based on infinite samples), yet individual or trial-by-trial results are often idiosyncratically variable (Goodman et al., 2008; Griffiths & Tenenbaum, 2006; Lewandowsky et al., 2009; Mozer et al., 2008). The tension between aggregate population behavior consistent with probabilistic inference and seemingly irrational individual behavior can be resolved by positing that individuals use very few samples to guide decisions. This produces high variance individual trial behavior that approximates full probabilistic inference over many trials. By modeling variation in individual behavior, researchers estimate how many samples people might be using and often find low numbers (Mozer et al., 2008). Although such reliance on few samples may seem surprising, in a broad set of decision tasks, under reasonable assumptions about the *cost* of samples, making quick decisions based on only one or a few samples is optimal (Vul et al., 2014).

As in data sampling models, considerations of sampling cost are critical for belief sampling. In some cases, generating belief samples can be highly burdensome, relying on the simulation of complex physical world models or generative processes. This, in combination with the effort to underscore the cognitive plausibility of belief sampling models, places the cost of samples front and center. The concept of *sample re-use* across inferences has therefore emerged as a relevant factor effecting the number of (novel) samples. When the cognitive costs of sampling are assumed to be large, the ability to remember previous samples offers a powerful opportunity to save time and computation across suitably similar decision contexts (Dasgupta & Gershman, 2021; Logan, 1988a).

First, sample reuse is obviously helpful when we might need to answer a more or less identical question again. For example, when asked to give an estimate for simple questions like, "What percent of the world's airports are in the United States?", people provided less correlated responses when they were asked the question again three weeks later than when they were asked immediately after the first response (Vul & Pashler, 2008), and the correlation between immediate repeated guesses was lower for people with lower memory capacity (Hourihan & Benjamin, 2010). Here, memory for the previous answer is assumed to bias the second response when people are prompted at close intervals, suggesting that people will re-use their initial sample in subsequent judgments. The broader question of how samples from memory align with the expected utility framework is something we address at greater length in the discussion.

A related body of work addresses the conditions under which people re-sample or continue to rely on a previous sample even as they receive *new data*. For example, across a number of category learning tasks, individual behavior can be fit by a particle filter algorithm with a single particle, a finding consistent with continuing to use a sampled hypothesis as long as it continues to fit the data (Sanborn et al., 2010). Similar work showed that a "win-stay, lose-sample" algorithm which retains an existing hypothesis and only re-samples when it fails to describe the data captures adult and toddler behavior in a causal learning task (E. Bonawitz, Denison, Gopnik, et al., 2014). Even more dramatically, Goodman et al. (2008) provide evidence that participants in sequential concept learning settings will continue to use a sampled rule even when other rules are more likely (or even fit the data perfectly). Together, these results suggest that people exhibit a strong tendency to re-use samples across repeated decisions, limiting the number of samples they need to draw when doing so is costly.

But how far does this tendency go? A growing body of work suggests that people will re-use costly samples to support multiple related decisions (Dasgupta & Gershman, 2021; Dasgupta et al., 2018; Dasgupta et al., 2020; S. Gershman & Goodman, 2014). For example, when people are asked to make judgments that can be supported by a previous inference due to conditional dependence, their responses can be strongly predicted by the previous response relative to people who did not make the previous inference first, and their response times are consistent with sample re-use (S. Gershman & Goodman, 2014). Though these results suggest that people make dynamic inferences about sample re-use in belief sampling, many questions remain about how people make such decisions and what kinds of limits or biases are introduced in the process.

The sample distribution in belief sampling

When sampling beliefs, the sample distribution reflects the challenges of generating representative samples from $P(s \mid x)$ in Equation 3 through purely cognitive processes. While sampling *data* is often a matter of exploring or receiving information from the world, sampling beliefs requires deploying machinery capable of imagining different possible worlds. In some cases, this may be as simple as imagining what kind of outcome we might receive from a particular gamble (Lieder, Griffiths, & Hsu, 2018) or what values a simple random variable in the environment might take (Vul et al., 2014). However, in many cases, this requires more sophisticated mental models or generative processes to produce the samples. To illustrate, consider “noisy physics engine” models of intuitive physics, in which inferences about whether a tower of blocks will fall or a ball will hit a target are based on forward simulation of a dynamic physics engine to produce samples of how different configurations of blocks might behave (Battaglia et al., 2013). Or, in category learning paradigms, generating a hypothesis about the rule that determines category membership (i.e., the belief that best describes the available data) may involve a rich generative process to produce sample rules (E. B. Bonawitz & Griffiths, 2010; Goodman et al., 2008) or draw on prior knowledge (Williams & Lombrozo, 2013).

In light of these challenges, bias in the sample distribution arises first from the difficulty of obtaining a sample at all. In such settings, sampling requires creative algorithmic solutions which simplify the process of generating a sample. However, as a result of these simplifications, such processes may not faithfully represent the underlying distribution with only a limited number of samples. One common approach relies on *Markov Chain Monte Carlo* (MCMC), in which each sample is generated through easily computable modifications to the *previous* sample. A sequence or "chain" of samples generated in this way has the property that in the limit, it approximates the underlying distribution with high fidelity (Gilks et al., 1995). MCMC is commonly used in machine learning applications to approximate complex distributions that cannot be analytically specified. However, because of the iterative sampling algorithm, MCMC samples have an autocorrelation that is more pronounced in small-sample regimes. The sample distribution may in turn be more homogeneous than would be expected, since each sample is correlated with the one that preceded it. If people implement a form of MCMC sampling in belief sampling, we might expect the autocorrelation of samples to have behavioral consequences; MCMC-like processes have been proposed as an account of perceptual switching in binocular rivalry (S. J. Gershman et al., 2012), sequential dependence in semantic memory search (Bourgin et al., 2014), and anchoring biases (Lieder, Griffiths, Huys, et al., 2018).

In a similar vein to MCMC, sequential Monte Carlo algorithms, or *particle filters*, provide a natural way of capturing *online* belief updating as new information comes in (Levy et al., 2009; Sanborn et al., 2010; Vul et al., 2009) rather than static inferences. These models maintain a set of sampled hypotheses that can change over time and be re-evaluated as new data is observed. For instance, the first few words of a sentence are consistent with many candidate parses, but subsequent words narrow the set of possibilities. A particle filter may sample such sentence parses and re-weight and re-sample plausible parses as new words are read, allowing for efficient, psychologically plausible approximation of the posterior distribution over parses (Levy et al., 2009). Similar dynamics have motivated the use of particle filters in other sequential tasks in humans (Sanborn et al., 2010; Vul et al., 2009), and animals (N. Daw & Courville, 2008). Notably, psychologically plausible particle filters only entertain a limited number of samples at any one time; as time goes on, these samples are "pruned" when they are inconsistent with the observed data. This may lead to scenarios where new data is entirely inconsistent with existing samples. In sentence parsing, such particle filter "collapse" has been proposed as an account of the experience of being stymied when parsing a garden path sentence such as, "the horse raced past the barn fell" (Levy et al., 2009).

The previous examples illustrate the potential for biases in the sample distribution driven by various algorithmic attempts to make sampling tractable in the first place. However, a second source of bias in the distribution of belief samples arises from attempts to generate the *right* samples. Just as data sampling may under-represent rare but high utility outcomes (Hertwig et al., 2004), belief sampling algorithms must account for states that have low probabilities but high-magnitude utilities, like winning the lottery or contracting a fatal illness. Such "black swan" states are unlikely to be sampled from the underlying distribution, but due to their high (positive or negative) utility, failure to consider them could lead to missed opportunities or highly undesirable outcomes. Empirically, these events are often given *disproportionate* attention and rated as more probable than they truly are (Kahneman & Tversky, 1979), suggesting that natural state sampling offers a poor psychological account of the ways in which people treat such events. How can a sample-based account of decision making—which would require taking potentially thousands of samples to assess particularly rare outcomes (Lieder, Griffiths, & Hsu, 2018)—address the readiness with which people consider possible black swan states when making decisions? As with the challenge of generating complex belief samples, answers to this question largely appeal to the sampling algorithm itself. Lieder, Griffiths, and Hsu (2018) propose that people use a form of *importance sampling* in which samples are drawn from a biased distribution that over-weights high-utility-variance states, and then corrects for this bias by weighting by the inverse of the utility variance. This algorithm ensures that extreme, albeit low probability, events are considered when making a decision. Further, the strategy avoids black swans in the ultimate decision, but causes people to overestimate the likelihood of these extreme, low probability events, thus offering an account of the availability bias (Kahneman & Tversky, 1979). Lieder, Griffiths, and Hsu (2018) show that this same strategy also accounts for other peculiarities of human decision making, including inconsistent risk preferences and certain memory biases. Broadly, these results illustrate that when using few samples and faced with the challenge of generating relevant samples at all, algorithms that solve this problem may introduce biases and distortions in the sample distribution which have downstream behavioral consequences.

Though research on the behavioral consequences of biased sample distributions has largely focused on adults, recent developmental work has highlighted the value of exploring similar questions in children. After all, young children, perhaps more than adults, are faced with a daunting task of assembling coherent representations of the state of the world given noisy and sparse data. A number of results suggest that children may indeed rely on sampling mechanisms similar to adults to update their beliefs about their environment (E. Bonawitz, Denison, Gopnik, et al., 2014; E. Bonawitz, Denison, Griffiths, et al., 2014; Denison et al., 2013). And just as biased sample distributions may lead to distinct biases in adult decision making, recent work has investigated whether patterns of sampling behavior might help explain key changes in cognitive development. For example, evidence that children are often more exploratory than adults and can sometimes learn novel relationships better than adults is consistent with children sampling a wider range of hypotheses and generalizing less (Gopnik et al., 2015; Gopnik et al., 2017; Lucas et al., 2014; Schulz et al., 2019). In this vein, developmental research offers a unique opportunity for further testing predictions of belief sampling models.

Action sampling

The decision making framework presented here posits that given a choice of possible actions, people will choose the one that maximizes expected utility. In the expected utility calculation in Equation 3, the process of choosing the best action is glossed over in a simple $\arg\max$ operator, but implementing it is a challenging optimization problem (Nocedal & Wright, 2006). In the controlled settings that dominate empirical work in psychology and decision making, people select among a pre-defined set of options (Kalis et al., 2013; Smaldino & Richerson, 2012), so the optimization problem is well-constrained; we can evaluate each of the few explicitly stated alternatives to find the best one. However, real-world situations present us with innumerably many possible actions to choose from. Our options about what to do for dinner include not only obvious choices like raiding the fridge or ordering take-out, but also novelties such as snaring a neighborhood squirrel, and non-sequiturs such as opening a pedicure salon. Intuitively, our deliberation process cannot evaluate *all* possible options, so the choice of best action necessarily involves choosing from a limited sample of alternatives, often called the *consideration set* (Hauser & Wernerfelt, 1990). This is illustrated by the example in Figure 1, where deciding what to wear to a party will in practice only involve explicitly sampling a handful of possible items. However, even remaining open to everything in our closet already represents a narrowing of the consideration set if we do not also consider buying a whole new outfit on the way to the party or fashioning new coverings from the contents of our kitchen pantry. In general, since we cannot consider all available actions when making decisions, sampling offers a potential solution.

Because our aim in considering actions is finding the best one, there are substantial differences in constraints on, and properties of, action sampling compared to belief or data sampling. On the one hand, optimization simplifies what we do with a sampled action: simply evaluate how good it is, and ultimately choose the best one. There is no need to re-sample an action multiple times to estimate frequency, or to weight it in some manner to correct for a biased sampling distribution. On the other hand, optimization underscores the need to sample actions that are likely to be valuable. With a potentially infinite, discontinuous set of possible actions, the number of terrible actions is vast, and if we sample a subset of actions from an irrelevant distribution, the best action in our consideration set might be quite bad indeed. In light of this, investigations into the number of samples and the sampling distribution for actions focuses on understanding how we manage to consider the right kinds of actions most of the time.

This description of decision making as choosing from among candidate (sampled) actions the one with the highest expected utility has been challenged by a diverse set of accounts. In particular, many alternatives emphasize the role of exogenous factors like potential regret (Loomes & Sugden, 1982) or disappointment (Loomes & Sugden, 1986), or affective responses more generally (Mellers et al., 1997) in decision making. Since these variables can be considered somewhat independent of an action's underlying utility, the extent to which they impact downstream decision making is important. Critically, the role that the number of action samples and the sample distribution play in decision making remains consistent with these theories. While our discussion focuses on choosing the action which maximizes approximate expected utility, we might just as easily replace this with the action that minimizes expected regret, or maximizes a more complex joint distribution of affect and utility. In either case, the motivating question remains how people sample the right actions from a practically infinite space of options.

The number of samples in action sampling

With belief and data sampling, the number of samples places an obvious constraint on our ability to make downstream inferences because the samples represent our estimated probability distributions $P(s \mid x)$. However, as Equation 3 illustrates, sampled actions do not play the same role; we simply choose the best from among them. Therefore, it could be the case that despite limiting the number of actions in the consideration set, this choice has no noteworthy consequences for behavior. Since we just pick the best action from the consideration set, so long as the best action is contained within, it doesn't matter how small that set is. In some cases, the first action that comes to mind is indeed likely to be the best. Johnson and Raab (2003) found that the first action considered by handball players faced with a possible game scenario was ultimately chosen 60% of the time and that in general, subsequently generated alternatives decreased in quality.

However, perfect calibration between the actions we consider and the utility function we aim to maximize not only seems implausible in theory, but is also not borne out in practice. First, ensuring that the best action is somewhere in the consideration set seems to require an omniscient procedure that already knows the outcome of the decision process; if we could choose a consideration set based on which actions are the best, we could just use that process to make our choice. Second, ensuring that the best answers are in the consideration set is not just *a priori* implausible. In many domains, people can identify the best action from among a set of alternatives, but fail to generate that action themselves (Adams et al., 2021). For instance, when proposing good questions in a modified battleship game, subjects could reliably identify the most informative questions when presented with a list, but rarely generated the optimal questions on their own (Rothe et al., 2018). This ability to recognize useful actions, despite failing to generate them, suggests a sub-optimality that arises not from action evaluation itself, but instead from relying on a subset of actions that may not always be calibrated to the problem at hand.

Given that the limited consideration set of actions may not always include the best one, the number of samples drawn plays a role in determining how likely the consideration set is to contain promising choices. As with data and belief sampling, additional samples will incur time costs (and thus a reduction in the rate of decisions) and are likely to also impose cognitive effort. In some cases, the effort required to evaluate sampled actions can be quite large. The familiar experience of *choice overload*, in which having more options can paradoxically lead to higher levels of regret, choice deferral, or dissatisfaction, makes these costs clear (Chernev et al., 2015).[5] The benefits of sampling more actions depend in part on the variation in payoffs across sampled alternatives; if all actions are equally beneficial, nothing is lost from poor optimization by considering too few samples. In addition, the value of additional sampled actions depends on the alignment between the sampling distribution and the true context-relevant payoffs. For sampled actions to be useful, they must be sampled according to a plausible approximation of their present expected utility, such as prior success in similar settings. However, the *actual* utility or reward obtained from a given action in the present context may not match the expectation. If the sample distribution is less likely to provide the best action for the given decision, then having more samples to choose from gives one better odds of choosing a high value action.

These considerations suggest that the expected utility calculation based on sampled actions will be sensitive to how many samples are taken and whether these samples are well calibrated to the environment. Indeed, in simulations testing these predictions, Morris et al. (2021) show that the expected reward rate varies as a function of the number of actions considered, as well as the correlation between action utilities and the probability that actions are sampled into the consideration set. But how many action samples should we take? With a perfect correlation, the first action considered will most often be the optimal action, so one sample yields the highest possible reward. However, as this correlation decreases, the consideration set will need to be larger to maintain a high probability of containing a good action; consequently, the optimal number of samples is

[5] While these results might seem counter to the claim that decision makers simply choose the best action from the consideration set, they primarily suggest that evaluating actions is indeed difficult and costly, so evaluating more choices can increase error or uncertainty in our choice. Factors that often lead to greater choice overload are exactly those we intuitively associate with making errors more likely, e.g., choice set complexity and preference uncertainty (Chernev et al., 2015).

larger. Morris et al. (2021) find that so long as the correlation between the distribution over actions and the true utilities is larger than 0.25, the optimal number of actions considered out of a possible 1000 is less than 10. These simulation results therefore suggest that in scenarios where our expectations map reasonably well onto the present decision, as with data and belief sampling, there may be little value in taking more than a few samples.

The sample distribution in action sampling

The expected utility maximization goal given by Equation 3 is to select the best action from the set of possible actions, but given that the space of actions is immense, the $\arg\max$ must be evaluated over a subset of actions. This subset represents the sample distribution over actions. In the previous sections on data and belief sampling, we discussed the ways in which biased sample distributions predict systematic deviations from optimal behavior. However, the potential for such deviations in action sampling is complicated by the fact that given a consideration set of actions, the decision maker need only pick the best one. If the globally best action is in the consideration set, any other bias in the sample distribution is essentially irrelevant. Thus, just as the number of samples in the consideration set only effects behavior if the best action is not guaranteed to be in the sample, the sample distribution will only bias behavior if the optimal choice is not present. This poses a challenge for researchers testing the role of samples in action selection: how to identify scenarios where people's consideration set will reliably fail to include the optimal choice? Intuitively, this depends on how the candidate actions are sampled.

Morris et al. (2021) propose that we consider actions based primarily on global average utilities that are not dependent on the local decision context. These actions are then sampled into the consideration set, and each one is evaluated more systematically given all that is known about the current context. On this account, a consideration set is constructed where the probability of including an option is proportional to its "cached" historical value, perhaps pulled from long-term memory (Kaiser et al., 2013), or obtained via model free reinforcement learning (Pezzulo et al., 2013). Critically, such a process might introduce a bias toward actions that have previously been selected or rewarded but which may be incongruous in novel circumstances. In new situations, if no prior circumstances are particularly relevant, we are left to consider only actions that have been previously successful on average. This predicts that people tend to disproportionately favor globally useful actions in novel circumstances. In support of this account, Morris et al. (2021) found that when asked to list all meal ideas following dental surgery, subjects tended to generate options that were high in general value (i.e., their favorite dishes), but were less likely to generate options of high utility in the specific context (i.e., suitable given dental restrictions). This finding is attributed to a process which samples broadly "useful" actions but, faced with the challenge of evaluating them in a relatively unfamiliar context (after dental surgery), does little additional filtering or optimization.

In contrast, Phillips et al. (2019) propose that we use the current context to determine which actions were frequently chosen, or had yielded high utility, in prior similar circumstances. These actions then form the basis of our sampled consideration set. On this account, the consideration set is more tailored to a given decision context, making it more likely that the optimal action is in the sample. Though intuitive, it can be difficult to extract behavioral predictions from this account. For one, the mechanism for identifying similar past situations in a given decision context is highly unconstrained; any past experience, behavioral tendency, observed contingency, or accrued reward could plausibly be used as input into the system that learns reward estimates, and any representation of the current context might be used to query this cache for good actions. Second, on this account, it is unclear when biases in the consideration set will emerge since the sample is predicted to be more context specific. Scenarios in which the context *seems* to invite a particular sample of actions as a useful consideration set but *in fact* are best approached with a different action require somewhat unique circumstances where previously learned adaptations are no longer useful or are incidentally low utility despite similar circumstances. If we walk out to somebody else's car with them and instinctively reach into our pocket for our own keys, this sort of action has the character of being a utility maximizing behavior under similar circumstances but one which, evaluated in the current setting, is not nearly as useful. Such paradigms are often used to study the tension between reflexive, *model-free* decisions and more deliberative, *model-based* action selection (Gläscher et al., 2010). On this view, model-free choices reflect the sort of behaviors that past experience alone predicts would be utility maximizing—actions selected from a potentially biased sample distribution without further consideration—while model-based action selection involves the more careful evaluation of all items in the consideration set and leads to different choices in the context at hand. In this way, action sampling might predict when the decision system as a whole will behave consistently with model-free habits or model-based plans, according to the interplay of the cost of action sampling and the correlation between the action sampling distribution and the underlying utility function.

Discussion

In this chapter, we have attempted to explicitly relate a broad array of sampling models in the literature as facets of an over-arching computation: approximating expected utility maximization. We argued that existing sample-based accounts in a diversity of settings can be succinctly described as sampling the underlying *data*, *belief*, and *action* distributions that are central to calculating expected utility. In this way, these models come together to support a unified process of making good decisions.

Previously, data, belief, and action sampling have largely been considered in isolation in tasks designed to make all but one of these components trivially simple. We begin with the premise that in real-world settings, all three aspects are non-trivial, and therefore an integrative account of decision making by sampling ought to consider them all. By jointly evaluating how samples of each type are used together, we see that subsets of data, belief, and actions play a fundamentally different role in the expected utility calculation: conditioning for data, marginalization for beliefs, and optimization for actions. These differences highlight discrepancies in the solutions and broader research goals in each domain. However, the present work suggests that there is room for optimism about a more integrated view. In support of this, we discuss major trends in these distinct literatures along two key lines of analysis which are critical to any mature sample-based theory: a precise account of how decision making in each setting reflects the *number of samples* and the *sample distribution*. We take as our starting point that in data, belief, and action sampling, decision making relies on only a few samples and that in each domain, the sample distribution that results from these limited samples must play a role in downstream behavior. This perspective provides a number of broad insights across these domains. Data sampling allows for a precise account of the number of samples, belief sampling models show a critical relationship between the sampling algorithm and behavior, and action sampling poses a challenge in specifying the relationship between the sample distribution or consideration set and decision making.

The differences between data, belief, and action sampling algorithms highlighted by our review offer a clear opportunity for synthesis. Because each field has emphasized different aspects of sampling, they have made progress on issues that may have been ignored by the other fields. In this vein, considering the different classes of sampling together provides a basis for development in one field to be translated to others. The formal tools for deriving optimal stopping rules in drift-diffusion models of data sampling may be fruitfully brought to bear on the meta-cognitive choices pertaining to belief sampling and might even inform the trade-off between implicit and deliberative action selection. Likewise, correlated sampling algorithms from belief sampling highlight methods for grappling with non-independent data sampling environments. Progress in both of these domains may offer a number of future directions for relatively nascent models of action selection. The synthesis offered in this chapter leaves open two large classes of questions which we briefly consider. First, the expected utility maximization framework for sampling models should be taken to its logical conclusion: how do people integrate data, belief, and action sampling when making complex decisions? Second, while this chapter attempts to bring a breadth of sample-based models into the framework of expected utility maximization, this is far from a complete survey. How might we consider still other sampling accounts, such as sampling utilities or memories?

Unifying sample-based expectation maximization

The motivation for unifying all the facets of sampling under one account of approximating expected utility maximization is that real-world tasks entail ambiguity, uncertainty, and approximation of every aspect of the calculation. Throughout this chapter, we have alluded to a simplistic example of deciding what to wear to a party which illustrates this (Figure 1). However, everyday life presents us with an abundance of similar scenarios. Critically, explicitly combining sampling of all of these variables creates at least two major complications: (a) determining which variables are sampled in what order, and (b) optimizing sample sizes given sample-based error along all variables.

First, the basic expression $\arg\max_A \sum_s U(A;s)P(s)$ entails evaluating the utility of every combination of sampled action and sampled state, meaning that the processing load grows as a product of sample sizes. Alternatively, perhaps not all combinations of action and state need to be evaluated; but then how can the system determine which combinations to consider? Equivalently, if we consider jointly sampling $(action, state)$ tuples, we face the same problem: what dependence structure ought to exist in the joint distribution? Another alternative is that we sample sequentially. Just as we condition on data to sample states, perhaps we condition on states to sample actions. Indeed, this would be one method to allow better generalization from past experience when determining the action sampling distribution. However, this conditional action sampling procedure does not find the action with the highest expected utility, but instead finds the action with the highest *optimistic* utility: the action that has the highest utility in the state for which it is well suited. In short, because action and belief sampling accounts typically work in isolation, little is known about how people ought to set the order and dependency structure when sampling both, and nothing is known about how they actually do this.

Second, the number of samples we ought to take for data, beliefs, or actions are inextricably linked. In the limit, where we sample only one action, it is irrelevant how much data we have to inform our latent state estimates, or how precisely we approximate our beliefs about states, since the solitary sampled action determines our choice. Likewise, if we sample no data, nothing is to be gained from sampling more than one action. If we do not inform our estimates about the current state, then our state-contingent utility calculation will not differ from a global average utility from which we might sample actions. However, not all samples are complements; additional data samples decrease the entropy of our belief distribution, but it is advantageous to take more belief samples in medium entropy scenarios (Vul et al., 2014). This means that the optimal number of belief samples does not change monotonically with the number of data samples; it initially rises when we obtain some data, then drops as data sufficiently constrains our beliefs. Given these interdependencies between sample sizes, the joint optimization over the number of data, belief, and action samples might yield several discrete classes of algorithms, depending on the balance of data, belief, and action sample costs.

Other forms of sampling

The framework proposed in this chapter emphasized data, state, and action sampling. However, two other interrelated forms of sampling have not been considered: sampling utilities (Stewart et al., 2006) and memories (Nosofsky & Palmeri, 1997). First, consider utility sampling. In the formalism in Equation 3, there is no uncertainty or error in utility evaluations: $U(a; s)$ is accessible directly and infallibly. However, some relaxations of this assumption can be accommodated in our formulation. An uncertain or stochastic utility function can be captured without explicitly sampling utilities by adding more state possibilities. For instance, instead of $U(a; s)$ yielding \$10 50% of the time, and \$0 the remainder of the time, we can posit a constant utility function defined over two new states: $U(a; u) = \$10$ and $U(a; v) = \$0$, and we split the probability associated with state s evenly across the two sub-states u and v. However, other challenges to the assumption of an available value function cannot be so addressed, and would instead require a new sampling term. For instance, if we do not have direct access to explicit value functions (Hayden & Niv, 2020), but must instead reconstruct them from past experience (Stewart et al., 2006), this reconstruction must have some sampling fidelity, and must be nested within state-action sampling. Such an approach is most relevant when we consider that value functions are likely to be learned over time, via reinforcement learning. Indeed, the Thompson Sampling (Thompson, 1933) strategy for striking a balance between exploitation and exploration in multi-armed bandit problems is to choose the action with the highest *sampled* payoff, suggesting that there is an important algorithmic role for sampling utility functions themselves.

Sampling from memory is a more complicated case, because memory sampling takes on many different flavors. The most basic example of memory sampling is conjuring past observations from memory—this is consistent with data sampling, but where the information search process proceeds through one's memory stores (Hills et al., 2012; Ratcliff, 1978). This form of memory sampling, although conceptually quite different from sampling information from the external world, is still consistent with sampling *data*; the sampled memories are conditioned on and yield noteworthy predictions insofar as they come from a biased distribution. Most memories however are not pure retrieval of facts but are instead *reconstructions* of the past (Bartlett, 1932). In that sense, a sampled memory lies somewhere between a sample of data, and a sample of the *beliefs*, i.e., the inferred world state. Finally, memory also serves to cache previous actions and calculations, as in the instance theory of automatization (Logan, 1988b), which seems most consistent with sampling *actions* based on previously calculated optimal choices. These different modes of memory sampling are sometimes quite tangled. For instance, sampling exemplars from memory is consistent with sampling beliefs in a categorization probability (Nosofsky & Palmeri, 1997; Shi et al., 2010), and more generally, sampling previously considered beliefs or interim calculations from memory (Dasgupta & Gershman, 2021; Dasgupta et al., 2018) is consistent with reusing past samples of beliefs, actions, or data. Altogether, when considered in our framework, sampling memory is a broad collection of strategies for reusing prior samples of either data, beliefs, or actions. In short, although utility and memory sampling may be accommodated within our framework, a serious treatment of either of these domains requires considerable elaboration of our framework, and thus offers a promising avenue for future work.

Conclusions

In summary, this chapter addresses the challenge of integrating the many sample-based models across diverse areas of psychology. We begin with the observation that in decision making contexts, a rational agent chooses the action that maximizes expected utility. Yet doing so poses major computational hurdles even for simple, everyday decisions. In light of this, we propose that several broad classes of sampling models in the literature can be viewed as approximating different components of the expected utility calculation: *data*, *beliefs*, and *actions*. Under this unifying framework, our review compares models of data, belief, and action sampling along two distinct dimensions which are central to any sampling account: the number of samples and the sample distribution. This comparison offers novel insights into the strengths of different sampling accounts as well as opportunities for future work. Finally, we advocate for further inquiry aimed at understanding how data, belief, and action sampling processes come together to support rich, grounded decision making.

References

Adams, G. S., Converse, B. A., Hales, A. H., & Klotz, L. E. (2021). People systematically overlook subtractive changes. *Nature*, *592*(7853), 258–261.

Baker, C. L., Saxe, R., & Tenenbaum, J. B. (2009). Action understanding as inverse planning. *Cognition*, *113*(3), 329–349.

Bartlett, F. C. (1932). *Remembering: A study in experimental and social psychology.* Cambridge University Press.

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.

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.

Bonawitz, E., Denison, S., Gopnik, A., & Griffiths, T. L. (2014). Win-stay, lose-sample: A simple sequential algorithm for approximating bayesian inference. *Cognitive psychology*, *74*, 35–65.

Bonawitz, E. B., & Griffiths, T. L. (2010). Deconfounding hypothesis generation and evaluation in bayesian models. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *32*(32).

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.

Bourgin, D., Abbott, J., Griffiths, T., Smith, K., & Vul, E. (2014). Empirical evidence for markov chain monte carlo in memory search. *Proceedings of the Annual Meeting of the Cognitive Science Society*, *36*(36).

Braddick, O. (1974). A short-range process in apparent motion. *Vision research*, *14*(7), 519–527.

Chater, N., Tenenbaum, J. B., & Yuille, A. (2006). Probabilistic models of cognition: Where next? *Trends in Cognitive Sciences*, *10*(7), 292–293.

Chernev, A., Böckenholt, U., & Goodman, J. (2015). Choice overload: A conceptual review and meta-analysis. *Journal of Consumer Psychology*, *25*(2), 333–358.

Coenen, A., & Gureckis, T. (2021). The distorting effects of deciding to stop sampling information.

Dasgupta, I., & Gershman, S. J. (2021). Memory as a computational resource. *Trends in Cognitive Sciences*.

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

Dasgupta, I., Schulz, E., Goodman, N. D., & Gershman, S. J. (2018). Remembrance of inferences past: Amortization in human hypothesis generation. *Cognition*, *178*, 67–81.

Dasgupta, I., Schulz, E., Tenenbaum, J. B., & Gershman, S. J. (2020). A theory of learning to infer. *Psychological review*, *127*(3), 412.

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

Daw, N. D., O'doherty, J. P., Dayan, P., Seymour, B., & Dolan, R. J. (2006). Cortical substrates for exploratory decisions in humans. *Nature*, *441*(7095), 876–879.

Dawes, R. M. (1993). Prediction of the future versus an understanding of the past: A basic asymmetry. *The American Journal of Psychology*, 1–24.

Denison, S., Bonawitz, E., Gopnik, A., & Griffiths, T. L. (2013). Rational variability in children's causal inferences: The sampling hypothesis. *Cognition*, *126*(2), 285–300.

Doucet, A., de Freitas, N., & Gordon, N. (2001). Sequential monte carlo methods in practice.

Fechner, G. T. (1860). *Elemente der psychophysik* (Vol. 2). Breitkopf u. Härtel.

Fiedler, K. (2000). Beware of samples! a cognitive-ecological sampling approach to judgment biases. *Psychological review*, *107*(4), 659.

Fiedler, K. (2008). The ultimate sampling dilemma in experience-based decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *34*(1), 186.

Fiedler, K., & Juslin, P. (2006). Taking the interface between mind and environment seriously. In K. Fiedler & P. Juslin (Eds.), *Information sampling and adaptive cognition* (pp. 3–32). Cambridge University Press.

Galesic, M., Olsson, H., & Rieskamp, J. (2012). Social sampling explains apparent biases in judgments of social environments. *Psychological Science*, *23*(12), 1515–1523.

Galesic, M., Olsson, H., & Rieskamp, J. (2018). A sampling model of social judgment. *Psychological review*, *125*(3), 363.

Gershman, S., & Goodman, N. (2014). Amortized inference in probabilistic reasoning. *Proceedings of the annual meeting of the cognitive science society*, *36*(36).

Gershman, S., Horvitz, E. J., & Tenenbaum, J. B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, *349*(6245), 273–278.

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

Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological review*, *103*(4), 650.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve bayesian reasoning without instruction: Frequency formats. *Psychological review*, *102*(4), 684.

Gilks, W. R., Richardson, S., & Spiegelhalter, D. (1995). *Markov chain monte carlo in practice*. CRC press.

Gittins, J. C. (1979). Bandit processes and dynamic allocation indices. *Journal of the Royal Statistical Society: Series B (Methodological)*, *41*(2), 148–164.

Gläscher, J., Daw, N., Dayan, P., & O'Doherty, J. P. (2010). States versus rewards: Dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*, *66*(4), 585–595.

Goodman, N. D., Tenenbaum, J. B., Feldman, J., & Griffiths, T. L. (2008). A rational analysis of rule-based concept learning. *Cognitive science*, *32*(1), 108–154.

Gopnik, A., Griffiths, T. L., & Lucas, C. G. (2015). When younger learners can be better (or at least more open-minded) than older ones. *Current Directions in Psychological Science*, *24*(2), 87–92.

Gopnik, A., O'Grady, S., Lucas, C. G., Griffiths, T. L., Wente, A., Bridgers, S., Aboody, R., Fung, H., & Dahl, R. E. (2017). Changes in cognitive flexibility and hypothesis search across human life history from childhood to adolescence to adulthood. *Proceedings of the National Academy of Sciences*, *114*(30), 7892–7899.

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.

Hau, R., Pleskac, T. J., & Hertwig, R. (2010). Decisions from experience and statistical probabilities: Why they trigger different choices than a priori probabilities. *Journal of Behavioral Decision Making*, *23*(1), 48–68.

Hauser, J. R., & Wernerfelt, B. (1990). An evaluation cost model of consideration sets. *Journal of consumer research*, *16*(4), 393–408.

Hayden, B., & Niv, Y. (2020). The case against economic values in the brain.

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological science*, *15*, 534–539.

Hertwig, R., & Erev, I. (2009). The description–experience gap in risky choice. *Trends in cognitive sciences*, *13*, 517–523.

Hertwig, R., & Pleskac, T. J. (2010). Decisions from experience: Why small samples? *Cognition*, *115*, 225–237.

Hills, T. T., Jones, M. N., & Todd, P. M. (2012). Optimal foraging in semantic memory. *Psychological review*, *119*(2), 431.

Hogarth, R. M., Lejarraga, T., & Soyer, E. (2015). The two settings of kind and wicked learning environments. *Current Directions in Psychological Science*, *24*(5), 379–385.

Hourihan, K. L., & Benjamin, A. S. (2010). Smaller is better (when sampling from the crowd within): Low memory-span individuals benefit more from multiple opportunities for estimation. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *36*(4), 1068.

Jazayeri, M., & Movshon, J. A. (2007). A new perceptual illusion reveals mechanisms of sensory decoding. *Nature*, *446*(7138), 912–915.

Johnson, J. G., & Raab, M. (2003). Take the first: Option-generation and resulting choices. *Organizational behavior and human decision processes*, *91*(2), 215–229.

Jones, M., & Love, B. C. (2011). Bayesian fundamentalism or enlightenment? on the explanatory status and theoretical contributions of bayesian models of cognition. *Behavioral and brain sciences*, *34*(4), 169.

Juni, M. Z., Gureckis, T. M., & Maloney, L. T. (2016). Information sampling behavior with explicit sampling costs. *Decision*, *3*, 147.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–292.

Kaiser, S., Simon, J. J., Kalis, A., Schweizer, S., Tobler, P. N., & Mojzisch, A. (2013). The cognitive and neural basis of option generation and subsequent choice. *Cognitive, Affective, & Behavioral Neuroscience*, *13*(4), 814–829.

Kalis, A., Kaiser, S., & Mojzisch, A. (2013). Why we should talk about option generation in decision-making research. *Frontiers in psychology*, *4*, 555.

Karnopp, D. C. (1963). Random search techniques for optimization problems. *Automatica*, *1*(2-3), 111–121.

Klayman, J., & Ha, Y.-W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological review*, *94*(2), 211.

Knill, D. C., & Richards, W. (1996). *Perception as bayesian inference*. Cambridge University Press.

Konovalova, E., & Le Mens, G. (2020). An information sampling explanation for the in-group heterogeneity effect. *Psychological review*, *127*(1), 47.

Kwisthout, J., Wareham, T., & van Rooij, I. (2011). Bayesian intractability is not an ailment that approximation can cure. *Cognitive Science*, *35*(5), 779–784.

Levy, R. P., Reali, F., & Griffiths, T. L. (2009). Modeling the effects of memory on human online sentence processing with particle filters. *Advances in neural information processing systems*, 937–944.

Lewandowsky, S., Griffiths, T. L., & Kalish, M. L. (2009). The wisdom of individuals: Exploring people's knowledge about everyday events using iterated learning. *Cognitive science*, *33*(6), 969–998.

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.

Lieder, F., Griffiths, T. L., Huys, Q. J., & Goodman, N. D. (2018). The anchoring bias reflects rational use of cognitive resources. *Psychonomic bulletin & review*, *25*(1), 322–349.

Logan, G. D. (1988a). Toward an instance theory of automatization. *Psychological review*, *95*(4), 492.

Logan, G. D. (1988b). Toward an instance theory of automatization. *Psychological review*, *95*(4), 492.

Loomes, G., & Sugden, R. (1982). Regret theory: An alternative theory of rational choice under uncertainty. *The economic journal*, *92*(368), 805–824.

Loomes, G., & Sugden, R. (1986). Disappointment and dynamic consistency in choice under uncertainty. *The Review of Economic Studies*, *53*(2), 271–282.

Lucas, C. G., Bridgers, S., Griffiths, T. L., & Gopnik, A. (2014). When children are better (or at least more open-minded) learners than adults: Developmental differences in learning the forms of causal relationships. *Cognition*, *131*(2), 284–299.

Mellers, B. A., Schwartz, A., Ho, K., & Ritov, I. (1997). Decision affect theory: Emotional reactions to the outcomes of risky options. *Psychological Science*, *8*(6), 423–429.

Morris, A., Phillips, J., Huang, K., & Cushman, F. (2021). Generating options and choosing between them depend on distinct forms of value representation. *Psychological Science*, *32*(11), 1731–1746.

Mozer, M. C., Pashler, H., & Homaei, H. (2008). Optimal predictions in everyday cognition: The wisdom of individuals or crowds? *Cognitive science*, *32*(7), 1133–1147.

Navarro, D. J., & Perfors, A. F. (2011). Hypothesis generation, sparse categories, and the positive test strategy. *Psychological review*, *118*(1), 120.

Nocedal, J., & Wright, S. (2006). *Numerical optimization*. Springer Science & Business Media.

Nosofsky, R. M., & Palmeri, T. J. (1997). An exemplar-based random walk model of speeded classification. *Psychological review*, *104*(2), 266.

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, *101*(4), 608.

Oaksford, M., Chater, N. et al. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford University Press.

Pezzulo, G., Rigoli, F., & Chersi, F. (2013). The mixed instrumental controller: Using value of information to combine habitual choice and mental simulation. *Frontiers in psychology*, *4*, 92.

Phillips, J., Morris, A., & Cushman, F. (2019). How we know what not to think. *Trends in cognitive sciences*, *23*, 1026–1040.

Ratcliff, R. (1978). A theory of memory retrieval. *Psychological review*, *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.

Ratcliff, R., & Smith, P. L. (2004). A comparison of sequential sampling models for two-choice reaction time. *Psychological review*, *111*(2), 333.

Robert, C. P., & Casella, G. (2004). *Monte carlo statistical methods* (Vol. 2). Springer.

Rothe, A., Lake, B. M., & Gureckis, T. M. (2018). Do people ask good questions? *Computational Brain & Behavior*, *1*, 69–89.

Sanborn, A. N. (2017). Types of approximation for probabilistic cognition: Sampling and variational. *Brain and cognition*, *112*, 98–101.

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.

Savage, L. J. (1954). *The foundations of statistics*. Wiley.

Schulz, E., Wu, C. M., Ruggeri, A., & Meder, B. (2019). Searching for rewards like a child means less generalization and more directed exploration. *Psychological science*, *30*(11), 1561–1572.

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.

Smaldino, P. E., & Richerson, P. J. (2012). The origins of options. *Frontiers in Neuroscience*, *6*, 50.

Smith, K., Huber, D. E., & Vul, E. (2013). Multiply-constrained semantic search in the remote associates test. *Cognition*, *128*(1), 64–75.

Stewart, N., Chater, N., & Brown, G. D. (2006). Decision by sampling. *Cognitive psychology*, *53*, 1–26.

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: Statistics, structure, and abstraction. *science*, *331*(6022), 1279–1285.

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.

Tversky, A., & Kahneman, D. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, *47*(2), 263–291.

Ullman, T. D., Spelke, E., Battaglia, P., & Tenenbaum, J. B. (2017). Mind games: Game engines as an architecture for intuitive physics. *Trends in cognitive sciences*, *21*(9), 649–665.

Von Neumann, J., & Morgenstern, O. (1947). *Theory of games and economic behavior*. Princeton university press.

Voss, A., Nagler, M., & Lerche, V. (2013). Diffusion models in experimental psychology: A practical introduction. *Experimental psychology*, *60*(6), 385.

Vul, E., Goodman, N., Griffiths, T. L., & Tenenbaum, J. B. (2014). One and done? optimal decisions from very few samples. *Cognitive science*, *38*, 599–637.

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

Vul, E., Frank, M. C., Tenenbaum, J. B., & Alvarez, G. A. (2009). Explaining human multiple object tracking as resource-constrained approximate inference in a dynamic probabilistic model. *Advances in neural information processing systems*.

Wald, A., & Wolfowitz, J. (1948). Optimum character of the sequential probability ratio test. *The Annals of Mathematical Statistics*, 326–339.

Watkins, C. J., & Dayan, P. (1992). Q-learning. *Machine learning*, *8*(3-4), 279–292.

Williams, J. J., & Lombrozo, T. (2013). Explanation and prior knowledge interact to guide learning. *Cognitive psychology*, *66*(1), 55–84.