

# The ABC of Heuristics: Rules of Thumb as Likelihood-free Approximate Bayesian Computation

Tianwei Gong<sup>1\*</sup>, Simon Valentin<sup>2\*</sup>, Neil R. Bramley<sup>3</sup> and Christopher G. Lucas<sup>1†</sup>

<sup>1</sup>Department of Experimental Psychology, University College London.

<sup>2</sup>School of Informatics, University of Edinburgh.

<sup>3</sup>Department of Psychology, University of Edinburgh.

\*Corresponding author(s). E-mail(s): [t-gong@ucl.ac.uk](mailto:t-gong@ucl.ac.uk);  
[s.valentin@ed.ac.uk](mailto:s.valentin@ed.ac.uk);

Contributing authors: [neil.bramley@ed.ac.uk](mailto:neil.bramley@ed.ac.uk); [c.lucas@ed.ac.uk](mailto:c.lucas@ed.ac.uk);

†These authors contributed equally to this work: Tianwei Gong and Simon Valentin.

## Abstract

Heuristics have long been viewed as simple decision rules that sacrifice optimality for cognitive efficiency. However, this perspective fails to fully account for their effectiveness in complex, real-world environments. This paper proposes a reinterpretation of heuristics as realizations of a universal algorithm for likelihood-free approximation to Bayesian inference, bridging the gap between heuristic and normative approaches to decision-making. We argue that many heuristic choice patterns are the consequence of a robust and adaptive inference process that involves simulating from one’s prior and using low-dimensional summary statistics (cues) to compare with observed data. This process sidesteps the typically-intractable Bayesian likelihood calculation, while still approximating the optimum conditional on the effectiveness of chosen summary statistics. We illustrate how this allows heuristics to be situated within a probabilistic framework via this process. Our proposal addresses several key puzzles in the field: It relaxes strong assumptions about cue-target relationships in heuristic models, accounts for human stochasticity in heuristic decision-making, and invites a more appropriate benchmark for evaluating heuristics’ performance. We provide two illustrative examples — a classic cue-based decision task and an ecological causal learning task, to demonstrate the practical applicability of our framework.

By unifying heuristics and Bayesian inference through the lens of likelihood-free methods, this work provides a more nuanced understanding of human inference and decision-making.

**Keywords:** Approximate Bayesian computation; Heuristics; likelihood-free inference; rejection sampling; bounded rationality.

Heuristics are defined as “rules of thumb”, or “mental shortcuts”, that allow people to make timely inferences or advantageous choices in familiar settings without expending excessive cognitive resources (Simon, 1945). For the past half century, a long and seemingly ever-growing list of heuristics has been proposed. These heuristics have typically been held as explanations for patterns of human behavior in tasks that resemble familiar kinds of real-world inference or decision making. A common theme is that they reduce effort and work surprisingly well in some settings, while they can also be pushed to failure by careful experimental manipulations (Gigerenzer & Todd, 1999; Shah & Oppenheimer, 2008; Simon, 1945; Tversky & Kahneman, 1974). Human empirical data have been aligned with both the successes and failures of heuristics, showing evidence that many proposed heuristics have explanatory power that goes beyond that of the normative models they are held up as alternatives to.

Shortcuts, or cognitively economical inference strategies, are arguably essential; reality often requires us to solve complex problems in real time (Christiansen & Chater, 2016) while relying on incomplete information (Oaksford & Chater, 2007). This has long been a challenge to rational, normative characterizations of those problems, which are typically very computationally expensive (Jones & Love, 2011; Marr, 1982). In fact, for many everyday problems that we solve heuristically, optimal solutions are not merely costly to compute but intractable in practice (Simon, 1945; Van Rooij, 2008).<sup>1</sup>

Even though the theory of heuristics provides a good explanation for the human ability to make rapid and reasonably accurate inferences, it says little about another hallmark of human cognition: flexibility. People have a remarkable ability to adapt their strategies when existing methods prove ineffective. When pushed far enough, people can abandon a malfunctioning heuristic and find a better approach (Newell, Weston, & Shanks, 2003). The heuristics literature, nonetheless, often focuses on fixed, inflexible strategies for making inferences and decisions. At the same time, this adaptiveness is not fully explained by a rational Bayesian perspective. Proposals about optimal strategy selection via rational meta-reasoning (Lieder & Griffiths, 2017, 2020) only go so far because they tend to assume that developmental or evolutionary-scale process produce strategies and an apparatus for understanding when to apply which one. This leaves us with a similar static toolkit to what is suggested by the heuristics

---

<sup>1</sup>The concept of intractability is central to our framework. We will engage with it in more detail later, but briefly: in computational theory and Bayesian inference, a problem is considered “intractable” when finding an exact solution is practically impossible within reasonable time or resource constraints. This can occur due to exponential growth in computational requirements as the problem size increases. In Bayesian contexts, exact calculation of conditional probability distributions is often infeasible due to high-dimensional parameter spaces, difficulties in evaluating likelihood functions, and the unavailability of a closed-form mathematical expression of a generative process. Among them, the difficulty in evaluating likelihoods is the focus of this paper.

literature, and seems at once reasonable and incomplete when it comes to the human ability to give snap answers to unfamiliar questions.

In this paper, we propose a new interpretation of heuristics by highlighting a deep commonality between the abstraction to simple indirect cues — the hallmark of heuristics literature — and the mathematical principles of likelihood-free inference. In particular, we focus on how, beyond the “lamplight” of small-world tasks and closed-form models, effective inferences about different complex domains can be flexibly made by comparing observations with the outputs of mental simulations based on abstracted simple, low-dimensional cues. This mathematical approach is known as Approximate Bayesian Computation (ABC), and we will address how the use of heuristics in the cognitive literature aligns with it. Ours is not the first attempt to unify the literature on heuristics with notions of approximate rationality (cf. [Binz, Gershman, Schulz, & Endres, 2022](#); [Dasgupta, Schulz, Tenenbaum, & Gershman, 2020](#); [Lieder & Griffiths, 2017](#)). However, we specifically advance the perspective that heuristics can be viewed as a form of likelihood-free inference, and that the ABC framework can be used to reinterpret the process of employing heuristics. We argue that the heuristics-as-likelihood-free-inference perspective offers a unique degree of generality and unlocks new insights into the functioning of the mind.

To foreshadow the advantages in using ABC to reinterpret heuristics, we highlight three challenges to previous research into heuristics that potentially undermine their power to explain human behavior: First, heuristic decision-making and inference — as typically conceived — relies on strong assumptions about the relation between cues (observable features used to make judgments) and targets (outcomes or variables being predicted). This limits their applicability to real-world problems, as it is unclear how the validity of these assumptions can be established and monitored in dynamic, complex environments where relationships may shift. Second, many standard accounts of heuristic decision-making and inference do not explicitly model either the systematic stochasticity at the aggregate level or uncertainty at the individual level. This stands in contrast to evidence showing systematic patterns of stochasticity in human responses, and people’s explicit and implicit consideration of uncertainty when combining cues and making risky decisions. The nature of this uncertainty representation in heuristic reasoning remains subject to ongoing debate ([Sanborn & Chater, 2016](#); [Szollosi, Donkin, & Newell, 2023](#)). Third, the “optimal” benchmarks to which many heuristics have traditionally been compared — weighted attribute averaging or linear regression — provide only a limited standard. A fairer evaluation would compare heuristics against models that are plausibly Bayes-optimal and grounded in reasonable priors.

Our perspective will preserve the central idea of heuristics while providing a means to address these challenges. Specifically, we will show heuristics are not independent rules, but particular *summary statistics* that, when combined with an ABC procedure, enable solutions to difficult or intractable inference problems, that can boundedly rational if the summary statistics are selected and used appropriately. The strong assumption about cue-target relationships can then be relaxed by viewing heuristics as flexible approximations to complex Bayesian computations. Besides, this offers an account of implicit uncertainty representation via the probabilistic framework of inference-by-sampling. As we will show later, sampling procedures within ABC can

naturally capture the variability of human responses, without the need to impose an additional decision rule (e.g., the softmax rule). Finally, by understanding heuristics under the approximate Bayesian approach, it becomes easier and more straightforward to use rational Bayesian inference as the benchmark for comparison. By reconceptualizing heuristics as likelihood-free approximations to Bayesian inference, we aim to provide a more holistic framework for understanding human decision-making.

## Heuristics via the Lens of Cognitive Economy

### From shortcuts to strategies

Heuristics, defined as the “shortcuts” to more elaborate solutions (Simon et al., 1990), can lead people to make imperfect inferences or decisions. Tversky and Kahneman’s work on “heuristics and biases” documented cases in which human decision making is suboptimal but consistent with simple heuristics (Kahneman, 2011; Kahneman & Tversky, 1973; Tversky & Kahneman, 1974). This includes examples such as favoring options that are more easily brought to mind (*availability heuristic*), more recently encountered (*recency bias*), or closer to an arbitrary anchor value (*anchoring*), categorizing someone based on how closely their description matches a category’s salient features, while overlooking other possibilities (*representativeness heuristic*). When people are asked to make decisions in artificial scenarios based on a bounded amount of newly learned evidence, they are found to combine multiple sources of evidence, or cues, with apparent neglect to their relative strengths (*Dawe’s rule*, Dawes & Corrigan, 1974) or by completely neglecting all but the strongest one cue (*take-the-best*, Gigerenzer, Hoffrage, & Kleinbölting, 1991). Experimental manipulations of option sets can lead followers of such heuristics to systematically sub-optimal judgments and decisions, a fact that a generation of marketing, business, bookmaking and political groups have exploited (John, Caldwell, McCoy, & Braganza, 2024; Newall, 2019; Thaler & Sunstein, 2009).

While the first wave of heuristics research primarily highlighted the potential pitfalls arising from humans’ reliance on heuristics, it also revealed how deeply such reasoning patterns are embedded in human cognition. As pointed out by Herbert Simon, heuristics serve to provide satisfactory solutions in the face of limited resources (Simon et al., 1990). Accordingly, the second wave of research emphasized that heuristics, when applied in their natural contexts, can be “fast and frugal” and “make us smart” (Gigerenzer & Goldstein, 1996; Gigerenzer & Todd, 1999). Behind these broad arguments were several more specific claims about heuristics. For example, they encode ecologically appropriate inductive biases (Brighton & Gigerenzer, 2012; Gigerenzer & Brighton, 2009; Todd & Gigerenzer, 2012) and embody forms of expert knowledge by distilling what truly matters and ignoring irrelevant or redundant information (Martignon, Vitouch, Takezawa, & Forster, 2003). At its zenith, research sometimes argued that heuristics could actually produce more accurate and more robust decisions than more complex, complete statistical inference benchmarks (Goldstein & Gigerenzer, 2008). For instance, the “take-the-best” heuristic, which makes decisions based solely on the most discriminative cue, was argued to outperform regression in predicting real-world features such as which of two cities is larger (Gigerenzer

& Goldstein, 1996). Adopting an  $1/N$  distribution of investments (allocating  $1/N$  of wealth to each of the  $N$  available assets) was argued to match or outperform more sophisticated investment strategies on average (DeMiguel, Garlappi, & Uppal, 2009; Hertwig & Hoffrage, 2013). Subsequent re-analyses have revealed that these “less-is-more” effects are due to either heuristics introducing good biases, i.e., encoding prior information unavailable to the statistical comparison model, or avoiding over-fitting to noise (Parpart, Jones, & Love, 2018). However, the legacy of the fast and frugal heuristics program has been a field-wide recognition of the adaptive value, and indeed the necessity, of approximation strategies and inductive biases, now often discussed under the headings of learning to learn, meta-learning, or amortization (essentially forms of pre-computation and reuse) strategies (Binz et al., 2023; Dasgupta et al., 2020; Kemp, Goodman, & Tenenbaum, 2010).

## Cues (attribute substitution) in heuristics

At the core of heuristic decision-making is the concept of “cues” — features that guide our inferences, judgments and choices. Importantly, our use of these cues often involves what is called “attribute substitution”, assessing an easier or more salient attribute in place of a more difficult target one (John et al., 2024; Kahneman, Frederick, et al., 2002): Rather than using raw sensory inputs, these cues often represent abstractions from our perceptual experiences, distilling complex information into more manageable forms.

Consider how we navigate our everyday environments. When choosing a restaurant in an unfamiliar city, we don’t process every detail about each establishment. Instead, we might focus on a few key cues: how busy the restaurant is and its exterior appearance. Each of these cues is a coarse-grained reflection of a wealth of underlying information. Similarly, when judging the quality of a yogurt brand, we might use its price as a cue. The price, in this case, serves as a summary of a collection of attributes, potentially encoding information about ingredients, production processes, and market positioning. Or when predicting which team will win a baseball match, a powerful approach has proven to be the calculation of the batters’ run creation average:  $(\text{Hits} + \text{Walks}) \times \text{Total Bases}$  (Lewis, 2004), as a cue that effectively summarizes a complex history of performances into a single, easily digestible statistic.

The way we abstract and use cues can vary widely depending on the context and our prior experiences. Sometimes, we primarily use a single, highly predictive cue (as in the “take-the-best” heuristic). Other times, we might combine multiple cues, either equally (as in Dawes’ rule; Dawes & Corrigan, 1974) or with simplified weighting schemes (Zhang, 2004). Notably, the way we apply these heuristic features outside of toy cases may rarely be a matter of a monotonic “more is better”. We might pick whatever wine falls *closest* to the price we had anticipated paying, go to the restaurant that is busy but not *too* busy, or predict the winning team to be the one whose characteristics are, overall, closer to those we imagine a winning team having (e.g. strikeout rates for pitchers or swing aggressiveness for batters). This non-monotonic relationship between cues and decisions distinguishes sophisticated heuristics from simple rule-based decision-making. The complexity of cue usage in real-world scenarios necessitates strategies for reducing cognitive effort, which we will explore next.

## Effort reduction

Given their effort-reducing nature (Shah & Oppenheimer, 2008), cue-based heuristics offer several advantages over alternative approaches to inference and inductive reasoning. They promote efficiency and adaptability in three main areas: (1) computational complexity: by focusing on a subset of available information or using simplified rules, heuristics can dramatically decrease the time required to reach a decision; (2) memory requirements: given the notorious limitations of human working memory (Cowan, 2010; Miller, 1956), heuristics provide a way to make effective use of available information while only storing or processing a subset of it; and (3) solution feasibility: heuristics are particularly valuable in enabling decision-making in situations where comprehensive approaches fail entirely, providing a way to *satisfice* — to navigate complex problems and arrive at satisfactory, though not provably optimal, solutions (Simon, 1945).

In the simplest cases, discrete-choice decisions may require nothing more than choosing options that maximize a heuristic cue. This direct mapping from features to decisions occurs without additional computational steps, contrary to the monotonic value-based processing often assumed in the heuristics and biases literature. This approach aligns with the well-established understanding that humans are selective in what they track and attend to (Chun, Golomb, & Turk-Browne, 2011). Such selectivity is not a limitation, but a feature of our cognitive systems, allowing us to function effectively in complex, information-rich environments. Returning to our earlier examples, one might immediately select the busier-looking restaurant for dinner, using the crowd size as a direct proxy for quality without further deliberation, or instinctively choose the more expensive yogurt brand, assuming price directly indexes quality.

By exploiting regularities of the environment, heuristics thus provide “shortcuts” to more elaborate reasoning and exhaustive consideration of alternatives. However, this efficiency can come at the cost of accuracy when there is a mismatch between the assumptions driving the heuristic and the environment one is actually in. The effectiveness of a heuristic cue ultimately depends on how reliably it correlates with the property we’re interested in across different contexts. Later, we will explore what makes certain heuristic cues more robust and universally applicable than others, providing insights into both the persistence of existing heuristics and the development of new ones.

Independent of commitments to how heuristics reduce effort, their definition as mechanisms of effort reduction implies the existence of an alternative, more effortful, normative benchmark. This benchmark typically involves more comprehensive information processing and complex computations, often based on principles of rationality or optimality. Understanding this benchmark is crucial for evaluating the effectiveness and limitations of heuristics in various decision-making contexts. We next look more closely at what that benchmark should be and how it relates to our understanding of heuristic decision-making.

# Bayesian Models as the Normative Benchmark

We will begin by discussing models that have been held up as higher-effort or full-information alternatives to heuristics in the past. We follow this with a brief summary of Bayesian inference as a standard for rational or optimal inferences and decisions, with comparisons to those earlier benchmarks, before turning to the challenges in approximating Bayesian inferences in practice.

## Past benchmarks: weighted additive rules

In evaluating heuristics — both in establishing their accuracy and their match with human behavior — it is crucial to compare them against appropriate normative benchmarks. Historically, “full-information” models like weighted additive decision (WAD) rules have often been used for this purpose (Shah & Oppenheimer, 2008).

$$V(i) = \sum_{n=1}^N w_n v_{i,n}, \quad (1)$$

where  $V(i)$  is the overall value of alternative  $i$ ,  $N$  is the number of cues,  $w_n$  is the weight of cue  $n$ , and  $v_{i,n}$  is the value of cue  $n$  for alternative  $i$ .

Despite their historical use, WAD models fall short of capturing the complexities of optimal inference. Given their assumptions of linearity, independence, and the absence of prior information, they fail to account for cases where the ground truth includes complex dependencies or non-linear relationships, or where prior knowledge would make a difference. From this perspective, WAD models are better regarded as another class of heuristics rather than as normative standards. In contrast, fully Bayesian models — when made sufficiently expressive — are, in theory, able to capture dependencies and non-linear patterns, and to incorporate appropriate priors for the problem domain. They thus provide a more suitable normative benchmark for rational reasoning and decision-making under uncertainty (e.g., Anderson & Matessa, 1992; Chater & Oaksford, 2008; Colombo & Seriès, 2012; Friston, 2010; Gong, Pacer, Griffiths, & Bramley, in press; Gopnik & Tenenbaum, 2007; Griffiths & Tenenbaum, 2006; Jaynes, 2003; Kemp, Perfors, & Tenenbaum, 2007; Körding & Wolpert, 2004; Lee, 2011; Lucas & Griffiths, 2010; Ray, King-Casas, Montague, & Dayan, 2008; Socher et al., 2009; Ullman & Tenenbaum, 2020; Xu & Tenenbaum, 2007; Zhao, Lucas, & Bramley, 2022).

## Bayesian inference and its intractability

Bayesian accounts offer a principled description of computational problems that involve integrating prior knowledge with new evidence, providing a formal framework for optimal inference under uncertainty. In Bayesian inference, the posterior distribution  $p(\boldsymbol{\theta}|\mathbf{y})$  of model parameters  $\boldsymbol{\theta}$  given observed data  $\mathbf{y}$  is proportional to the product of the likelihood  $p(\mathbf{y}|\boldsymbol{\theta})$  and the prior  $p(\boldsymbol{\theta})$ , normalized by the evidence  $p(\mathbf{y})$ :

$$p(\boldsymbol{\theta}|\mathbf{y}) = \frac{p(\mathbf{y}|\boldsymbol{\theta})p(\boldsymbol{\theta})}{p(\mathbf{y})} \quad (2)$$



Despite its simple mathematical form and its capability for optimal inference, a direct application of Bayesian inference can often be challenging. To illustrate it, consider an example that encapsulates many of the computational difficulties faced in everyday reasoning: Identifying a distant moving light. Let  $\theta$  represent the identity of the light source (e.g., airplane, helicopter, cyclist, satellite, flashlight) and  $\mathbf{y}$  represent the light’s appearance over time. The Bayesian inference task is to compute  $P(\theta|\mathbf{y})$ :

$$P(\theta|\mathbf{y}) = \frac{p(\theta)p(\mathbf{y}|\theta)}{\sum_{\theta' \in \Theta} p(\theta')p(\mathbf{y}|\theta')} \quad (3)$$

The process involves two major challenges:

1. **Intractable Marginal Likelihood (Denominator):** The denominator involves summing over all possible light sources  $\theta$ , which could be enormous or even infinite.

2. **Intractable Conditional Likelihood:** The conditional likelihood  $p(\mathbf{y}|\theta)$  itself is often intractable, depending on numerous nuisance parameters  $\boldsymbol{\eta}$ :

$$p(\mathbf{y}|\theta) = \int p(\mathbf{y}|\boldsymbol{\eta}, \theta)p(\boldsymbol{\eta}|\theta)d\boldsymbol{\eta} \quad (4)$$

These parameters might include the object’s direction, distance, plausible paths, and many other factors — numerous dimensions that cannot be easily factorized. The intractability of this likelihood is the key motivation for likelihood-free inference methods, which perform inference without explicitly computing these complex likelihoods.

The intractability problem renders the full Bayesian inference impractical for real-time decision-making in complex, real-world scenarios. The computational demands far exceed human cognitive capacities (Lieder & Griffiths, 2020; Van Rooij, Blokpoel, Kwisthout, & Wareham, 2019), and can quickly reach the limits of computational tractability altogether. As such, it stands to reason that both human cognition and artificial systems must rely on efficient approximations or heuristics. However, understanding these approximations in relation to the full Bayesian ideal allows us to characterize how they trade off optimality for computational efficiency. Heuristics, in this light, can be viewed as computationally efficient, likelihood-free approximations to Bayesian inference, tailored to specific environmental structures and computational constraints.

## Approximate Bayesian Computation (ABC)

Likelihood-free inference (LFI) addresses a class of problems where direct evaluation of the likelihood function is not feasible, but we have a generative model of the domain. The key innovation of LFI is to bypass likelihood evaluation by working directly with simulations from the model. LFI had early roots in population genetics (Tavaré, Balding, Griffiths, & Donnelly, 1997) and quickly gained popularity in other scientific fields. Recent years have seen a surge of interest in likelihood-free inference methods with the development of increasingly sophisticated techniques (for



reviews, see [Cranmer, Brehmer, & Louppe, 2020](#); [Lintusaari, Gutmann, Dutta, Kaski, & Corander, 2017](#)) such as sequential Monte Carlo likelihood-free methods ([Sisson, Fan, & Tanaka, 2007](#)), synthetic likelihood methods ([Wood, 2010](#)) and neural density estimation ([Papamakarios, Sterratt, & Murray, 2019](#)).

A particularly prominent LFI method we focus on here is Approximate Bayesian Computation (ABC), which implements the broader concept of inference by comparison (see Section 26.2 in [Murphy, 2023](#)). Intuitively, ABC can be thought of as a “guess and check” approach to inference. Its simplest form, rejection ABC, repeatedly samples possible parameter values from a prior distribution, simulates data using these guesses, and checks how well the simulated data matches the observed data. If the match is *close enough*, the guessed parameters are accepted as plausible explanations for the observed data. This process is repeated many times to build up a collection of plausible (i.e., accepted) parameter values, which approximate the posterior distribution. The “closeness” of the match is typically judged using a distance function over summary statistics of the data rather than the raw data itself, making the comparison more efficient and often more meaningful. By leveraging this simulation-based approach, ABC opens up new possibilities for cognitive modeling, particularly in scenarios involving complex, high-dimensional processes that resist traditional analytical approaches or approximate methods that require access to a likelihood function.

---

**Algorithm 1** ABC Rejection Sampling with Summary Statistics

---

```

1: procedure ABCREJECTION( $p(\theta)$ ,  $p(y|\theta)$ ,  $y_{obs}$ ,  $s$ ,  $d$ ,  $\epsilon$ ,  $N$ )
2:    $S \leftarrow \emptyset$ 
3:    $s_{obs} \leftarrow s(y_{obs})$ 
4:   while  $|S| < N$  do
5:      $\theta^* \sim p(\theta)$ 
6:      $y^* \sim p(y|\theta^*)$ 
7:      $s^* \leftarrow s(y^*)$ 
8:     if  $d(s^*, s_{obs}) \leq \epsilon$  then
9:        $S \leftarrow S \cup \{\theta^*\}$ 
10:    end if
11:  end while
12:  return  $S$ 
13: end procedure

```

---

## Rejection sampling

A more formal understanding of rejection ABC is provided by the pseudocode of this approach in Algorithm 1, which contains the key elements of the choice of summary statistic function  $s(\cdot)$ , distance metric  $d$ , the sample size  $N$ , and tolerance threshold  $\epsilon$ .

We illustrate ABC using an example in Figure 1. Suppose an observer is considering three hypotheses to explain a moving light in the sky: a spotlight, a helicopter, and an unidentified flying object (UFO). These hypotheses differ in their prior probabilities

(Figure 1). For instance, the prior probability of a UFO might be low, while that of a spotlight could be relatively high, especially if assuming there was a forecast for a light show that night. To perform inference, a data point is simulated from each hypothesis, with sampling frequency proportional to the prior. If the distance between the summary statistics of the simulated data and the observed data falls below a tolerance threshold  $\epsilon$ , the sample is accepted; otherwise, it is rejected. This process continues until  $N$  samples (set to 10 in the figure) are accepted. The final decision is made by selecting the hypothesis that appears most frequently among the accepted samples. As  $\epsilon \rightarrow 0$ , the ABC posterior approaches the true posterior distribution. The approximation is asymptotically exact if the summary statistics are sufficient, but non-sufficient statistics can still yield good approximations in practice (Blum, Nunes, Prangle, & Sisson, 2013; Joyce & Marjoram, 2008; Lintusaari et al., 2017; Murphy, 2023; Sunnåker et al., 2013). However, the acceptance rate decreases as  $\epsilon$  does, making inference more costly. Conversely, as  $\epsilon \rightarrow \infty$ , the ABC posterior approaches the prior distribution, since all simulated data points are accepted regardless of their similarity to the observed data.

We focus on rejection ABC due to its simplicity and interpretability, without committing to rejection sampling as a process-level model. While alternative ABC methods, including MCMC-ABC (Murphy, 2023) can offer better efficiency or accuracy and align with ongoing research into “inference by sampling” (Sanborn, Griffiths, & Navarro, 2010; Zhu, Sanborn, & Chater, 2020), we leave a detailed reconciliation of that research and our ideas to future work; the simple rejection sampling procedure provides new insights into how decisions are made with a minimum of auxiliary hypotheses.

In using rejection ABC to predict responses, we assume that people make decisions by choosing the best or most-likely option, without adding any variability to that already created by the sampling process. For discrete options, this amounts to picking the most frequent  $\theta$  in  $S$ . Given different tolerance thresholds  $\epsilon$  and total sample sizes  $N$ , the procedure can capture a spectrum of responses — ranging from falling back to the prior, to behavior that looks like probability matching, to maximizing under the true posterior.

## Summary statistics

Summary statistics are an essential part of ABC and central to our account of how heuristics are usefully understood in terms of ABC. They provide a way to compress complex, high-dimensional data into lower-dimensional representations that ideally retain the most relevant information for the inference task at hand. In our example of identifying a moving light in the darkness, the summary statistics could include the light’s average warmth (hue) and the average speed of its movement (Figure 1). Formally, a summary statistic  $\mathbf{s}(\mathbf{y})$  is a function of observed or simulated data  $\mathbf{y}$ , occupying a lower-dimensional space. If  $\mathbf{y} \in \mathcal{Y}$  is the observed data, then  $\mathbf{s} : \mathcal{Y} \rightarrow \mathcal{S}$  where typically  $\dim(\mathcal{S}) < \dim(\mathcal{Y})$ .

A key concept in the theory of summary statistics is sufficiency. A statistic  $\mathbf{s}(\mathbf{y})$  is sufficient for parameter  $\theta$  if it captures all relevant information about  $\theta$  contained in

$\mathbf{y}$ . From an information theoretic perspective, if  $\mathbf{s}(\mathbf{y})$  is sufficient for  $\boldsymbol{\theta}$ , then:

$$I(\mathbf{s}(\mathbf{y}); \boldsymbol{\theta}) = I(\mathbf{y}; \boldsymbol{\theta}) \quad (5)$$

where  $I(\cdot; \cdot)$  denotes mutual information. In Bayesian inference, using sufficient statistics simplifies the computation of the posterior distribution. This means we can perform Bayesian inference using only the sufficient statistic  $\mathbf{s}(\mathbf{y})$  instead of the full data  $\mathbf{y}$ , without any loss of information:

$$p(\boldsymbol{\theta}|\mathbf{y}) = p(\boldsymbol{\theta}|\mathbf{s}(\mathbf{y})) \propto p(\mathbf{s}(\mathbf{y})|\boldsymbol{\theta})p(\boldsymbol{\theta}) \quad (6)$$

While sufficient statistics are well understood for many simple parametric distributions (for example, we are all familiar with the mean and standard deviation as sufficient statistics for a Gaussian distribution), finding strictly sufficient statistics for complex simulator models is difficult, and low-dimensional sufficient statistics need not exist for any particular model. In practice, ABC often works with summary statistics that are not strictly sufficient but capture enough of the relevant information for the modeler’s purposes. For instance, it often assumes that the behavior of some phenomenon is *close enough* to Gaussian that we can summarize it with a mean and a standard deviation.

Instead of comparing observed data  $\mathbf{y}$  directly with simulated data  $\mathbf{y}^{\text{sim}}$ , ABC compares the summary statistics of observed data to the summary statistics of simulated data:

$$d(\mathbf{s}(\mathbf{y}^{\text{sim}}), \mathbf{s}(\mathbf{y})) \leq \epsilon \quad (7)$$

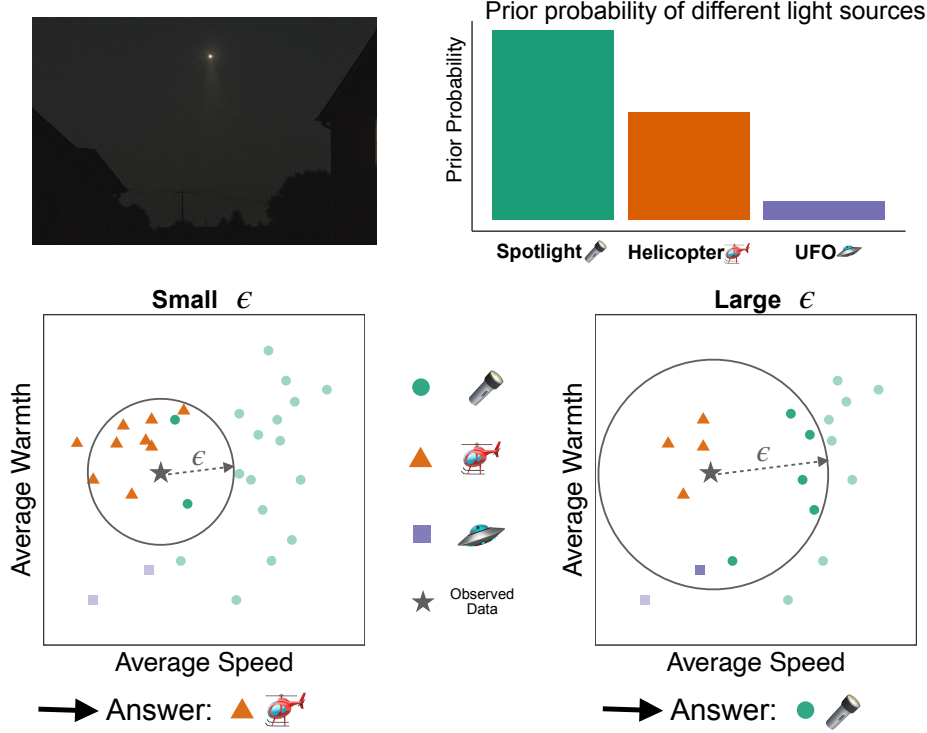
where  $d$  is a distance function (e.g. the Euclidean distance) and  $\epsilon$  is a tolerance threshold. Given the importance and challenge of choosing good summary statistics, there has been growing interest in methods for learning summary statistics from data. These approaches often use neural networks (Chen, Zhang, Gutmann, Courville, & Zhu, 2020) or other flexible function approximators to learn mappings from the data to lower-dimensional representations that are informative for the inference task.

## Heuristics as Approximate Bayesian Computation

While the theoretical developments of ABC and human heuristics were independent, there are intriguing similarities in their approaches to complex inference problems. We review recent cognitive research suggesting at least four connections between the two fields: (1) the motivation for resource-rational approximation, (2) the reliance on sampling-based processes, (3) the ability to simulate observations, and (4) the similarity between cues (attribute substitution) and summary statistics.

### Resource rationality

Resource rationality (Bhui, Lai, & Gershman, 2021; Lieder & Griffiths, 2020) is a current framework in cognitive science that offers a new view for understanding heuristics within the broader context of Bayesian cognition. Moving one step beyond earlier



**Fig. 1:** A light source inference example for rejection ABC. Three hypothetical sources (Spotlight, Helicopter, UFO) differ in their priors. The reasoner compares the summary statistics (average warmth and average speed of the light) between the observation (the gray star) and data simulated based on the generative models and hypothesis priors. Samples that fall within the tolerance threshold ( $\epsilon$ ) are accepted until the pre-set sample size  $N=10$ . When  $\epsilon$  is small, more simulated data points are needed, and the final answer tends to reflect the ground truth (Helicopter). When  $\epsilon$  is large, the final answer tends to reflect the priors.

approaches such as bounded rationality (Simon, 1955) and fast and frugal heuristics (Gigerenzer & Todd, 1999), resource rationality explicitly formalizes the trade-off between accuracy and computational cost. It posits that a rational agent optimizes the total benefit after considering both the utility of accurate inferences and the costs of computation (Griffiths, Lieder, & Goodman, 2015). Resource rationality differs from its predecessors in its emphasis on deriving process-level models from first principles. While bounded rationality and fast and frugal heuristics often propose computationally efficient strategies and then justify their effectiveness post hoc, resource rationality starts by formalizing the total loss function, including both accuracy and computational costs (Lieder, Griffiths, & Hsu, 2018). This formalization can then be used to derive specific cognitive processes or heuristics that are optimal under given resource constraints and environmental structures (Gershman, Horvitz, & Tenenbaum, 2015).

From this viewpoint, heuristics are supposed to be seen as potentially resource-rational solutions to the challenges posed by real-world complexity and the computational constraints of human cognition. Although some research has attempted to reframe heuristics, such as the anchoring effect, within the resource-rational framework (Lieder, Griffiths, M. Huys, & Goodman, 2018), the relationship between heuristics based on attribute substitution and fast, approximately normative inferences has not previously been given a systematic, formal foundation. Connecting heuristics to ABC can clarify how computational costs are quantitatively reduced and the price one pays in faithfulness to exact posterior distributions, measured by the sample size in the sampling process and the trade-off between the quantity and quality of the summary statistics used.

## Decision by sampling

The ABC approach, as well as other approximate techniques (e.g., Monte Carlo methods) used when the Bayesian prior is intractable, relies on generating samples to estimate probabilities or make decisions. This aligns with the observation that human judgments often exhibit variability consistent with sampling-based processes (Bramley & Xu, 2023; Daw & Courville, 2008; Griffiths, Vul, & Sanborn, 2012; Sanborn & Chater, 2016; Vul, Goodman, Griffiths, & Tenenbaum, 2014; Zhu, Sundh, Spicer, Chater, & Sanborn, 2023), especially when assuming a small sample size (Vul et al., 2014). The success of sampling-based models extends across various cognitive domains, from categorization to causal reasoning, and shows compatibility with neural data (Haefner, Berkes, & Fiser, 2016). Additionally, sequential models like drift diffusion models (DDMs) align with the idea that decision accuracy improves with evidence accumulation over time, interpretable as generating more samples from an internal probability distribution (Bitzer, Park, Blankenburg, & Kiebel, 2014; Ratcliff, Smith, Brown, & McKoon, 2016).

## Mental simulation

The ABC approach relies on a generative model to sample potential observations, which brings us to another relevant line of cognitive research: mental simulation. Mental simulation has often been proposed as a key mechanism in human reasoning, particularly in domains like intuitive physics. When reasoning about physical systems, it has been shown that humans may simulate the behavior of objects in their minds without explicitly computing probabilities (Battaglia, Hamrick, & Tenenbaum, 2013; Bramley, Gerstenberg, Tenenbaum, & Gureckis, 2018; Gerstenberg, Goodman, Lagnado, & Tenenbaum, 2021; Hamrick, Battaglia, Griffiths, & Tenenbaum, 2016; Ullman, Stuhlmüller, Goodman, & Tenenbaum, 2018). This can be understood as a form of sampling trajectories from an intuitive model, where the mind generates plausible trajectories or outcomes based on its knowledge of principles and constraints in a given domain (Wang et al., 2024). By running multiple simulations and aggregating the results, people might make probabilistic inferences and predictions about complex systems without explicitly computing likelihoods. Recent studies demonstrate simulation’s potential to drive inference wherever we have imagination-supporting domain

knowledge, including the familiar Newtonian dynamics of the physical world but also familiar social situations (Baker, Jara-Ettinger, Saxe, & Tenenbaum, 2017; Gerstenberg & Tenenbaum, 2017; Ho, Saxe, & Cushman, 2022). Mental simulation provides a route to inference that does not require explicit computation or representation of likelihoods. Instead, it relies on a generic capacity to generate concrete realizations of a model of the world, which directly links to the likelihood-free idea in ABC.

### Cues and summary statistics in cognition

As discussed in the “cues (attribute substitution) in heuristics” section, cues play a central role in accounts of heuristic reasoning. After walking through the framework of rejection ABC, we can see the idea of cues, or attribute substitution, satisfy the definition of summary statistics, as both articulate the reduction of evidence by abstracting the useful information. There is ongoing debate about the level of granularity at which cues can and should be understood. The notion of natural assessments (Tversky & Kahneman, 1983) suggests that people often treat combinations of basic information as single units due to frequent exposure to these information patterns. For instance, when judging a restaurant’s quality, “ambiance” might be considered a single cue, despite combining multiple sensory inputs and prior experiences. Shah and Oppenheimer (2008) point out the ambiguity in determining whether a heuristic that considers only one cue is actually a natural assessment or a combination of multiple cues. This ambiguity contributes to the challenge of defining “basic” cues and the risk of reducing cognitive processes to absurdly granular levels. The parallel between how humans process environmental cues and how statistical methods reduce high-dimensional data to summary statistics is exactly what we are pointing out here.

## Applying ABC to Heuristic Paradigms

We here provide two practical examples to illustrate how the ABC approach can be applied to tasks that study heuristics. Notably, the purpose of this paper is to provide a new interpretation of heuristics rather than a new approach to outperform them. As such, we do not expect our modeling results to surpass those proposed in previous literature. Instead, we focus on how the approach can (1) align with previous predictions of heuristics, (2) explain the systematic stochasticity patterns in human judgment that are missing from traditional heuristic accounts, (3) approximate the normative Bayesian benchmark, while (4) better capturing human performance than the normative Bayesian benchmark. We use the first task to illustrate points (1), (2), and (3), and the second task to illustrate points (3) and (4).

The first example is a classic binary comparison task, in which the decision-maker chooses between two alternatives based on a set of cues. This has been the core task in the literature on classic heuristic research (Bergert & Nosofsky, 2007; Gigerenzer et al., 1991; Lee & Cummins, 2004; Mata, von Helversen, & Rieskamp, 2010; Newell et al., 2003; Parpart et al., 2018; Rieskamp & Otto, 2006; Todd & Gigerenzer, 2012), especially for demonstrating the take-the-best and tallying heuristics (see below).

The second example is an ecological causal learning task, in which learners infer the causal structure between events based on how the events unfold over time. This represents a current line of research that investigates how people can learn about or interact with the complex world within limited time. It extends the simple decision-making task by focusing more on heuristics in situations where people must extract useful information from rich perceptual observations in complex everyday tasks including causal learning (Gong & Bramley, 2023; Gong, Gerstenberg, Mayrhofer, & Bramley, 2023; Rehder, Davis, & Bramley, 2022), physical learning (Bramley et al., 2018; Ullman et al., 2018), decision making (Callaway et al., 2022; Song, Bnaya, & Ma, 2019), and multi-step planning (Van Opheusden et al., 2023). Similar to traditional decision-making paradigms, numerous domain-specific heuristics have been proposed in these studies. However, because researchers tend to highlight the flexibility of human cognition, individual heuristics often fail to capture human performance. Nevertheless, combinations of cues have frequently been shown to outperform the Bayesian benchmark in accounting for human behavior (Gong & Bramley, 2023; Song et al., 2019; Van Opheusden et al., 2023). Here, we do not aim to develop a domain-general method for combining heuristics, but rather to show that any combination of cues can be systematically nested within the ABC approach, which may yield predictions that better capture human judgments than the Bayesian benchmark in similar ways to non-ABC cue combination methods used in previous papers.

Code of the modeling process in this paper can be found at <https://github.com/tianweigong/ABCHeuristics>

## Binary comparison

Take-the-best is one of the most well-known heuristics that have been proposed. It applies to situations where a decision maker needs to choose an option from a consideration set. In research settings, it is often simplified into a binary comparison procedure, where the decision-maker chooses the better of two alternatives (e.g., which city has a larger population) based on a set of binary cues (e.g., whether the city has an underground system; whether it has a university). The importance of cues can be provided directly or learned through a training procedure in which examples are accompanied by feedback indicating which option is better. This importance is typically quantified as cue validity, which measures how often a correct decision is made when the decision maker relies solely on the state of the cue (i.e., one stimulus has the cue, and the other does not). Cue validity is calculated as:  $\frac{N_{\text{correct}}}{N_{\text{total}}}$ , where  $N_{\text{correct}}$  is the number of correct decisions made using the cue, and  $N_{\text{total}}$  is the total number of decisions for which the cue discriminated between the options. A Bayesian version was later proposed to account for the influence of base rates (i.e., how often the cue is present in the training examples), in which cue validity is calculated as  $\frac{N_{\text{correct}}+1}{N_{\text{total}}+2}$  (Lee & Cummins, 2004). Nevertheless, two calculations yield the same validity ranking of cues as long as each cue occurs a similar number of times.

As a “one-reason” decision strategy, *take-the-best* relies on the single most valid cue and ignores the rest. If this cue discriminates between the options, the favored one is chosen. If not, the next most valid cue is considered, and so on, until a decision is made. *Tallying* demonstrates another heuristic approach that uses additive decision



rule. It chooses the option supported by a larger number of cues, treating each cue as equally important. Formally, both heuristics consider positive cues (where presence indicates a better option) and negative cues (where absence indicates a better option), with the calculation rule flipped for negative cues. However, many studies focus only on positive cues, or even just the most recognizable ones, and make random choices when cues do not discriminate, which is better aligned with the theoretical assumption that heuristics should remain simple and cognitively bounded.

Both take-the-best and tallying have been framed as deterministic decision rules in heuristics literature, meaning they do not account for the randomness in participants' decisions. We will here aim to replicate the previous phenomena by framing them as summary statistics under the ABC approach, meanwhile providing a way to explain the randomness in responses.

### *The Task*

We perform simulations based on the task in [Lee and Cummins \(2004\)](#) and [Bergert and Nosofsky \(2007\)](#), who used the same stimuli with different cover stories. Using [Bergert and Nosofsky \(2007\)](#) for demonstration here, the task asks to decide which of two insects is more poisonous, based on six binary body features (body, eyes, legs, antennae, fangs, and tail). The training phase provided feedback on pairwise comparisons of 16 insects, whose rankings were predetermined by the experimenters. The testing phase asked to choose between pairs of novel insects.

This task setup is ideal for our illustration here because (1) given the artificial context, numerous environments can be created by randomly swapping the rankings of the 16 training objects; and (2) all participants go through the same training phase, ensuring they receive the same amount of evidence, which is unlike other tasks that require actively searching for information (e.g., [Newell et al., 2003](#)).

We later created 100 environments by swapping the rankings of 16 training objects from [Lee and Cummins \(2004\)](#) and [Bergert and Nosofsky \(2007\)](#). In the training phase, all 120 pairwise comparisons of 16 training objects were shown to the observers. In the testing phase, observers were asked to make judgments on the remaining 1896 possible pairs ( $2^6 - 120$ ). As in the original research, we focus on environments in which all cues are positive (i.e., the presence of a cue indicates a better option, *ceteris paribus*).<sup>2</sup>

We focus on two key phenomena identified in previous research. First, the heuristics take-the-best and tallying often yield correct decisions across a wide range of situations. While many earlier simulations used real-world environments — where the correct answer is determined by actual ground truths — we demonstrate that these heuristics also perform well in artificial environments. Specifically, when Bayesian rationality is used as the normative benchmark, both the original heuristic algorithms and the ABC versions show strong performance as the approximation.

Second, the applicability of heuristics depends on the environments. For instance, when cues have similar levels of importance, relying on a single cue may be insufficient for identifying the best option. In such cases, the effectiveness of the take-the-best heuristic diminishes as cue validities converge. Conversely, when cue importance varies

---

<sup>2</sup>As they are sufficient to address the point we aim to make in this paper without requiring customized rules for handling negative cues.

greatly, disregarding cue rankings may lead to suboptimal decisions. In such cases, the effectiveness of the tallying heuristic decreases as cue validities diverge. We will show that the ABC version of take-the-best and tallying replicates these observations.

Apart from general ecological characteristics, performance could also vary across the testing trials under some specific environments. For example, the cues can be non-compensatory, where the most important cue carries more importance than the sum of all other cues, or compensatory, where the combined importance of the less important cues exceed that of the best cue. In compensatory environments, take-the-best may make errors in certain trials—specifically when the unchosen option would have been better if all cues had been taken into account. We will show that the ABC version of take-the-best will generally replicate this observation while retaining its accuracy in certain special cases due to its consideration of uncertainty.

### ***Bayesian Benchmark***

In general, the probability that alternative  $A$  is better than  $B$ , based all cues  $F$  can be written as

$$p(A > B|F) \propto p(F|A > B)p(A > B) \quad (8)$$

where  $A > B = \mathbb{1}[V[A] > V[B]]$ . Assuming the alternatives have the same prior, and the cues contribute independently to the likelihood function, the Bayesian benchmark decision simplifies to the log-odds that one alternative is better (Lee & Cummins, 2004):

$$L_{AB} = \sum_{i \in F_A} \ln\left(\frac{\hat{v}_i}{1 - \hat{v}_i}\right) - \sum_{i \in F_B} \ln\left(\frac{\hat{v}_i}{1 - \hat{v}_i}\right) \quad (9)$$

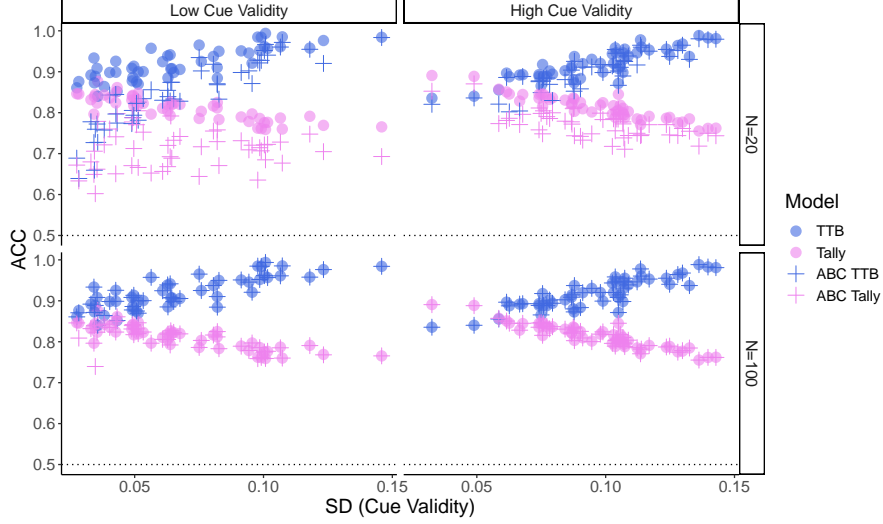
where  $\hat{v}_i$  is the estimated validity of cue  $i$  given the training evidence,  $F_A$  contains the features that are present in  $A$  but absent in  $B$ , and  $F_B$  contains the features that are present in  $B$  but absent in  $A$ . The benchmark answer will be set as  $A$  when  $L_{AB} > 0$  and  $B$  when  $L_{AB} < 0$ .

### ***Summary Statistics***

We consider take-the-best and tallying as two different types of summary statistics, both of which can be applied within the ABC framework. Take-the-best corresponds to a summary statistic that first sorts the cues in descending order of cue validity, then returns the value difference of the first discriminating cue (i.e., a cue that appears in one alternative but not the other). Given that all cues are usable in the current set up, the take-the-best summary statistic will take on one of two values,  $\{1, -1\}$ , indicating whether  $A$  or  $B$  wins, respectively.<sup>3</sup> tallying, on the other hand, corresponds to the difference in the count of positive cues (i.e., cues with  $\hat{v} > 0.5$ ) for both  $A$  and  $B$ , and calculates the difference between the two counts. It will range from -6 to 6 in the current set up.

---

<sup>3</sup>This summary value can be computed directly for an observation in various ways. One example is  $\mathbf{s}(\mathbf{y}) = 1 \times \text{Sign}(\text{Binary}(\text{Sort}(V_A, \text{rank}_{V_B})) - \text{Binary}(\text{Sort}(V_B, \text{rank}_{V_A}))) - \frac{1}{2}$ , which treats the ranked cue values as a binary number, enforcing Take-the-best's non-compensatory property — e.g. 100000 (32) is larger than 011111 (31).

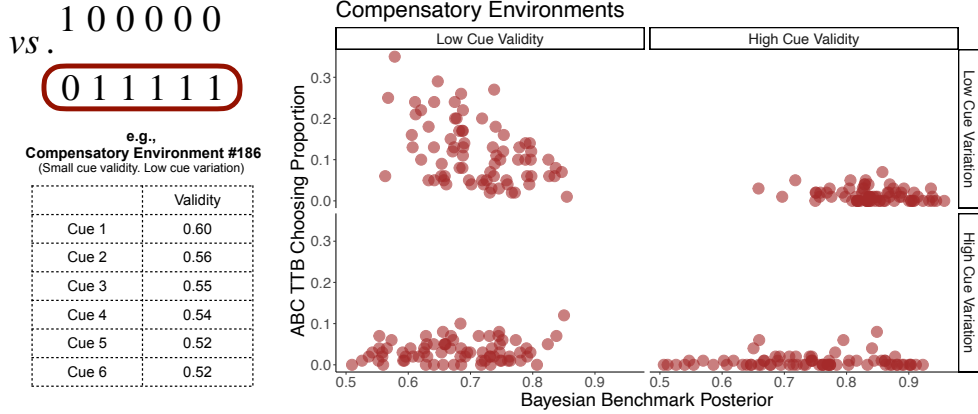


**Fig. 2:** Scatterplots of the standard deviation of cue validity and performance accuracy in four models: Take-the-best, tallying, ABC version of Take-the-best, and ABC version of tallying.  $N$  indicates the sample size for acceptance in rejection ABC. Dotted lines indicate the change level (50%).

### Results

Because the possible values of the take-the-best and tallying summary statistics are limited and discrete, we fix the acceptance threshold to 0 for take-the-best ABC and 1 for tallying ABC. This means that the value must match exactly for take-the-best, and can differ by at most 1 for tallying summary statistics. We here use samples of  $N = \{20, 100\}$  to capture randomness in the response. 100 ABC observers were simulated for each environment.

We first simulated 100 environments and divided them into low or high cue validity groups based on the median of the average validity of six cues. As shown in Figure 2, the heuristic algorithms achieved high accuracies (i.e., closely aligned with the Bayesian benchmark choices), significantly above random decision-making (0.5). Accuracy was influenced by cue validity variation: The accuracy of take-the-best increased as the variation increased (i.e. when one cue was more likely to dominate the diagnosis), whereas the accuracy of tallying increased as the variation decreased. Importantly, the ABC versions of take-the-best and tallying closely tracked the performance of their original deterministic counterparts. That is, they align with the deterministic versions when the ABC observer makes decisions based on a large sample size ( $N = 100$ ), but exhibit more stochasticity when the sample size is small ( $N = 20$ ). This stochasticity increases when cue validity is generally low and variation is minimal (top-left panel, and the left side within the panel) — environments under which even the Bayesian benchmark struggles to make accurate decisions. This suggests that ABC heuristics not only replicate the behaviors of traditional heuristics, but also make further predictions about how the responses become uncertain in



**Fig. 3:** Scatterplots showing the proportion that the ABC take-the-best algorithm selected the compensatory answer in compensatory environments. The examples on the left illustrate a non-compensatory choice (top) versus a compensatory choice (bottom, highlighted in the brown box) in a compensatory environment where a combination of non-best cues can outweigh the absence of the best cue.

particular situations. These predictions are also consistent with the human tendency to behave more stochastically in highly noisy environments as suggested in previous research (Newell, Koehler, James, Rakow, & Van Ravenzwaaij, 2013; Shanks, Tunney, & McCarthy, 2002; Vulkan, 2000; Wozny, Beierholm, & Shams, 2010).

One type of trial used to diagnose whether an observer follows Bayesian rational decision-making or the take-the-best heuristic involves situations where the option containing the best cue is actually worse than the alternative, which lacks the best cue but includes other informative cues. This scenario can arise in compensatory environments, where a combination of non-best cues can outweigh the absence of the best cue. We simulate 300 such compensatory environments to evaluate how the ABC-take-the-best model (ABC-TTB) performs. We focus on trials in which Option A (the upper option in Figure 3) is positive only on the best cue, whereas Option B (the lower option in Figure 3) is positive on all cues except the best one. In these trials, the take-the-best heuristic will always select Option A. In contrast, a Bayesian decision-maker will consistently favor Option B, although the posterior probability will vary depending on the specific environment.

As shown in Figure 3, the ABC take-the-best observer often replicates the classic take-the-best heuristic by not choosing Option B — the the Bayesian benchmark’s preference. However, there remains a small probability for it to Option B — especially in situations where cue validity and cue variation are both low, indicating a high degree of uncertainty even from the Bayesian perspective. These results highlight that the ABC take-the-best model captures not only the deterministic behavior of the take-the-best heuristic but also predicts stochasticity reflecting rational uncertainty in the responses.

## Continuous-time causal learning

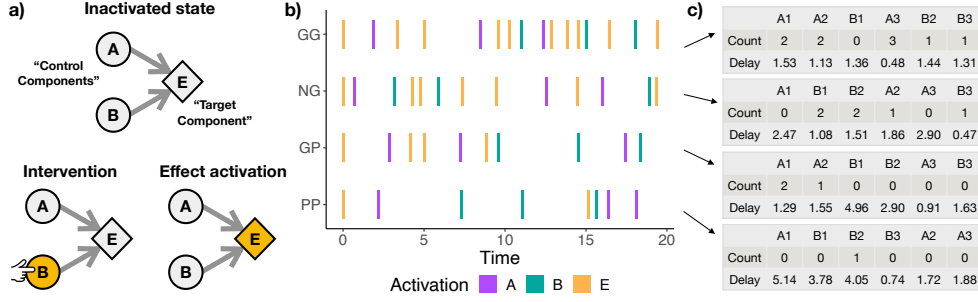
In the second case example, we move from the traditional, discrete-trial-based paradigm, to a more ecological paradigm that requires people to process rich, real-time evidence to make judgments. This approach emerges from recent studies examining people’s intuitive understanding of the world, such as how they infer the physical properties and relationships of objects (Bramley et al., 2018; Ullman et al., 2018), the mental states and social relationships of agents (Baker et al., 2017; Gerstenberg & Tenenbaum, 2017; Liu, Ullman, Tenenbaum, & Spelke, 2017), and the causal structure of events in general (Btesh, Bramley, Speekenbrink, & Lagnado, in press; Davis, Bramley, & Rehder, 2020; Gong & Bramley, 2023; Gong et al., 2023). Given the complexity of the task structure, a single heuristic can seldom capture human responses (Callaway et al., 2022; Gerstenberg et al., 2021). However, research also shows that rational Bayesian inference does not capture human responses well either (Gong & Bramley, 2023; Song et al., 2019; Van Opheusden et al., 2023).

A key feature of these tasks is that, given a candidate structural hypothesis, one can fairly easily imagine a scene or sequence of events it might produce. However, there exists an infinity of such possible scenes or sequences, with much of the underlying generative process remaining unobserved. It creates a situation in which the likelihood is extremely difficult or impossible to calculate, providing another fertile ground for studying heuristic decision-making strategies. Here, we use Gong and Bramley (2023) as a representative example of this line of research and reanalyze its data under the ABC approach.

### *The Task*

In the task, participants observed causal systems that produce series of point events in continuous time, and then judged the causal relationships linking control components to a target component based on these temporal patterns of events. Each causal system included two control components,  $A$  and  $B$ , which were intervened upon by a pre-set program, and one target component,  $E$  (Figure 4a). Participants were asked to use the available temporal information to whether the control component  $A$  or  $B$  was a generative cause (G) or a preventative cause (P) of the target component, or non-causal (N, see Figure 4b for stimulus examples).

Participants were instructed on  $E$ ’s base rate and how a generative or preventative component would work. For instance, a generative component, when activated, would produce one extra  $E$  activation with a delay of around 1.5 seconds. A preventative component, when activated, would block all  $E$ ’s activations for around 3 seconds. Even though the rules were simple, the evidence could involve complex dynamics: An effect event expected to be generated by one cause might be prevented by another cause. An effect event in fact generated by one cause might appear after the intervention of another cause. These possibilities dramatically compound the complexity of calculating the exact likelihood that a causal structure produces a particular observation.



**Fig. 4:** Illustration of task completed by participants in Gong and Bramley (2023). (a) The two “control components” were shown to be intervened on, and participants can observe the following effects. (b) Example stimulus sequences. (c) The Count and Delay summary statistics. If no effect follows an intervention, the Delay will be calculated as the delay between it and the next intervention (or the end of the observation).

### Bayesian Benchmark

A normative Bayesian analysis here would rely on enumerating and marginalizing over all possible token-level causation (i.e., which particular event actually caused which particular effect), to explain all effects observed, as well as all effects that were expected but not observed (e.g., due to the existence of a preventative cause; see Gong & Bramley, 2023, for modeling details). Given that any effect event could potentially be attributed to any cause event that happened before, the computational complexity grows rapidly with observation duration, and quickly become intractable once the observation period lasts longer than the particular experimental settings used in Gong and Bramley (2023).

### Summary Statistics

Gong and Bramley (2023) proposed people chunk the continuous time evidence and focuses on summary statistics of what happened in the period between two interventions (i.e., A’s or B’s activations). Two summary statistics are proposed: Count (how many effect events occur after an intervention and before the next intervention) and Delay (how long is the delay between the intervention and its following effect event, regardless of what truly generated that effect). Those cues have an intuitive relationship with diagnosing the causal types (Gong et al., in press; Valentin, Bramley, & Lucas, 2023). For example, one will intuitively expect to see more effect activations following the intervention of a generative cause (due to the combination of base rate activations and generation). Conversely, fewer activations are likely to follow the intervention of a preventative cause.

We here implement the proposed summary statistics to the general ABC rejection sampling framework as shown in Algorithm 1. This will provide a parsimonious explanation of the variations in participants’ responses, as directly arising from the

randomness inherent in the sampling process.<sup>4</sup> Given that the study involved stimuli with three interventions to each component  $A$  and  $B$ , there are different ways to summarize the data. We here test which one of the two ways could better capture people’s judgments. One is to use the “Full Vector”, which includes delay and count features for all six interventions —  $A_1$ ,  $A_2$ ,  $A_3$ ,  $B_1$ ,  $B_2$ , and  $B_3$  (see Figure 4c). The other uses the “Mean Vector” which averages each feature across components, resulting in a reduced total of four features.

## Results

We compared our Full Vector and Mean Vector ABC models with the Bayesian benchmark, examining which can better explain participants’ response. Each model contains one fit parameter. A tolerance threshold parameter was fit for the ABC models and a softmax temperature parameter was fit for the Bayesian benchmark. We use the Wasserstein distance as the metric for model comparison for the ability to capture both the magnitude and distribution of differences between predicted and observed outcomes.

Results are shown in Figure 5. The lowest Wasserstein distance is achieved by the ABC Full Feature model (0.050), compared to the Bayesian normative benchmark (0.062) and the ABC Mean Feature model (0.053). Qualitatively, the full normative benchmark model also provides a considerably poorer explanation of human inferences compared to the more heuristic accounts, as the latter better capture participants’ responses—both in terms of the degree to which they chose the correct structure (upper right) and the degree to which they made specific errors (bottom left). This helps demonstrate the idea that, instead of following the highly demanding rational Bayesian inference, people tend to rely on simple features in event streams, the heuristics, to make judgments. The advantage of using the ABC framework here also lies in its ability to predict the distribution of human responses beyond simple point estimates (Figure 5), something that cannot be easily achieved under the classic heuristic approach.

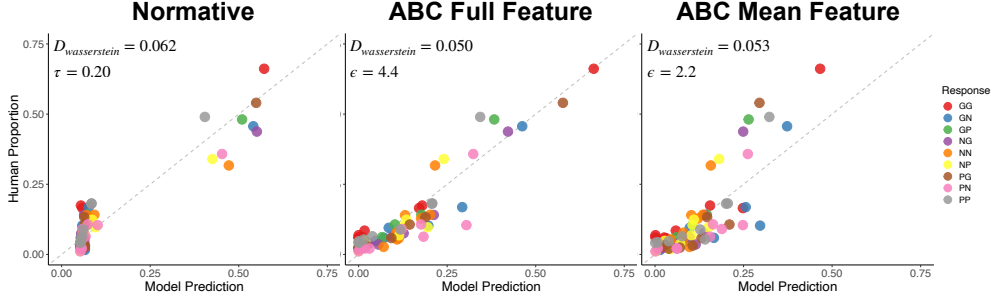
## General Discussion

In this paper, we have proposed that reliance on simple heuristics can be framed as a resource-rational approximation to a more complete Bayesian inference. By drawing parallels between heuristics and likelihood-free inference methods, particularly ABC, we offer a new insights into the nature and adaptive value of heuristics. It maintains the normative Bayesian framework while acknowledging and working within real-world computational constraints, providing a bridge between ideal observer models and the heuristics employed in human cognition. As briefly mentioned in the introduction, it addresses three key challenges in understanding human decision-making: applicability, uncertainty representation, and relationship with optimality.

---

<sup>4</sup>This idea has driven the modeling in [Gong and Bramley \(2023\)](#), but the original paper does not involve the sampling or decision process scheme we use here. Instead it focuses more on a computational level analysis of the summary statistics’ distributions and kernel based likelihood estimates from this.





**Fig. 5:** Scatterplots of model predictions and human judgments show the proportion of choices for different causal types under various ground truths. Each ground truth is represented by nine data points, corresponding to the nine possible response types. The data are fitted at the individual stimulus level and visualized at the aggregate level by averaging responses within the corresponding categories.

## Applicability of heuristics

The ABC framework significantly extends the applicability of heuristic models to complex, real-world scenarios by relaxing strong assumptions about cue-target relationships. For example, the reanalysis of [Gong and Bramley \(2023\)](#) demonstrates this enhanced applicability in at least two ways: First, our ABC model successfully captured human judgments in a continuous-time causal inference task characterized by intricate temporal dynamics, which presents a challenge for traditional heuristic models due to its complex cue-outcome mappings. By employing summary statistics within an ABC framework, we demonstrated how ABC can extend heuristic modeling beyond the typical domain of static cue structures. Second, our approach exhibited considerable flexibility in cue selection. By exploring various summary statistics, including features per action and overall mean features, we illustrated how ABC facilitates adaptive cue selection without requiring strong a priori assumptions about cue validity. This flexibility allows heuristic models to adapt to diverse task structures, addressing a key limitation of traditional approaches that often rely on fixed, predefined cues.

## Uncertainty representation

A common criticism of heuristic models is their limited capacity to represent uncertainty. Our ABC framework naturally provides predictions about how variability in human responses tracks uncertainty, given the stochastic nature of the sampling procedure. In the first illustrative example, we captured the phenomenon that humans tend to behave more stochastically when environmental uncertainty is high ([Newell et al., 2013](#); [Shanks et al., 2002](#); [Vulkan, 2000](#); [Wozny et al., 2010](#)). As shown in the simulation (Figure 3), this behavior ultimately leads to better payoffs than the traditional take-the-best strategy in environments characterized by low cue validity and low cue variation. In the second example, it demonstrated the ability not only to capture the most frequent human judgments but also to reflect the systematic error patterns in

responses (Figure 5). This suggests that our LFI approach is modeling the full distribution of human judgments, not merely the average behavior indicated by heuristic models.

## Relationship with optimality

The LFI framework provides a nuanced perspective on the relationship between heuristics and optimal inference. Our ABC model can approximate normative inferences (given sufficient statistics, large numbers of samples, and  $\epsilon \rightarrow 0$ ), but anticipates different kinds of departures from Bayesian norms, due to insufficient summary statistics, larger  $\epsilon$ , and fewer samples. The performance of our computationally simpler ABC model over the normative model exemplifies the principle of computational rationality. By framing heuristics as LFI, we establish a formal link between heuristic strategies and Bayesian inference, positioning heuristics not as departures from Bayesian rationality, but as rational approximations to likelihood calculations under computational constraints. Moreover, the varying performance of different summary statistics in our ABC model provides a formal framework for understanding why certain heuristics are more effective or widely used than others.

## Towards process-level explanations

One promising line of work is given by integrating the ideas of likelihood-free inference with the literature on inference-by-sampling (Castillo, León-Villagr , Chater, & Sanborn, 2024; Chater & Oaksford, 2008; Daw & Courville, 2008; Griffiths et al., 2012; Sanborn, 2017; Sanborn & Chater, 2016; Sanborn et al., 2010; Shi & Griffiths, 2009; Stewart, Chater, & Brown, 2006; Zednik & J kel, 2016; Zhu, Chater, et al., 2023; Zhu et al., 2020; Zhu, Sundh, et al., 2023). In common with the inference-by-sampling (IBS) accounts of human cognition, we argue that while the generative model is the true latent representation of one’s belief we generically have to rely on samples to access our beliefs and do inferences that respect them. The crucial distinction lies in how these samples are generated and evaluated: standard IBS approaches typically assume the mind can calculate unnormalized likelihoods to determine whether to accept or reject samples. In contrast, we contend that likelihoods, even unnormalized ones, are generally intractable in complex, real-world scenarios. By eliminating the need for explicit likelihood calculations, the ABC approach offers a more realistic model of how mental samples can be evaluated. At the same time, our proposal is compatible with recent IBS models: To implement ABC for a particular choice of summary statistics and distance function, models like the autocorrelated Bayesian sampler (Zhu, Sundh, et al., 2023) generally need only replace their unnormalized likelihoods with  $\mathbb{1}(d(\mathbf{s}(\mathbf{y}^{\text{sim}}), \mathbf{s}(\mathbf{y})) \leq \epsilon)$  indicator functions.

The ABC framework can be naturally extended to incorporate key features of inference-by-sampling models, enhancing its explanatory power and cognitive plausibility. Rejection sampling is one of the simplest methods available for likelihood-free inference. We implement this basic form in ABC, but there are several alternatives that could be implemented and may serve as cognitively plausible sampling procedures.

Future work may benefit from investigating different versions of ABC, examining subtle differences in the responses, and exploring which variant better explains people’s behavior. For example, by adapting our approach to use a Markov Chain Monte Carlo (MCMC) sampler with a local proposal distribution, instead of sampling directly from the prior, we can create a more efficient model than rejection-ABC (Marjoram, Molitor, Plagnol, & Tavaré, 2003). This extension would not only maintain the core strength of our likelihood-free approach but also capture patterns of autocorrelation and anchoring observed in human learning. This extended ABC model would predict a range of phenomena already associated with IBS models, such as autocorrelation effects, speed-accuracy trade-offs, and calibration effects under speeded judgment. Moreover, it could generate novel predictions unique to our likelihood-free framework. For instance, our model anticipates differences in speed-accuracy trade-offs based on the surprisingness of observations. Unlike tractable-likelihood models, which are indifferent to how unusual evidence is in absolute terms, our ABC approach predicts that surprising examples would require more processing time. This is because unusual observations would be less likely to match the summary statistics of simulated data, necessitating more simulations to achieve a successful match. This distinctive prediction offers a testable hypothesis that could differentiate our framework from traditional IBS approaches.

In this paper, we assume that the reasoner simulates concrete samples and extracts summary statistics from them. However, at the process level, the simulator itself may be more abstract, neglecting nuisance variables and focusing on what is more relevant to the task (Ho, Abel, et al., 2022). This links back to the core idea of the classic representativeness heuristic, which argues that people judge the probability of an event by how similar it is to their mental prototype (Gigerenzer, 1991; Griffiths, Daniels, Austerweil, & Tenenbaum, 2018; Kahneman et al., 2002; Kahneman & Tversky, 1972). Within the ABC framework, the prototype can be understood as an imperfect simulator that encodes only the most salient features of a category. When a characteristic is atypical for a category, the simulator may fail to represent it and thus never generate it, making it absent in samples from that category. In contrast, categories for which the characteristic is stereotypical will faithfully include it in their samples, giving them an advantage in perceived similarity to the observation. Future work could relax the constraints that are still considered normative (e.g., a normative simulator) in the ABC framework to examine whether this better captures human behavior.

## Where do summary statistics come from?

The origin and selection of summary statistics are crucial aspects of this framework, as summary statistics underpin heuristics in likelihood-free inference. Although we do not propose a general solution here, we identify several sources and mechanisms by which summary statistics can arise, each offering insights into human cognition and decision-making.

In traditional statistical approaches, summary statistics derive from domain knowledge and the mathematical structure the data-generating process. For instance, in Gaussian settings, the sample mean and variance are natural choices due to their sufficiency and simplicity. Likewise, exponential family distributions use natural sufficient

statistics. These approaches are effective in well-understood domains but struggle with complex, high-dimensional data where relevant features are not obvious.

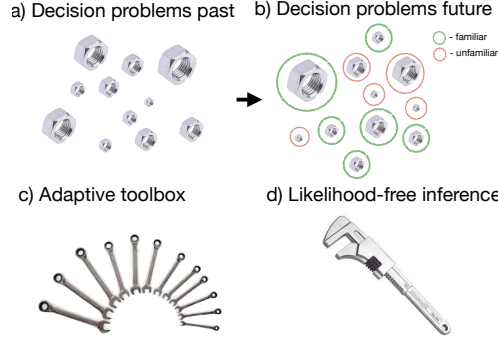
This limitation shifts the challenge of discovering useful summary statistics to earlier stages of learning, e.g., through iterative trial-and-error of domain knowledge within and across generations (see next section). A more general, data-driven approach, which is common in modern machine learning, learns summary statistics directly from data via deep neural networks (Chen, Gutmann, & Weller, 2023; Chen et al., 2020; Jiang, Wu, Zheng, & Wong, 2017). Such methods can capture information, non-obvious features while discarding irrelevant or redundant ones.

Humans, too, seem to learn summary statistics through experience and exposure to data. This can be viewed as *amortized inference*: the effort of identifying relevant features is distributed across multiple encounters with similar problems. For example, medical professionals learn to quickly identify key symptoms and risk factors for particular diseases (see, e.g., Gobet, 2017), effectively constructing internal mappings from observations to summary statistics — analogous to the abstraction process in ABC.

In addition to learning from real-world experiences, humans can also learn summary statistics through mental simulation (e.g., Wittkuhn, Chien, Hall-McMaster, & Schuck, 2021). By generating data from internal models of the world, individuals can train their recognition systems, offline, to extract informative features based on replay of past experiences. This could be related to a presumed core feature of cognition: the wake-sleep cycle, in which offline replay during rest and sleep facilitates semanticisation and abstraction (Dayan, Hinton, Neal, & Zemel, 1995; Ellis et al., 2023; Mattar & Daw, 2018; Sablé-Meyer, Ellis, Tenenbaum, & Dehaene, 2022).

The acquisition of summary statistics can occur across various timescales. Over developmental time, children learn to extract increasingly sophisticated statistical regularities, as demonstrated by infants’ ability to use transitional probabilities for speech segmentation (Saffran, Aslin, & Newport, 1996) and older children’s capacity to infer causal structures from data (Gopnik & Schulz, 2004). Throughout a lifetime, individuals accumulate domain-specific expertise, refining their ability to identify relevant features in familiar contexts, as seen in chess masters’ rapid pattern recognition (Chase & Simon, 1973) or radiologists’ efficient identification of diagnostic features (Kundel, Nodine, Conant, & Weinstein, 2007). On the shortest timescale, individuals can rapidly adapt to their current task through trial-and-error or instruction, as evidenced by studies on human categorization showing quick learning of task-relevant stimulus dimensions (Nosofsky, 1986).

This multi-timescale view helps highlight both the supreme flexibility of human cognition in its ability to adapt to momentarily novel situations and the remarkable efficiency of expert intuition in familiar domains. The origin of summary statistics in human cognition likely involves a complex interplay between innate predispositions, individual learning experiences, and cultural transmission of knowledge. The ABC framework suggests that heuristics, viewed as applications of summary statistics, are best thought of as a general purpose learning architecture rather than a set of fixed strategies, allowing individuals to adapt to new situations and accumulate experience in a complex environment that cannot be represented exactly.



**Fig. 6:** (a) A range of nuts of different gauges representing a history of decision problems. (b) Newly encountered nuts in a mixture of familiar and unfamiliar gauges. (c) A set of gauged wrenches representing an Adaptive Toolbox of fixed heuristics, purchased to solve the problems in a, but now hard to select from and unsuitable for some problems in b. (d) A monkey wrench, representing general likelihood-free inference mechanism such as ABC. While in use, it gives the appearance of a fixed strategy (fixed gauge), but can be endlessly adjusted to new problems.

This perspective inverts the classical heuristics literature’s view of the mind as containing an “adaptive toolbox” filled with situation-specific heuristics (Gigerenzer & Selten, 2002). The adaptive toolbox view notoriously leads to the meta-cognitive challenges of selecting, switching between strategies (Lieder & Griffiths, 2017) and not to mention inventing new ones (Krueger, Callaway, Gul, Griffiths, & Lieder, 2024). While likelihood-free view does not make the problem of generalization go away, nor does it imply that the mind solves it by storing and selecting from proliferating set of fixed strategies (i.e. the ever-growing list of heuristics documented by psychologists). It also makes heuristic invention less mysterious, by associating this with a generic cognitive process of abstraction of inputs to lower dimensional representations. We illustrate this conceptual inversion with an analogy to the problem of provisioning a toolbox for adaptivity flexibility (Figure 6).

## Amortized inference and heuristics

The concept of amortized inference provides a complementary perspective on how heuristics might be learned and applied in complex real-world domains. This idea aligns naturally with our proposal and connects to previous accounts of heuristics, though it partially focuses on different aspects than what we explore in this work. We suggest that a mixed strategy combining amortization and likelihood-free inference-by-sampling is both plausible and cognitively efficient.

Binz et al. (2022) propose Bounded Meta-Learned Inference (BMLI), an account of how people discover and select heuristics in decision-making, which addresses the challenges implied by the adaptive toolbox metaphor above. Their work suggests that through meta-learning, individuals exploit environment structure to select

heuristics while managing computational resources. Different heuristics emerge as resource-rational strategies depending on available information: one-reason decision-making when feature rankings are known, equal weighting when feature directions are known, and weighted combinations when neither are known. In contrast to our work, BMLI assumes that people optimize for future predictions, rather than approximating Bayesian inference. Related to this, [Dasgupta et al. \(2020\)](#) propose a theory of “learning to infer”, to explain systematic deviations from Bayesian optimality. Their model uses a recognition network that maps queries to probability distributions. It allocates computational resources (as implemented by restricted the size of the network) preferentially to frequently encountered queries, leading to characteristic biases observed in human inference.

In the context of our framework, these approaches could be reconciled by considering the mechanisms they postulate for learning summary statistics. The notion of focusing computational effort on the most informative aspects of data aligns closely with our proposal that people select appropriate summary statistics for the task at hand. Further, for frequently encountered problems, people may not simulate a number of explicit possibilities, but rather perform a direct mapping to an inference through pattern-matching, or perform a small number of very informed simulations.

It is important, however, to distinguish between these variational-optimization approaches and our sampling-based framework. While both offer explanations for how the mind might approximate Bayesian inference, they differ in their underlying mechanisms and constraints ([Sanborn, 2017](#)), and we hope the present proposal provides a useful framework for investigating these differences further.

This raises the question of why we do not simply amortize everything, given the efficiency gains of amortization. Several considerations argue against this. First, amortized inference can struggle with novel or out-of-distribution situations, as it generalizes poorly beyond its training domain. In contrast, Approximate Bayesian Computation (ABC) methods are often more robust under such conditions ([Kelly et al., 2025](#)). Second, amortization requires substantial upfront computational and memory costs (analogous to “buying many gauged wrenches” in Figure 6). For rare or rapidly changing environments, these costs may outweigh the benefits, especially considering memory costs. Maintaining some capacity for online inference based on rich simulator models allows for handling unexpected scenarios more effectively. Finally, pure amortization would lead to overfitting specific problem structures, reducing capacity to generalize to new situations (Figure 6b), especially when there are only a few examples to generalize from.

In light of these considerations, a mixed strategy of amortization and inference-by-sampling seems like a sensible trade off. It would balance the computational efficiency of amortized inference with the flexibility and expressiveness of simulation-based inference. In our framework, people might perform mental simulations using a simulator model and subsequently applying summary statistics learned through amortization to these simulations. Adaptation could involve adjusting which summary statistics are used, how many samples are generated, or the acceptance threshold. Furthermore, parts of the simulation process itself might be amortized, with full simulation engaged only when necessary, as a form of multi-fidelity simulation. It aligns well

with the notion of learning heuristics or intuitive theories through experience: As people encounter various inferential problems within a domain, they may gradually learn to identify the most informative regularities that can guide their mental simulations and support rapid, approximate inferences. For instance, an experienced chef might develop heuristics for cooking decisions, such as adding a dash of soy sauce to bland dishes or reducing sautéing time by two minutes for smaller onions. These learned heuristics can then be generalized to novel but related problems, allowing for efficient generation of plausible hypotheses or predictions.

## Conclusion

In this paper, we developed a unifying framework that views heuristics as resource-rational approximations to Bayesian inference, drawing connections between cognitive shortcuts, mental simulation, and likelihood-free inference methods. This perspective can improve our understanding of how heuristics operate and also illuminate why they emerge and how they relate to broader principles of cognition. Our approach reconciles seemingly contradictory findings in the literature, such as the tension between the simplicity of heuristics and the complexity of the computational problems minds face in uncertain choice. This work opens up several promising avenues for future research, such as examining how heuristics are learned or adapted over time within this framework, potentially providing new insights into cognitive development and expertise acquisition. By integrating insights from cognitive science, machine learning, and statistics, we provide a novel perspective on the nature and origin of heuristics in human cognition.

## Acknowledgments

TG is supported by a British Academy Postdoctoral Fellowship Scheme (PFSS24\24009). This work was done prior to SV joining Amazon.

## References

- Anderson, J.R., & Matessa, M. (1992). Explorations of an incremental, bayesian algorithm for categorization. *Machine Learning*, 9, 275–308,
- Baker, C.L., Jara-Ettinger, J., Saxe, R., Tenenbaum, J.B. (2017). Rational quantitative attribution of beliefs, desires and percepts in human mentalizing. *Nature Human Behaviour*, 1(4), 0064,
- 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,



- Bergert, F.B., & Nosofsky, R.M. (2007). A response-time approach to comparing generalized rational and take-the-best models of decision making. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 33(1), 107,
- Bhui, R., Lai, L., Gershman, S.J. (2021). Resource-rational decision making. *Current Opinion in Behavioral Sciences*, 41, 15–21,
- Binz, M., Dasgupta, I., Jagadish, A.K., Botvinick, M., Wang, J.X., Schulz, E. (2023). Meta-learned models of cognition. *Behavioral and Brain Sciences*, 1–38,
- Binz, M., Gershman, S.J., Schulz, E., Endres, D. (2022). Heuristics from bounded meta-learned inference. *Psychological Review*, 129(5), 1042,
- Bitzer, S., Park, H., Blankenburg, F., Kiebel, S.J. (2014). Perceptual decision making: drift-diffusion model is equivalent to a bayesian model. *Frontiers in Human Neuroscience*, 8, 102,
- Blum, M.G., Nunes, M.A., Prangle, D., Sisson, S.A. (2013). A comparative review of dimension reduction methods in approximate bayesian computation. *Statistical Science*, 28(2), 189–208,
- Bramley, N.R., Gerstenberg, T., Tenenbaum, J.B., Gureckis, T.M. (2018). Intuitive experimentation in the physical world. *Cognitive Psychology*, 105, 9–38,
- Bramley, N.R., & Xu, F. (2023). Active inductive inference in children and adults: A constructivist perspective. *Cognition*, 238, 105471,
- Brighton, H., & Gigerenzer, G. (2012). Homo heuristicus and the bias–variance dilemma. J. Schulkin (Ed.), *Action, perception and the brain: Adaptation and cephalic expression* (pp. 68–91). Springer.
- Btsh, V.J., Bramley, N., Speekenbrink, M., Lagnado, D. (in press). Less is more: Local focus in continuous time causal learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, ,
- Callaway, F., van Opheusden, B., Gul, S., Das, P., Krueger, P.M., Lieder, F., Griffiths, T.L. (2022). Rational use of cognitive resources in human planning. *Nature Human Behaviour*, , <https://doi.org/10.1038/s41562-022-01332-8>

- Castillo, L., León-Villagr , P., Chater, N., Sanborn, A. (2024). Explaining the flaws in human random generation as local sampling with momentum. *PLOS Computational Biology*, 20(1), e1011739,
- Chase, W.G., & Simon, H.A. (1973). Perception in chess. *Cognitive psychology*, 4(1), 55–81,
- Chater, N., & Oaksford, M. (2008). *The probabilistic mind: Prospects for bayesian cognitive science*. Oxford University Press, USA.
- Chen, Y., Gutmann, M.U., Weller, A. (2023). Is learning summary statistics necessary for likelihood-free inference? *International conference on machine learning* (pp. 4529–4544).
- Chen, Y., Zhang, D., Gutmann, M., Courville, A., Zhu, Z. (2020). Neural approximate sufficient statistics for implicit models. *arXiv preprint arXiv:2010.10079*, ,
- Christiansen, M.H., & Chater, N. (2016). The now-or-never bottleneck: A fundamental constraint on language. *Behavioral and Brain Sciences*, 39, e62,
- Chun, M.M., Golomb, J.D., Turk-Browne, N.B. (2011). A taxonomy of external and internal attention. *Annual Review of Psychology*, 62(1), 73–101,
- Colombo, M., & Seri s, P. (2012). Bayes in the brain—on bayesian modelling in neuroscience. *The British Journal for the Philosophy of Science*, 63(3), 697–723,
- Cowan, N. (2010). The magical mystery four: How is working memory capacity limited, and why? *Current Directions in Psychological Science*, 19(1), 51–57,
- Cranmer, K., Brehmer, J., Louppe, G. (2020). The frontier of simulation-based inference. *Proceedings of the National Academy of Sciences*, 117(48), 30055–30062,
- Dasgupta, I., Schulz, E., Tenenbaum, J.B., Gershman, S.J. (2020). A theory of learning to infer. *Psychological Review*, 127(3), 412–441,

- Davis, Z., Bramley, N.R., Rehder, B. (2020). Causal structure learning in continuous systems. *Frontiers in Psychology*, 11, 244,
- Daw, N., & Courville, A. (2008). The pigeon as particle filter. *Advances in Neural Information Processing Systems*, 20, 369–376,
- Dawes, R.M., & Corrigan, B. (1974). Linear models in decision making. *Psychological Bulletin*, 81(2), 95–106,
- Dayan, P., Hinton, G.E., Neal, R.M., Zemel, R.S. (1995). The helmholtz machine. *Neural Computation*, 7(5), 889–904,
- DeMiguel, V., Garlappi, L., Uppal, R. (2009). Optimal versus naive diversification: How inefficient is the 1/n portfolio strategy? *The review of Financial studies*, 22(5), 1915–1953,
- Ellis, K., Wong, L., Nye, M., Sable-Meyer, M., Cary, L., Anaya Pozo, L., ... Tenenbaum, J.B. (2023). Dreamcoder: growing generalizable, interpretable knowledge with wake-sleep bayesian program learning. *Philosophical Transactions of the Royal Society A*, 381(2251), 20220050,
- Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, 11(2), 127–138,
- Gershman, S.J., Horvitz, E.J., Tenenbaum, J.B. (2015). Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science*, 349(6245), 273–278,
- Gerstenberg, T., Goodman, N.D., Lagnado, D.A., Tenenbaum, J.B. (2021). A counterfactual simulation model of causal judgments for physical events. *Psychological Review*, 128(5), 936–975,
- Gerstenberg, T., & Tenenbaum, J.B. (2017). Intuitive theories. M. Waldmann (Ed.), *The oxford handbook of causal reasoning* (pp. 515–548). New York: Oxford University Press.
- Gigerenzer, G. (1991). How to make cognitive illusions disappear: Beyond “heuristics and biases”. *European Review of Social Psychology*, 2(1), 83–115,

- Gigerenzer, G., & Brighton, H. (2009). Homo heuristicus: Why biased minds make better inferences. *Topics in Cognitive Science*, 1(1), 107–143,
- Gigerenzer, G., & Goldstein, D.G. (1996). Reasoning the fast and frugal way: models of bounded rationality. *Psychological Review*, 103(4), 650–669,
- Gigerenzer, G., Hoffrage, U., Kleinbölting, H. (1991). Probabilistic mental models: a brunswikian theory of confidence. *Psychological Review*, 98(4), 506–528,
- Gigerenzer, G., & Selten, R. (2002). *Bounded rationality: The adaptive toolbox*. MIT press.
- Gigerenzer, G., & Todd, P.M. (1999). *Simple heuristics that make us smart*. Oxford University Press, USA.
- Gobet, F. (2017). *Understanding expertise: A multi-disciplinary approach*. Bloomsbury Publishing.
- Goldstein, D.G., & Gigerenzer, G. (2008). The recognition heuristic and the less-is-more effect. *Handbook of experimental economics results*, 1, 987–992,
- Gong, T., & Bramley, N.R. (2023). Continuous time causal structure induction with prevention and generation. *Cognition*, 240, 105530,
- Gong, T., Gerstenberg, T., Mayrhofer, R., Bramley, N.R. (2023). Active causal structure learning in continuous time. *Cognitive Psychology*, 140, 101542,
- Gong, T., Pacer, M., Griffiths, T.L., Bramley, N.R. (in press). Rational causal induction from events in time. *Psychological Review*, ,
- Gopnik, A., & Schulz, L. (2004). Mechanisms of theory formation in young children. *Trends in Cognitive Sciences*, 8(8), 371–377,
- Gopnik, A., & Tenenbaum, J.B. (2007). Bayesian networks, bayesian learning and cognitive development. *Developmental Science*, 10(3), 281–287,

- Griffiths, T.L., Daniels, D., Austerweil, J.L., Tenenbaum, J.B. (2018). Subjective randomness as statistical inference. *Cognitive Psychology*, 103, 85–109,
- Griffiths, T.L., Lieder, F., Goodman, N.D. (2015). Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Topics in Cognitive Science*, 7(2), 217–229,
- 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,
- Haefner, R.M., Berkes, P., Fiser, J. (2016). Perceptual decision-making as probabilistic inference by neural sampling. *Neuron*, 90(3), 649–660,
- Hamrick, J.B., Battaglia, P.W., Griffiths, T.L., Tenenbaum, J.B. (2016). Inferring mass in complex scenes by mental simulation. *Cognition*, 157, 61–76,
- Hertwig, R., & Hoffrage, U. (2013). *Simple heuristics in a social world*. Oxford University Press.
- Ho, M.K., Abel, D., Correa, C.G., Littman, M.L., Cohen, J.D., Griffiths, T.L. (2022). People construct simplified mental representations to plan. *Nature*, 606(7912), 129–136,
- Ho, M.K., Saxe, R., Cushman, F. (2022). Planning with theory of mind. *Trends in Cognitive Sciences*, 26(11), 959–971,
- Jaynes, E.T. (2003). *Probability theory: The logic of science*. Cambridge university press.
- Jiang, B., Wu, T.-y., Zheng, C., Wong, W.H. (2017). Learning summary statistic for approximate bayesian computation via deep neural network. *Statistica Sinica*, 1595–1618,

- John, Y.J., Caldwell, L., McCoy, D.E., Braganza, O. (2024). Dead rats, dopamine, performance metrics, and peacock tails: proxy failure is an inherent risk in goal-oriented systems. *Behavioral and Brain Sciences*, 47, e67,
- 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), 169–188,
- Joyce, P., & Marjoram, P. (2008). Approximately sufficient statistics and bayesian computation. *Statistical Applications in Genetics and Molecular Biology*, 7(1), ,
- Kahneman, D. (2011). *Thinking, fast and slow*. Farrar, Straus and Giroux.
- Kahneman, D., Frederick, S., et al. (2002). Representativeness revisited: Attribute substitution in intuitive judgment. T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (p. 49-81). Cambridge University Press.
- Kahneman, D., & Tversky, A. (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology*, 3(3), 430–454,
- Kahneman, D., & Tversky, A. (1973). On the psychology of prediction. *Psychological Review*, 80(4), 237–251,
- Kelly, R.P., Warne, D.J., Frazier, D.T., Nott, D.J., Gutmann, M.U., Drovandi, C. (2025). Simulation-based bayesian inference under model misspecification. *arXiv preprint arXiv:2503.12315*, ,
- Kemp, C., Goodman, N.D., Tenenbaum, J.B. (2010). Learning to learn causal models. *Cognitive Science*, 34(7), 1185–1243,
- Kemp, C., Perfors, A., Tenenbaum, J.B. (2007). Learning overhypotheses with hierarchical bayesian models. *Developmental Science*, 10(3), 307–321,
- Körding, K.P., & Wolpert, D.M. (2004). Bayesian integration in sensorimotor learning. *Nature*, 427(6971), 244–247,

- Krueger, P.M., Callaway, F., Gul, S., Griffiths, T.L., Lieder, F. (2024). Identifying resource-rational heuristics for risky choice. *Psychological Review*, ,
- Kundel, H.L., Nodine, C.F., Conant, E.F., Weinstein, S.P. (2007). Holistic component of image perception in mammogram interpretation: gaze-tracking study. *Radiology*, 242(2), 396–402,
- Lee, M.D. (2011). How cognitive modeling can benefit from hierarchical bayesian models. *Journal of Mathematical Psychology*, 55(1), 1–7,
- Lee, M.D., & Cummins, T.D. (2004). Evidence accumulation in decision making: Unifying the “take the best” and the “rational” models. *Psychonomic Bulletin & Review*, 11(2), 343–352,
- Lewis, M. (2004). *Moneyball: The art of winning an unfair game*. WW Norton & Company.
- Lieder, F., & Griffiths, T.L. (2017). Strategy selection as rational metareasoning. *Psychological Review*, 124(6), 762–794,
- Lieder, F., & Griffiths, T.L. (2020). Resource-rational analysis: Understanding human cognition as the optimal use of limited computational resources. *Behavioral and Brain Sciences*, 43, e1,
- Lieder, F., Griffiths, T.L., Hsu, M. (2018). Overrepresentation of extreme events in decision making reflects rational use of cognitive resources. *Psychological Review*, 125(1), 1–32,
- Lieder, F., Griffiths, T.L., M. Huys, Q.J., Goodman, N.D. (2018). The anchoring bias reflects rational use of cognitive resources. *Psychonomic Bulletin & Review*, 25(1), 322–349,
- Lintusaari, J., Gutmann, M.U., Dutta, R., Kaski, S., Corander, J. (2017). Fundamentals and recent developments in approximate bayesian computation. *Systematic Biology*, 66(1), e66–e82,
- Liu, S., Ullman, T.D., Tenenbaum, J.B., Spelke, E.S. (2017). Ten-month-old infants infer the value of goals from the costs of actions. *Science*, 358(6366), 1038–1041,



- Lucas, C.G., & Griffiths, T.L. (2010). Learning the form of causal relationships using hierarchical bayesian models. *Cognitive Science*, *34*(1), 113–147,
- Marjoram, P., Molitor, J., Plagnol, V., Tavaré, S. (2003). Markov chain monte carlo without likelihoods. *Proceedings of the National Academy of Sciences*, *100*(26), 15324–15328,
- Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. Cambridge: MIT Press.
- Martignon, L., Vitouch, O., Takezawa, M., Forster, M.R. (2003). Naive and yet enlightened: From natural frequencies to fast and frugal decision trees. D. Hardman & L. Macchi (Eds.), *Thinking: Psychological perspectives on reasoning, judgment and decision making* (pp. 189–211). Wiley Online Library.
- Mata, R., von Helversen, B., Rieskamp, J. (2010). Learning to choose: Cognitive aging and strategy selection learning in decision making. *Psychology and Aging*, *25*(2), 299–309,
- Mattar, M.G., & Daw, N.D. (2018). Prioritized memory access explains planning and hippocampal replay. *Nature Neuroscience*, *21*(11), 1609–1617,
- Miller, G.A. (1956). The magical number seven, plus or minus two: Some limits on our capacity for processing information. *Psychological Review*, *63*(2), 81,
- Murphy, K.P. (2023). *Probabilistic machine learning: Advanced topics*. MIT Press. Retrieved from <http://probml.github.io/book2>
- Newall, P.W. (2019). Dark nudges in gambling. *Addiction Research & Theory*, *27*(2), 65–67,
- Newell, B.R., Koehler, D.J., James, G., Rakow, T., Van Ravenzwaaij, D. (2013). Probability matching in risky choice: The interplay of feedback and strategy availability. *Memory & Cognition*, *41*, 329–338,
- Newell, B.R., Weston, N.J., Shanks, D.R. (2003). Empirical tests of a fast-and-frugal heuristic: Not everyone “takes-the-best”. *Organizational Behavior and Human Decision Processes*, *91*(1), 82–96,

- Nosofsky, R.M. (1986). Attention, similarity, and the identification–categorization relationship. *Journal of Experimental Psychology: General*, 115(1), 39–57,
- Oaksford, M., & Chater, N. (2007). *Bayesian rationality: The probabilistic approach to human reasoning*. Oxford University Press, USA.
- Papamakarios, G., Sterratt, D., Murray, I. (2019). Sequential neural likelihood: Fast likelihood-free inference with autoregressive flows. *The 22nd international conference on artificial intelligence and statistics* (pp. 837–848).
- Parpart, P., Jones, M., Love, B.C. (2018). Heuristics as bayesian inference under extreme priors. *Cognitive Psychology*, 102, 127–144,
- Ratcliff, R., Smith, P.L., Brown, S.D., McKoon, G. (2016). Diffusion decision model: Current issues and history. *Trends in Cognitive Sciences*, 20(4), 260–281,
- Ray, D., King-Casas, B., Montague, P., Dayan, P. (2008). Bayesian model of behaviour in economic games. *Advances in Neural Information Processing Systems*, 21, ,
- Rehder, B., Davis, Z.J., Bramley, N. (2022). The paradox of time in dynamic causal systems. *Entropy*, 24(7), 863,
- Rieskamp, J., & Otto, P.E. (2006). Ssl: a theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, 135(2), 207–236,
- Sablé-Meyer, M., Ellis, K., Tenenbaum, J., Dehaene, S. (2022). A language of thought for the mental representation of geometric shapes. *Cognitive Psychology*, 139, 101527,
- Saffran, J.R., Aslin, R.N., Newport, E.L. (1996). Statistical learning by 8-month-old infants. *Science*, 274(5294), 1926–1928,
- Sanborn, A.N. (2017). Types of approximation for probabilistic cognition: Sampling and variational. *Brain and Cognition*, 112, 98–101,

- Sanborn, A.N., & Chater, N. (2016). Bayesian brains without probabilities. *Trends in Cognitive Sciences*, 20(12), 883–893,
- Sanborn, A.N., Griffiths, T.L., Navarro, D.J. (2010). Rational approximations to rational models: alternative algorithms for category learning. *Psychological Review*, 117(4), 1144–1167,
- Shah, A.K., & Oppenheimer, D.M. (2008). Heuristics made easy: an effort-reduction framework. *Psychological Bulletin*, 134(2), 207–222,
- Shanks, D.R., Tunney, R.J., McCarthy, J.D. (2002). A re-examination of probability matching and rational choice. *Journal of Behavioral Decision Making*, 15(3), 233–250,
- Shi, L., & Griffiths, T. (2009). Neural implementation of hierarchical bayesian inference by importance sampling. *Advances in Neural Information Processing Systems*, 22, 1669–1677,
- Simon, H.A. (1945). *Administrative behavior*. Simon and Schuster.
- Simon, H.A. (1955). A behavioral model of rational choice. *The Quarterly Journal of Economics*, 69(1), 99–118,
- Simon, H.A., et al. (1990). Invariants of human behavior. *Annual review of psychology*, 41(1), 1–20,
- Sisson, S.A., Fan, Y., Tanaka, M.M. (2007). Sequential monte carlo without likelihoods. *Proceedings of the National Academy of Sciences*, 104(6), 1760–1765,
- Socher, R., Gershman, S.J., Sederberg, P., Norman, K., Perotte, A., Blei, D. (2009). A bayesian analysis of dynamics in free recall. *Advances in Neural Information Processing Systems*, 22, 1714–1722,
- Song, M., Bnaya, Z., Ma, W.J. (2019). Sources of suboptimality in a minimalistic explore–exploit task. *Nature human behaviour*, 3(4), 361–368,

- Stewart, N., Chater, N., Brown, G.D. (2006). Decision by sampling. *Cognitive Psychology*, 53(1), 1–26,
- Sunnåker, M., Busetto, A.G., Numminen, E., Corander, J., Foll, M., Dessimoz, C. (2013). Approximate bayesian computation. *PLoS Computational Biology*, 9(1), e1002803,
- Szollosi, A., Donkin, C., Newell, B.R. (2023). Toward nonprobabilistic explanations of learning and decision-making. *Psychological Review*, 130(2), 546–568,
- Tavaré, S., Balding, D.J., Griffiths, R.C., Donnelly, P. (1997). Inferring coalescence times from dna sequence data. *Genetics*, 145(2), 505–518,
- Thaler, R.H., & Sunstein, C.R. (2009). *Nudge: The final edition*. Yale University Press.
- Todd, P.M., & Gigerenzer, G. (2012). *Ecological rationality: Intelligence in the world*. OUP USA.
- Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases: Biases in judgments reveal some heuristics of thinking under uncertainty. *Science*, 185(4157), 1124–1131,
- Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, 90(4), 293–315,
- Ullman, T.D., Stuhlmüller, A., Goodman, N.D., Tenenbaum, J.B. (2018). Learning physical parameters from dynamic scenes. *Cognitive psychology*, 104, 57–82,
- Ullman, T.D., & Tenenbaum, J.B. (2020). Bayesian models of conceptual development: Learning as building models of the world. *Annual Review of Developmental Psychology*, 2(1), 533–558,
- Valentin, S., Bramley, N.R., Lucas, C.G. (2023). Discovering common hidden causes in sequences of events. *Computational Brain & Behavior*, 6(3), 377–399,

- Van Opheusden, B., Kuperwajs, I., Galbiati, G., Bnaya, Z., Li, Y., Ma, W.J. (2023). Expertise increases planning depth in human gameplay. *Nature*, 618(7967), 1000–1005,
- Van Rooij, I. (2008). The tractable cognition thesis. *Cognitive Science*, 32(6), 939–984,
- Van Rooij, I., Blokpoel, M., Kwisthout, J., Wareham, T. (2019). *Cognition and intractability: A guide to classical and parameterized complexity analysis*. Cambridge University Press.
- Vul, E., Goodman, N., Griffiths, T.L., Tenenbaum, J.B. (2014). One and done? optimal decisions from very few samples. *Cognitive Science*, 38(4), 599–637,
- Vulkan, N. (2000). An economist’s perspective on probability matching. *Journal of Economic Surveys*, 14(1), 101–118,
- Wang, H., Jedoui, K., Venkatesh, R., Binder, F.J., Tenenbaum, J., Fan, J.E., . . . Smith, K.A. (2024). Probabilistic simulation supports generalizable intuitive physics. *Proceedings of the annual meeting of the cognitive science society* (Vol. 46, pp. 1953–1960).
- Wittkuhn, L., Chien, S., Hall-McMaster, S., Schuck, N.W. (2021). Replay in minds and machines. *Neuroscience & Biobehavioral Reviews*, 129, 367–388,
- Wood, S.N. (2010). Statistical inference for noisy nonlinear ecological dynamic systems. *Nature*, 466(7310), 1102–1104,
- Wozny, D.R., Beierholm, U.R., Shams, L. (2010). Probability matching as a computational strategy used in perception. *PLoS computational biology*, 6(8), e1000871,
- Xu, F., & Tenenbaum, J.B. (2007). Word learning as bayesian inference. *Psychological Review*, 114(2), 245–272,
- Zednik, C., & Jäkel, F. (2016). Bayesian reverse-engineering considered as a research strategy for cognitive science. *Synthese*, 193, 3951–3985,

- Zhang, H. (2004). The optimality of naive bayes. *Proceedings of the the 17th international florida artificial intelligence research society conference* (pp. 562–567).
- Zhao, B., Lucas, C.G., Bramley, N.R. (2022). How do people generalize causal relations over objects? a non-parametric bayesian account. *Computational Brain & Behavior*, 5(1), 22–44,
- Zhu, J.-Q., Chater, N., León-Villagr , P., Spicer, J., Sundh, J., Sanborn, A. (2023). An introduction to psychologically plausible sampling schemes for approximating bayesian inference. K. Fiedler, P. Juslin, & J. Denrell (Eds.), *Sampling in judgment and decision making* (pp. 467–489). Cambridge University Press.
- Zhu, J.-Q., Sanborn, A.N., Chater, N. (2020). The bayesian sampler: Generic bayesian inference causes incoherence in human probability judgments. *Psychological Review*, 127(5), 719–748,
- Zhu, J.-Q., Sundh, J., Spicer, J., Chater, N., Sanborn, A.N. (2023). The autocorrelated bayesian sampler: A rational process for probability judgments, estimates, confidence intervals, choices, confidence judgments, and response times. *Psychological Review*, 131(2), 456–493,