Heuristics and optimal solutions to the breadth–depth dilemma

Rubén Moreno-Bote[a,b,c,d,1], Jorge Ramírez-Ruiz[a,b], Jan Drugowitsch[e], and Benjamin Y. Hayden[f,g,h]

[a]Center for Brain and Cognition, Universitat Pompeu Fabra, 08002 Barcelona, Spain; [b]Department of Information and Communication Technologies, Universitat Pompeu Fabra, 08002 Barcelona, Spain; [c]Serra Húnter Fellow Programme, Universitat Pompeu Fabra, 08002 Barcelona, Spain; [d]Catalan Institution for Research and Advanced Studies–Academia, Universitat Pompeu Fabra, 08002 Barcelona, Spain; [e]Department of Neurobiology, Harvard Medical School, Boston, MA 02115; [f]Department of Neuroscience, University of Minnesota, Minneapolis, MN 55455; [g]Center for Magnetic Resonance Research, University of Minnesota, Minneapolis, MN 55455; and [h]Center for Neural Engineering, University of Minnesota, Minneapolis, MN 55455

Edited by Michael Woodford, Columbia University, New York, NY, and accepted by Editorial Board Member David J. Heeger July 7, 2020 (received for review March 15, 2020)

In multialternative risky choice, we are often faced with the opportunity to allocate our limited information-gathering capacity between several options before receiving feedback. In such cases, we face a natural trade-off between breadth—spreading our capacity across many options—and depth—gaining more information about a smaller number of options. Despite its broad relevance to daily life, including in many naturalistic foraging situations, the optimal strategy in the breadth–depth trade-off has not been delineated. Here, we formalize the breadth–depth dilemma through a finite-sample capacity model. We find that, if capacity is small (~10 samples), it is optimal to draw one sample per alternative, favoring breadth. However, for larger capacities, a sharp transition is observed, and it becomes best to deeply sample a very small fraction of alternatives, which roughly decreases with the square root of capacity. Thus, ignoring most options, even when capacity is large enough to shallowly sample all of them, is a signature of optimal behavior. Our results also provide a rich casuistic for metareasoning in multialternative decisions with bounded capacity using close-to-optimal heuristics.

decision making | risky choice | bounded rationality | breadth–depth dilemma | metareasoning

The breadth–depth (BD) dilemma is a ubiquitous problem in decision making. Consider the example of going to graduate school, where one can enroll in many courses in many topics. Let us assume that the goal is to determine the single area of research that is most likely to result in an important discovery. One cannot know, even in a few weeks of enrollment, whether a course is the most exciting one. Should I enroll in few courses in many topics—breadth search—at the risk of not learning enough about any topic to tell which one is the best? Or should I enroll in many courses in very few topics—depth search—at the risk of not even taking the course with the really exciting topic for the future? One crucial element of this type of decision is that the resources (time, in this case) need to be allocated in advance, before feedback is received (before classes start). Also, once decided, the strategy cannot be changed on the fly, as doing so would be very costly.

The BD dilemma is important in tree search algorithms (1, 2) and in optimizing menu designs (3). It is also one faced by humans and other foragers in many situations, such as when we plan, schedule, or invest with finite resources while lacking immediate feedback. Furthermore, it is a dilemma that a large number of distributed decision-making systems have to tackle. These include, for example, ant scouts searching for a new colony settlement (4), stock market investors, or soldiers in an army during battle. Evidence suggests that distributed processing with limited resources is also a valid model of brain computations (5, 6). In face of this, it is remarkable that the bulk of research on the BD has been in fields outside of psychology and neuroscience (e.g., refs. 7–9). We believe that one reason for this is the lack of models and formal tools for thinking about the BD dilemma and separating it from other dilemmas.

Many features of the BD dilemma warrant its study in isolation. First, BD decisions are about how to divide finite resources, with the possibility of oversampling specific options and ignoring others, e.g., one can select several courses on the same topic while ignoring other topics. Second, the BD dilemma is about making strategic decisions, that is, decisions that need to be planned in advance and cannot be changed on the fly once initiated, e.g., it is very costly to change courses once they have started, at least during the first semester. Finally, BD decisions need to be made before the relevant feedback is received, e.g., enrollment happens before courses start, and thus before knowing the true relevance of the courses and topics. One can easily imagine replacing courses by ant scouts or neurons, and topics by potential new settlements or sensory functions, and so on, in the above example to reveal new relevant BD dilemmas pertaining to distributed decision making or brain anatomy, respectively.

The identifying features of the BD dilemma are distinct from those of the well-known exploitation–exploration (EE) dilemma (10–14) and its associated formalization in multiarmed bandits (15–17). Specifically, whereas in the EE dilemma samples are

Significance

From choosing among the many courses offered in graduate school to dividing budget into research programs, the breadth–depth is a commonplace dilemma that arises when finite resources (e.g., time, money, cognitive capabilities) need to be allocated among a large range of alternatives. For such problems, decision makers need to trade off breadth—allocating little capacity to each of many alternatives—and depth—focusing capacity on a few options. We found that little available capacity (less than 10 samples for search) promotes allocating resources broadly, and thus breadth search is favored. Increased capacity results in an abrupt transition toward favoring a balance between breadth and depth. We finally describe a rich casuistic and heuristics for metareasoning with finite resources.

Author contributions: R.M.-B., J.R.-R., J.D., and B.Y.H. designed research; R.M.-B. and J.R.-R. performed research; R.M.-B. and J.R.-R. analyzed data; and R.M.-B., J.R.-R., J.D., and B.Y.H. wrote the paper.

The authors declare no competing interest.

This article is a PNAS Direct Submission. M.W. is a guest editor invited by the Editorial Board.

[1]To whom correspondence may be addressed. Email: ruben.moreno@upf.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2004929117/-/DCSupplemental.

First published August 5, 2020.

allocated sequentially, one by one, to gather information and reward after each sample, in the BD dilemma multiple samples can be allocated in parallel at once to multiple options (possibly allocating multiple samples to some) without immediate feedback to gather information and maximize future reward. It is worth pointing out that EE and BD are not mutually exclusive aspects of decision making, and therefore they are expected to appear hand-in-hand in many realistic situations.

Past work in multialternative choice has revealed that humans appear to carefully trade off the benefits of examining many options broadly and examining a smaller number of options deeply. For example, when faced with a large number of options, we often focus—even if arbitrarily—on a subset of them (18–21) with the presumable benefit that we can more precisely evaluate them. Likewise, we may consider all options, but arbitrarily reject value-relevant dimensions (22, 23), as if contemplating them all is too costly. Option narrowing appears to be a very general pattern, one that is shared with both human and nonhuman animals, despite the fact that rejecting options can reduce experienced utility (18, 21). It is often proposed that such heuristics reflect bounded rationality (24), which is likely correct in principle, but the exact processes underlying that boundedness remain to be identified. Why do we so often consider a very small number of options when considering more would a priori improve our choice? One possibility is that this pattern reflects an evolved response to an empirical fact: that when capacity is constrained, optimal search favors consideration of a small number of options.

Because cognitive capacity is limited in many ways, the BD dilemma has direct relevance to many aspects of cognition as well. For example, executive control is thought to be limited in capacity, such that control needs to be allocated strategically (25–28). Likewise, attentional focus and working memory capacity are limited, such that, during search, we often foveate only a single target or hold a few items in memory (29). Although the effective numbers are low, each contemplated option is encoded with great detail (30–32). Furthermore, it seems clear that recollection of information from memory can be thought of as a search-like process (33–35). That is, to retrieve a memory we must attend to a recollection processes, with its associated limited capacity. Thus memory-guided decisions presumably involve BD trade-offs too.

Although the relevance of the BD dilemma is clear, tractable models are lacking, and thus, optimal strategies for BD decisions are largely unknown. Here, we develop and solve a model for multialternative decision making endowed with the prototypical ingredients of the BD dilemma. Our model consists of a reward-optimizing yet bounded decision maker (24, 36, 37) confronted with multiple alternatives with unknown subjective values. The first critical element of the model is "finite-sample capacity," which enforces a trade-off between sampling many options with few samples each (breadth) and sampling few options with many samples each (depth). The second critical element is that samples need to be allocated across alternatives before sampling starts and, thus, before feedback is available. This strategic decision with the finite-sample capacity constraint implies a metareasoning problem (37, 38) where deliberation about the multiple possible allocations of resources (meta-actions) need to be made in advance to optimize expected utility of a future choice.

Despite the simplicity of the model, it features nontrivial behaviors, which are characterized analytically. When capacity is low (less than 4 to 10 samples can be probed), it is best to sample as many alternatives as possible, but only once each; that is, breadth search is favored. At larger capacities, there is a qualitative and sharp change of behavior (a "phase transition") and the optimal number of sampled alternatives roughly grows with the square root of sample capacity ("square root sampling law"), balancing breadth and depth. Therefore, in the high-capacity regime, it is best to ignore the vast majority of potentially accessible options. We considered globally optimal allocations in comparison to even allocation of samples across sampled alternatives and found that the square root sampling law, obtained for the latter, provides a close-to-optimal heuristic that is simpler to implement. We also study limit cases where the above rules break down, as well as generalizations to dynamic allocation of finite resources with feedback that illustrate the generality of the results. Our results are also robust to strong variations of the environments where the probability of finding good options widely varies.

Results

Finite-Sample Capacity Model. We assume that a decision maker can choose how to allocate a finite resource among options of unknown status to determine the best option (Fig. 1). The environment generates a large number of options, each characterized by the probability of delivering a successful outcome. The success probabilities, unknown to the decision maker, determine the quality of each of the options, with better options having higher success probabilities (e.g., options with a higher probability of delivering a large reward if they are sampled). The goal of the decision maker is to infer which of the options has the highest success probability (and, thus, highest expected value). The success probabilities of the options are generated randomly from an underlying prior probability distribution, modeled as a beta distribution with parameters (α, β). We assume that this distribution is known by the decision maker, due, for example, to previous experience with the environment. The prior distribution determines the overall difficulty of finding successful options in the environment.

The decision maker is endowed with a finite-sample capacity, C, i.e., a finite number of samples that she can allocate to any option and to as many options as desired. Within the allowed flexibility, it is possible that the decision maker decides to oversample some options by allocating more than one sample to them, and it is also possible that she decides to ignore some options by not sampling them at all. Feedback is not provided at the allocation stage, so this decision is based purely on the expected quality of options in the environment. After allocation has been determined, the outcomes of the samples are revealed, constituting the only feedback that the decision maker receives about the fitness of her sample allocation. Outcomes for each of the sampled alternatives are modeled as a Bernoulli variable, where a successful outcome (corresponding to a large reward) has probability equal to the success probability of that option (see below for a generalization in which we consider Gaussian outcomes). The inferred best alternative is the one with the largest inferred success probability based on the observed outcomes from the allocated samples to each of the options (39–41). Choosing this alternative maximizes expected utility (see below and *SI Appendix*).

While making a choice based on the observed outcomes is a trivial problem, deciding how to allocate samples over the options to maximize expected future reward is a hard combinatorial problem. There are many ways a finite number of samples can be allocated among a very large number of alternatives. At the breadth extreme, one can split capacity to over as many alternatives as possible, sampling each just once. In this case, the decision maker will likely identify a few promising options, but will lack the information for choosing well between them. At the depth extreme, the search could allocate all samples only to a couple of alternatives. The decision maker's estimate of the success probability of those options will be accurate, but that of the other alternatives will remain unknown. It would seem that an intermediate strategy is better than either extreme. Specifically, the optimal allocation of samples should balance the diminishing

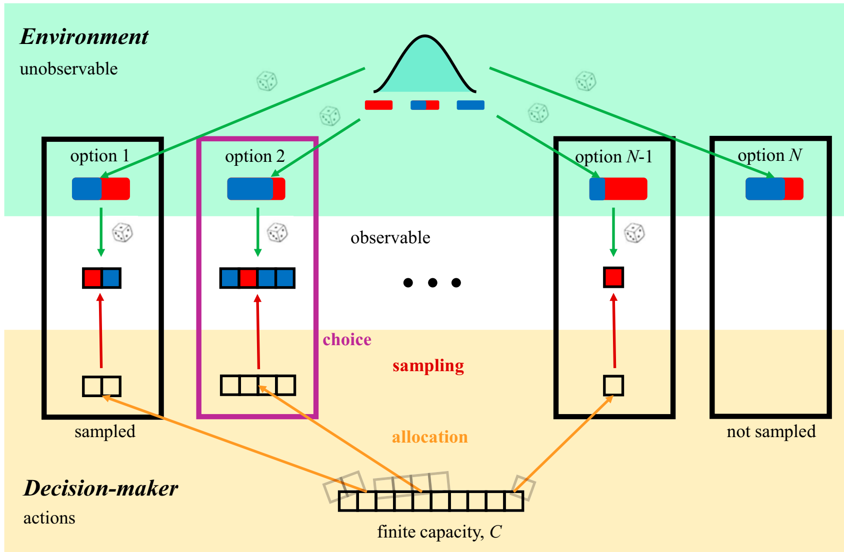

Fig. 1. Finite-sample capacity model. The environment (*Top*, green) contains a large number N of options, and choosing any of them might lead to a successful outcome (e.g., a large vs. a small reward). For each option, the probability of success (blue fraction of red/blue bar) is a priori unknown to the decision maker and is drawn independently across options from an underlying prior probability distribution, modeled as a beta distribution (top distribution). The prior distribution defines the overall difficulty of finding successful options in the environment. Options are characterized by the probability of delivering a successful outcome (e.g., a large reward), and the outcomes are modeled as Bernoulli variables. The decision maker (*Bottom*, orange) has a finite capacity C, i.e., a finite number of samples (bar of squares) that can be allocated to any option in any possible way. The decision maker can decide to oversample options by allocating more than one sample to them (e.g., options on the left), and also ignore some options by not sampling them at all (e.g., rightmost option). All samples need to be allocated in advance, and allocation cannot be changed thereafter. Therefore, feedback is not provided at this stage. After allocation, sampling starts (*Center*, white), in which the decision maker observes a number of successes and failures for each of the sampled options (colored squares; blue: success, large reward, red: failure, small reward). Once this evidence is collected, the decision maker chooses the option that is deemed to have the highest probability of success (in this case, option 2; purple box).

marginal gains of sampling a new alternative and those of drawing an additional sample from an already sampled alternative.

To formalize the above model, let us assume that the decision maker can sample and choose from $N = C$ alternatives. That is, we consider scenarios where the number of alternatives N is as large as the decision maker's sampling capacity—if the number of alternatives is larger than capacity, the only difference is that there would be a larger number of ignored alternatives. The allocation of samples over the alternatives is described by the vector $\vec{L}$, with components L_i representing the number of samples allocated to alternative $i = 1, \ldots, N$. The finite-sample capacity of the decision maker imposes the constraint $\sum_i L_i = C$. Upon drawing L_i samples from each alternative i, the decision maker observes the number of successes (1's), denoted n_i, of each of the Bernoulli variables. The best option is then the one with the highest posterior mean probability $E(p_i|n_i,\alpha,\beta) = (n_i+\alpha)/(L_i+\alpha+\beta)$ after observing these successes, such that the utility for a given allocation $\vec{L}$ and associated outcomes $\vec{n}$ becomes $U(\vec{n},\vec{L}) = \max_i \left[(n_i+\alpha)/(L_i+\alpha+\beta)\right]$. Because the number of successes is only revealed after selecting the sample allocation strategy $\vec{L}$, the decision maker's utility for using that strategy, $U(\vec{L})$, is an average of $U(\vec{n},\vec{L})$ over all possible outcomes given $\vec{L}$,

$$U\left(\vec{L}\right) = \sum_{\vec{n}} p\left(\vec{n}|\vec{L},\alpha,\beta\right) \max_i \frac{n_i+\alpha}{L_i+\alpha+\beta}, \quad [1]$$

where $p(\vec{n}|\vec{L},\alpha,\beta)$ is the joint probability distribution of the outcomes $\vec{n}$ given the allocation $\vec{L}$ and the prior distribution parameters. As each alternative is sampled independently, the distribution of success counts factorizes as $p(\vec{n}|\vec{L},\alpha,\beta) = \prod_i p(n_i|L_i,\alpha,\beta)$, where $p(n_i|L_i,\alpha,\beta)$ is a beta-binomial distribution (42). This distribution specifies the probability of observing exactly n_i successes from a Bernoulli variable that is drawn L_i times, and whose success probability p_i follows a beta distribution with parameters α and β. These two parameters control the skewness of the distribution: If both parameters are equal, the distribution is symmetric around one-half, while for α larger (smaller) than β the distribution is negatively (positively) skewed.

Finally, the optimal allocation of samples across options $\vec{L}^*$ is the one that maximizes the decision maker's expected utility $U(\vec{L})$ in Eq. **1** over all allocations of samples $\vec{L}$,

$$\vec{L}^* = \underset{\vec{L}}{\operatorname{argmax}}\ U\left(\vec{L}\right), \quad [2]$$

with the above finite-sample capacity constraint (see *SI Appendix* for details). The optimal expected utility then becomes $U^* = \max_{\vec{L}} U(\vec{L})$, which involves a double maximization over the expected success probabilities of the sampled alternatives and the allocation of samples over the alternatives, effectively solving the two-stage decision process (i.e., first allocate samples, then observe outcomes, and then choose) in reverse order (i.e., first optimize choices given outcomes and allocation, then optimize allocation).

This maximization allows total flexibility over how many samples to allocate to each alternative. However, for the sake of tractability, let us first consider the best even allocation of samples, that is, a subfamily of allocation strategies where the same number of samples L are allocated to each of M sampled alternatives, while the remaining alternatives $(C - M)$ are not sampled, subject to the standard capacity constraint $M \times L = C$.

Indeed, finding the optimal even allocation of samples is easier than finding the globally optimal allocation, which might be uneven in general (see below). As we show in *SI Appendix*, a particularly simple expression for the optimal even sample allocation, L^*, arises when the prior distribution over success probabilities is uniform ($\alpha = \beta = 1$),

$$L^* = \underset{L}{\operatorname{argmin}} \frac{\sum_{s=0}^{L} (s+1)^M}{(L+1)^M (L+2)}, \qquad [3]$$

where the right-hand side is related to utility by

$$U(M = C/L) = 1 - \frac{\sum_{s=0}^{L} (s+1)^M}{(L+1)^M (L+2)}. \qquad [4]$$

Note that only $M^* = C/L^* \leq C$ alternatives are sampled in the optimal allocation, while the remaining options are given zero samples, thus effectively being ignored. The sampled alternatives can be chosen randomly, as they are indistinguishable before sampling. Using extreme value theory (*SI Appendix*), we show that the optimal number of sampled alternatives M^* and optimal number of samples per alternative L^* both follow a power law with exponent 1/2 for large capacity C:

$$\lim_{C\to\infty} M^* = \sqrt{C}, \quad \lim_{C\to\infty} L^* = \sqrt{C}, \qquad [5]$$

which corresponds to perfectly balancing breadth and depth.

In the next section, we analyze this case in detail. After that, we consider optimal even allocations of samples for arbitrary prior distributions, and finally we provide results for the globally optimal allocations, not necessarily even.

Sharp Transition of Optimal Sampling Strategy at Low Capacity. We first analyze the expected utility $U(M)$ as a function of the number of evenly sampled alternatives M, each sampled L times (such that $M \times L = C$) (Fig. 2*A*). At low capacity ($C = 4$, lighter gray line), the utility increases monotonically from sampling just one alternative ($M = 1$) four times, to sampling four alternatives ($M = 4$) one time each. Thus, a pure breadth strategy is favored. At intermediate capacity ($C = 10$, medium gray line), the maximum occurs at an intermediate number of alternatives (specifically, $M = 5$), reflecting an increasing emphasis on depth. At large capacity ($C = 100$, black line), the maximum expected utility occurs when sampling few different alternatives ($M = 10$ sampled alternatives with $L = 10$ samples each), reflecting a tight balance between breadth and depth. For such large capacities, a breadth search that samples most of the alternatives (rightmost point of the black line) would lead to a reward that approaches 2/3, which is the lowest expected reward one would obtain if at least one sampled alternative has a positive outcome (*SI Appendix*).

The model displays a sharp transition when capacity crosses the critical value of 5 (Fig. 2*B*). Below this critical capacity, the optimal number of sampled alternatives equals capacity. That is, one should follow a breadth strategy and distribute one sample to each alternative. Above 5, the optimal number of sampled alternatives is much smaller than the capacity, with the temporary exception of capacity equal to 7. That is, one should balance the number of sampled alternatives with the depth of sampling each of them. Specifically, the optimal number of sampled alternatives follows a power law with exponent 1/2 (log-log linear regression, power = slope = 0.49, 95% CI = [0.48, 0.50]), as predicted by Eq. **5**, which implies that the fraction of sampled alternatives decreases with the square root of capacity. This means that breadth and depth are tightly balanced in the optimal strategy. The sharp transition at around 5 becomes clearer when plotting the ratio between the optimal number of sampled alternatives and capacity as a function of capacity (Fig. 2*C*).

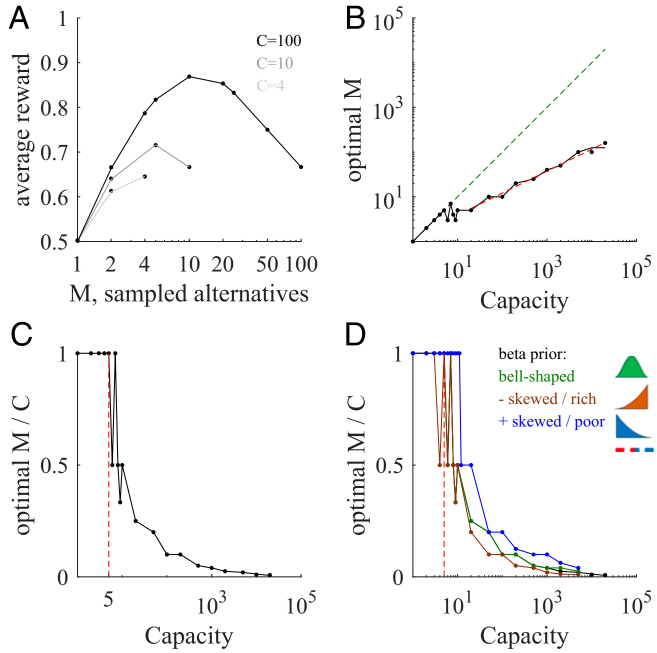

Fig. 2. Sharp transitions in optimal number of sampled alternatives at low capacity, and power law behavior at large capacity. (*A*) Average reward (points, simulations; lines, theoretical expressions; Eq. **4**) as a function of the number of sampled alternatives M for three different capacities ($C = 4, 10, 100$; light, intermediate, and dark lines, respectively) for the flat environment (uniform prior). The maximum occurs at the large extreme for low capacity but at a relatively low value for large capacity. Note log horizontal scale. (*B*) Optimal number of sampled alternatives as a function of capacity. When capacity is smaller than around 5, a linear trend of unit slope is observed (dashed green line), but when capacity is above 7, a sublinear behavior is observed (dashed red line corresponds to the best power law fit, with exponent close to 1/2). The black line corresponds to analytical predictions. The jagged nature of this prediction and simulation lines in this and other panels is due to the discrete values that the optimal M can only take, not due to numerical undersampling. (*C*) The sharp transition is clearer when plotting the optimal number of sampled alternatives to capacity ratio as a function of capacity. For low capacity, the ratio is 1, but for large capacity the ratio decreases very rapidly. The last point below which the optimal ratio is always 1 (critical capacity) corresponds to capacity equal to 5 (indicated with a vertical red line). (*D*) Number of sampled alternatives to capacity ratios for different prior distributions ($\alpha = \beta = 3$, bell-shaped, green line; $\alpha = 3, \beta = 1$, negatively skewed prior modeling a "rich" environment, brown line; $\alpha = 1, \beta = 3$, positively skewed distribution modeling looking for a "needle in a haystack," that is, a "poor" environment, blue line). Lines correspond to analytical predictions from *SI Appendix*, Eq. **9**; points correspond to numerical simulations; error bars are smaller than data points in all panels.

In summary, if the capacity of a decision maker increases by a factor of 100, the decision maker will roughly increase the number of samples alternatives just by a factor of 10, one order of magnitude smaller than the capacity increase. Because the optimal number of sampled alternatives increases with capacity with an exponent 1/2, we call this the "square root sampling law." A remarkable implication of this law is that the vast majority of potentially accessible alternatives should be ignored (e.g., for $C = 100$, $C - M = 90$ options are "rationally" ignored).

Generalizing to Variations in Beta Prior Distributions. The above critical capacity for optimal even sample allocation changes when, instead of using a uniform prior of success probabilities, we allow for variations of the prior distribution (Fig. 2*D*). However, the critical capacity consistently lies again at around low values (~10) with the specific value depending on the environment. By changing the prior's parameters, we can vary the difficulty of finding a good extreme alternative, and thus can compare different scenarios. For the uniform prior that we have used previously (a "flat" environment), a decision maker is equally likely to find an alternative with any success probability. Consider a prior distribution that is concentrated and symmetric around a success probability of 0.5 (approximately as a Gaussian, corresponding to the beta prior parameters $\alpha = \beta = 3$). In this environment, unusually good (high success probability) and unusually bad (low success probability) options are rarer than medium ones (Fig. 2*D*, green line). In this case, the BD trade-off as a function of C is remarkably similar to the uniform prior case, with a transition at $C = 5$.

We also consider a negatively skewed prior distribution ($\alpha = 3, \beta = 1$). This distribution refers to environments with rare bad options, as, for example, a tree whose fruits are mostly ripe but that has a few unripe ones. In this "rich" environment, one can afford sampling a smaller number of options, and as they are sampled more deeply, it is possible to better detect the really excellent ones. A sharp transition occurs even in this condition, exactly when the critical capacity equals 3 (brown line). As expected in this environment, the decay of the ratio between the optimal number of sampled alternatives and capacity after this transition is (slightly) faster than that of the symmetric prior. Therefore, negative skews engender a modest bias toward depth over breadth.

Finally, consider the opposite scenario, in which the prior distribution is concentrated at low success probability values ($\alpha = 1, \beta = 3$, positively skewed beta distribution), which corresponds to looking for a "needle in a haystack" or a "poor" environment. In this scenario, one ought to sample more alternatives less deeply to allow for the possibility of finding the rare good alternatives, and thus breadth should be emphasized over depth (Fig. 2*D*, blue line). In this scenario, the sharp transition occurs at capacities around 10 (blue line).

Despite the large variations of prior distributions, a fast transition occurs in all conditions at around a small capacity value, like in the uniform prior case. In addition, a power law behavior is observed at larger values of capacity regardless of skew, with exponents close to 1/2 in all cases (uniform prior, exponent = 0.49; negatively skewed prior, 0.49; positively skewed prior, 0.64; SEM = 0.01). These behaviors are observed over a larger range of parameters of the prior distribution (Fig. 3).

One interesting limit scenario arises for strongly positively skewed prior distributions, e.g., by taking β to infinity while fixing $\alpha = 1$. In this limit, the prior mean probability $\alpha/(\alpha + \beta)$ decreases to zero, and the critical capacity rises very steeply to infinity (Fig. 3*A* as one moves leftward). Increasing the prior's skewness makes finding good options less likely, as most of the options are very likely to be very bad, akin to an extreme case of the haystack environment considered before. As expected, this makes breadth search optimal for increasingly larger values of capacity, as indicated by the increasing values of critical capacity. However, for large enough capacities, a transition is still observed above which a roughly balanced mix between breadth and depth becomes optimal. More precisely, in this regime the optimal number of sampled alternatives features a power behavior with exponents close but above 1/2, indicating a bias toward breadth (leftmost points in Fig. 3*B*). When the prior mean probability exceeds values as low as 0.1, the critical capacity plateaus to low values below 10, and the exponent drops to values smaller but close to 1/2, indicating a weak preference toward depth.

To test how robust the behaviors we explored are, we furthermore considered Gaussian rather than Bernoulli samples (*SI Appendix*, Fig. S1). Strikingly, for a large range of the samples' noisiness, we again observed a sharp transition occurring at low critical capacities (~ 10). Below the critical capacity, breadth search is preferred, while above it a mix between breadth and depth is optimal, characterized by a power law behavior (exponent = 0.35, 95% CI = [0.30, 0.41]). Thus, the resulting strategy was qualitatively identical, and numerically similar, to the Bernoulli samples case.

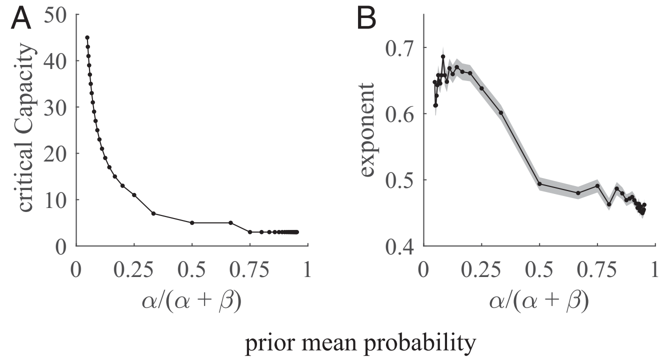

Fig. 3. Sharp transitions at relatively low-capacity values and close to square root sampling behavior for large capacity are observed for a broad range of parameters of the prior distribution. (*A*) Critical capacity decreases very steeply to low values (~10) as a function of the prior mean probability $\alpha/(\alpha + \beta)$. (*B*) Exponents decrease as a function of the prior mean probability and cluster around 1/2. The exponents are obtained from log-log linear regression fits of the optimal number of alternatives samples vs. capacity ($M \propto C^{exponent}$) for C values ranging from 1,000 to 2,000 in steps of 1. Shaded areas correspond to 95% CIs. In both panels, points are obtained by theoretical predictions from *SI Appendix*, Eq. 9. For prior mean probabilities smaller than or equal to 1/2, we fix $\alpha = 1$ while β varies from 1 to 20 in steps of 1, and for values larger than 1/2, we fix $\beta = 1$ while α varies from 1 to 20.

Optimal Choice Sets and Sample Allocations. So far, we have focused on optimal even sample allocation. Let us now consider the payoffs for decision makers willing to consider all possible allocation strategies. The number of all possible allocations equals the number of partitions of integers in number theory, which grows exponentially with the square root of capacity (43). This makes finding the globally optimal sample allocation a problem that is intractable in general. For small capacity values $C \leq 7$ and uniform prior distributions, we compute the exact optimal sample allocation by exhaustive search and rely on a stochastic hill climbing method for larger capacities and other priors. The latter finds a local maximum for the utility, which is likely to be a global maximum, as we found it to coincide with the one provided by exhaustive search for small capacities $C \leq 7$, and the optimal utility did not significantly change across different initializations and random seeds for larger capacity values.

Globally optimal sample allocation (which defines optimal choice sets) for a uniform prior beta distribution tends to sample all or most of the alternatives when the capacity is small, but as capacity increases the number of sampled alternatives decrease (Fig. 4, *Left*). For instance, for capacity equal to 5 samples, the optimal sample allocation is (2, 1, 1, 1, 0). In general, in optimal allocations, the decision maker adopts a local balance between oversampling a few alternatives and sparsely sampling others—a local compromise between breadth and depth—even though all options are initially indistinguishable. This further level of specialization and distinction between alternatives might be able to better break ties between similar alternatives when compared to an even sampling strategy.

We also studied optimal sample allocation for positively and negatively skewed prior distributions. In a rich environment

PSYCHOLOGICAL AND COGNITIVE SCIENCES
NEUROSCIENCE

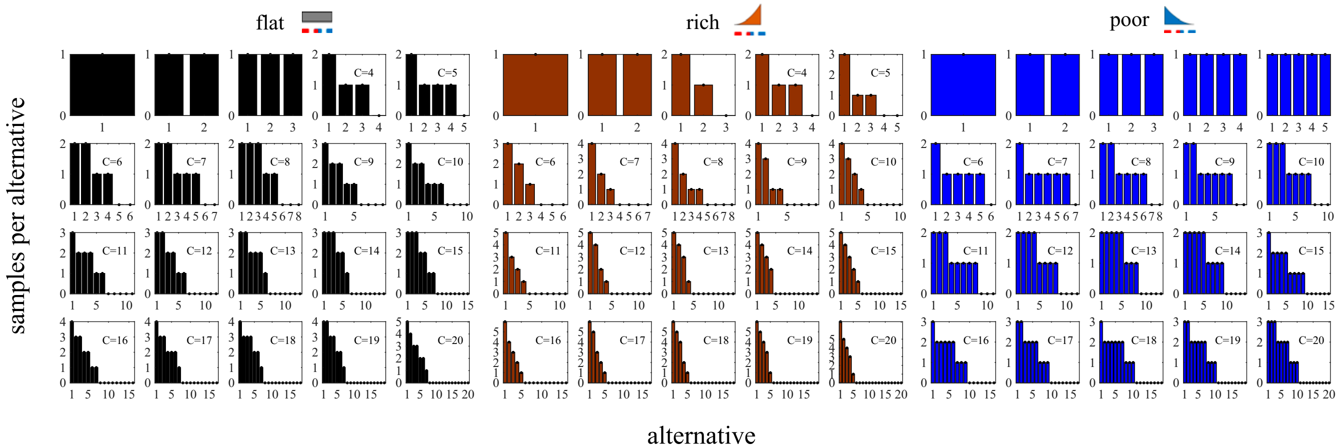

Fig. 4. Optimal sample allocations and choice sets. Optimal sample allocation for flat, rich, and poor environments from capacity $C = 1$ up to $C = 20$. The environments correspond, respectively, to uniform, negatively and positively skewed prior distributions (top icons). Optimal sample allocations are represented as bar plots, indicating the number of samples allocated to each alternative ordered from the most to the least sampled alternative.

(center panel), the optimal sample allocation is uneven for capacities as low as $C = 3$. In contrast, in a poor environment (right), the optimal sample allocation remains even up to capacity $C = 5$, which was not the case for the flat environment (compare with left panel). For higher capacities, decision makers in rich environments ought to sample less broadly but more deeply. For instance, for capacity $C = 20$, only around 5 alternatives are sampled, while the remaining 15 potentially accessible alternatives are neglected. In the poor environment, in contrast, about half of the alternatives are sampled, but not very deeply (only a maximum of 3 samples are allocated to the most sampled alternatives).

Even Sample Allocation Is Close to Optimal. Three principles stand out. First, globally optimal sample allocation almost never coincides with optimal even allocation. Second, at low capacity optimal allocation favors breadth, while at large capacity a BD balance is preferred (Fig. 5*A*). Third, a fast transition is observed between the two regimes happening at a relatively small capacity value. The last two features are shared by the optimal even allocations as well (cf. Fig. 2*C*).

Optimal even and globally optimal sample allocations share some important features, but are they equally good in terms of average reward obtained? We compared the average reward from globally optimal and even optimal sample allocations. For comparison, we always used even sampling based on a uniform prior over each alternatives' success probabilities, that is, we sample $M = C$ alternatives with one sample each if capacity is $C \leq 7$ and $M = \sqrt{C}$ alternatives with $L = \sqrt{C}$ samples each if capacity is larger (square root sampling law; see *SI Appendix* for details). This heuristic produced comparable performances to the optimal ones (Fig. 5*B*). The worst-case scenario occurred in the poor environment (blue line) when capacity is close to 10, which led to a drop in reward by close to 10%, but the maximum discrepancy value was even smaller for the flat and rich environments. Indeed, for the flat environment, the maximum drop in reward was only around 5%.

For large capacity $C > 100$, the square root sampling law produced results that were very close to the performance of the optimal solutions (as found by stochastic hill climbing). Therefore, the gain of globally optimal sample allocation over optimal even sampling at low capacity, and over the square root sampling law for high capacity, is at most marginal.

We also compared the merits of the square root sampling law to other sensible heuristics: pure breadth, pure depth, random sampling of options, and a triangular approximation. Pure breadth search allocates just one sample per alterative, such that

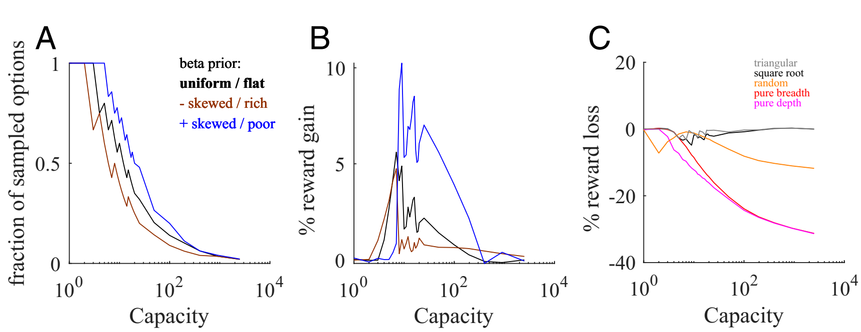

Fig. 5. Globally optimal and optimal even sample allocations share similar features and have similar performances. (*A*) Fraction of sampled options (compared to the maximum number of potentially accessible alternatives, equal to *C*) as a function of capacity *C*. The fraction is close to 1 for small values for all environments (flat, black line; rich, brown; and poor, blue). The fraction decays rapidly to zero from a critical value that depends on the prior. The jagged nature of the lines is due to the discrete nature of capacity. (*B*) Percentage increase (gain) in averaged reward by using globally optimal sample allocation compared to even allocation (see *SI Appendix*). Color code as in the previous panel. (*C*) Percentage loss in averaged reward by using triangular (gray), square root sampling law (black), random (orange), pure breadth (red), and pure depth (pink) heuristics compared to optimal allocation, in a flat environment. For the square root and triangular heuristics, we used pure breadth search when $C \leq 5$.

the number of sampled alternatives equals capacity. The pure depth heuristic randomly chooses two alternatives that are each allocated half of the sampling capacity. Random search randomly assigns each of the C samples to any alternative with replacement. A final heuristic, called "triangular," is inspired by the seemingly isosceles right triangle shape of the optimal allocations (Fig. 4). It splits capacity by sampling the first alternative with $\lfloor 2\sqrt{C} - 1 \rfloor$ samples and any further alternative with one sample less than the previous one until capacity is exhausted ($\lfloor x \rfloor$ is the floor function). All heuristics finally choose the alternative with the highest posterior mean probability. While the loss relative to optimal allocation is smallest for triangular allocation, the square root sampling law performs similarly, and much better than random, pure breadth and pure depth heuristics (Fig. 5*C*).

Dynamic Allocation of Capacity. Thus far, we have considered "static" allocations whereby no feedback is provided before all samples are allocated. In a less rigid "dynamic" sample allocation strategy, some basic form of interim feedback might be available, based upon which further alternatives can be sampled. To model such a scenario, assume that the capacity can be divided into a sequence of a maximum of C waves $k = 1, \ldots, C$, such that in each wave a number of alternatives M_k, no larger than in the previous wave, is sampled just once. The number of alternatives sampled at each wave can be chosen freely, but has to be allocated before sampling starts, that is, the decision maker has to determine the policy at the start knowing she will receive feedback in the future. However, to dynamically react to past sampling outcomes, the kth wave allocates its M_k samples to only those M_k alternatives with the largest number of successes so far (with random allocation in case of ties). This implies that, in wave $k+1$, one can only sample a subset of the alternatives sampled in wave k. Once sampling has been completed across all waves, the alternative with the highest posterior mean probability is chosen among the M_1 sampled alternatives in the first wave. We restricted the final choice to this initial set of alternatives sampled in the first wave to handle the unlikely case that the lastly sampled alternatives turned out to be worse than (our a priori belief about) the initially sampled ones. In that case, the dynamic strategy could lead to worse performance than the static one. We call the above flexible allocation of the predefined sample waves dynamic sample allocations. As for static allocations, we find the best-performing $M_1, M_2, \ldots$sequence by stochastic hill climbing.

Optimal dynamic sample allocations share many features with optimal static sample allocations (Fig. 6). At low capacity, pure breadth search is again optimal. That is, it is best to allocate all samples in the first wave, assigning just one sample per alternative (Fig. 6*A*). For capacities larger than the critical capacity $C = 3$, it is best to mix breadth with depth search, and for very large capacity most accessible alternatives are again ignored. The optimal dynamic and static sample allocations have, however, important differences (Fig. 6*B* and cf. Fig. 4). Specifically, the initial wave tends to sample many alternatives to identify good ones, and follow-up waves narrow down the search to the potentially best ones. This results in broader sample allocations (Fig. 6*C*) that, overall, sample more alternatives than for static allocations (cf. Fig. 4). Finally, we test how the static square root sampling law performs against the optimal dynamic allocations, finding that the former is worse by less than 9% for all capacity values (Fig. 6*D*). We also confirm that static random, pure breadth, and pure depth strategies are substantially worse than the square root sampling law, while the triangular strategy is similar to the simple square root sampling heuristic.

Discussion

We delineate a formal mathematical framework for thinking about a commonplace decision-making problem. The BD dilemma occurs when a decision maker is faced with a large set of possible alternatives, can query multiple alternatives simultaneously with arbitrary intensities, and has overall a limited search capacity. In such situations, the decision maker will often have to balance between allocating search capacity to more (breadth) or to fewer (depth) alternatives. We develop and use a finite-sample capacity model to analyze optimal allocation of samples as a function of capacity. The model displays a sharp transition of behavior at a critical capacity corresponding to just a small set of available samples (~10). Below this capacity, the optimal strategy is to allocate one sample per alternative to access as many alternatives as possible (i.e., breadth is favored). Above this capacity, BD balance is emphasized, and the square root sampling law, a close-to-optimal heuristic, applies. That is, capacity should be split into a number of alternatives equal to the square root of the capacity. This heuristic provides average rewards that are close to those from the optimal allocation of samples. As it is easy to implement, it can become a general rule of thumb for strategic allocation of resources in multialternative choices. The same results roughly apply to a wide variety of environments, including flat, rich, and poor ones, characterized by very different difficulties of finding good alternatives.

Despite the billions of neurons in the brain, our processing capacity seems quite limited. This strict limit applies to attention, where it is sometimes called the attentional bottleneck (44–46), including spatial attention, where the limit is best characterized (47), over working memory (29, 31, 32, 48–50), to executive control (28, 51, 52), and to motor preparation (53). These narrow limits, which often number only a handful of items (although see ref. 32), suggest some sort of bottleneck. However, another interpretation is that capacity is much larger than it appears, and instead, observed capacity reflects the strategic allocation of resources according to the compromises that our model identifies as optimal. The square root sampling law, in other words, suggests that the apparently narrow bandwidth of cognition may reflect the optimal allocation across very few alternatives of a relatively large capacity.

This is particularly likely to be true for economic choice. We are especially interested in the apparent strict capacity limits of prospective evaluation (54–58). Indeed, the failures of choice with choices sets over a few items are striking and have been a major part of the heuristic literature (59, 60). These strict limits are ostensibly difficult to explain. They do not appear to derive, for example, from the basic computational or biophysical properties of the nervous system, as is evident from the fact that our visual systems are an exception to the general pattern and can process much information in parallel. Nor do these limits appear to relate to any desire to reduce the extent of computation, as large numbers of brain regions coordinate to implement these cognitive processes (61–64). Our results presented above offer an appealing explanation for this problem: Economic choice can be construed as BD search problems, and even when capacity is large, the optimal strategy is to focus on a very small region of the search space. Thus, our results can also help to understand why many cognitive systems operate in a regime of low sampling size, thus resolving the paradox of why low breadth sampling and large brain resources can coexist.

We believe that these results are particularly relevant to behavioral economics. Research has shown that consumers often consider just a small number of brands from where to purchase a specific product out of the many brands that exist in the market (65, 66). The prevailing notion is that decision makers hold a consideration choice set from where to make a final choice rather than contemplating all possibilities. Several reasons for

PSYCHOLOGICAL AND COGNITIVE SCIENCES

NEUROSCIENCE

this behavior have been provided. First, choice overload has been shown to produce suboptimal choices in certain conditions (60, 67). Second, selecting a small number of options from where to choose can be actually optimal if there is uncertainty about the value of the options and there is cost for exploring and sampling further options (68–70). Estimating the overall benefits of considering larger sets has to be balanced with the associated cost of exploring further options.

This research has provided a relevant line of thought for understanding low sampling behavior within the context of bounded rationality by formally assuming the presence of linear costs of time for searching for new options. Time costs come in their models at the expense of unknown parameters, which often are difficult to fit (68, 69). Furthermore, linear time costs always permit unlimited number of sampled options, as they do not impose a strict limit in the number of options that can be sampled. In our approach, in contrast, allocating finite resources imposes a strict limit to the number of options that can be sampled and, as resources are limited, there is a trade-off between sampling more options with less resources or sampling fewer options with more resources, directly addressing the BD dilemma. This difference could be the main reason why the consideration set literature has not reported sharp transitions of behavior as a function of model parameters (costs) nor power sampling laws, which are the main features of our finite-sample capacity model.

A number of extensions would be required to fully address more realistic problems associated to the BD dilemma. So far, we have considered a two-stage decision process, where the first metareasoning decision is about optimally distributing limited sampling capacity. We have also considered a sequential problem where some basic form of feedback can be used, but the allocation strategy needs to be chosen before the gathering of information and remains fixed thereafter. By construction, these optimal dynamic allocations at large capacity sample more deeply those alternatives that have largest values, in line with experimental work (55, 71). Perhaps a more relevant observation is that the depth of processing of the best alternatives increases with capacity and that more samples are allocated to the top alternatives than for optimal static sample allocations (cf. Fig. 6). Furthermore, if capacity increases, relatively more samples are allocated to the most-sampled than the second-most–sampled alternative. Both predictions are currently untested.

It would be interesting to extend these results to truly sequential processes where the decision of how many samples to allocate per wave is flexible and depends on intermediate feedback. An advantage of this more general setup (72) is that a full-fledged interaction between the BD and EE dilemmas could be studied. In particular, a relevant direction is relating our square root sampling law with Hick's law (73) for multialternative choices. The two approaches touch different aspects of multialternative decision making: While Hick's law refers to the problem of how long options should be sampled in a multialternative setting, it does so by sampling all available options; the square root sampling law, by contrast, applies to situations where there are many alternatives and a large fraction of them are to be ignored due to limited capacity, directly facing the BD

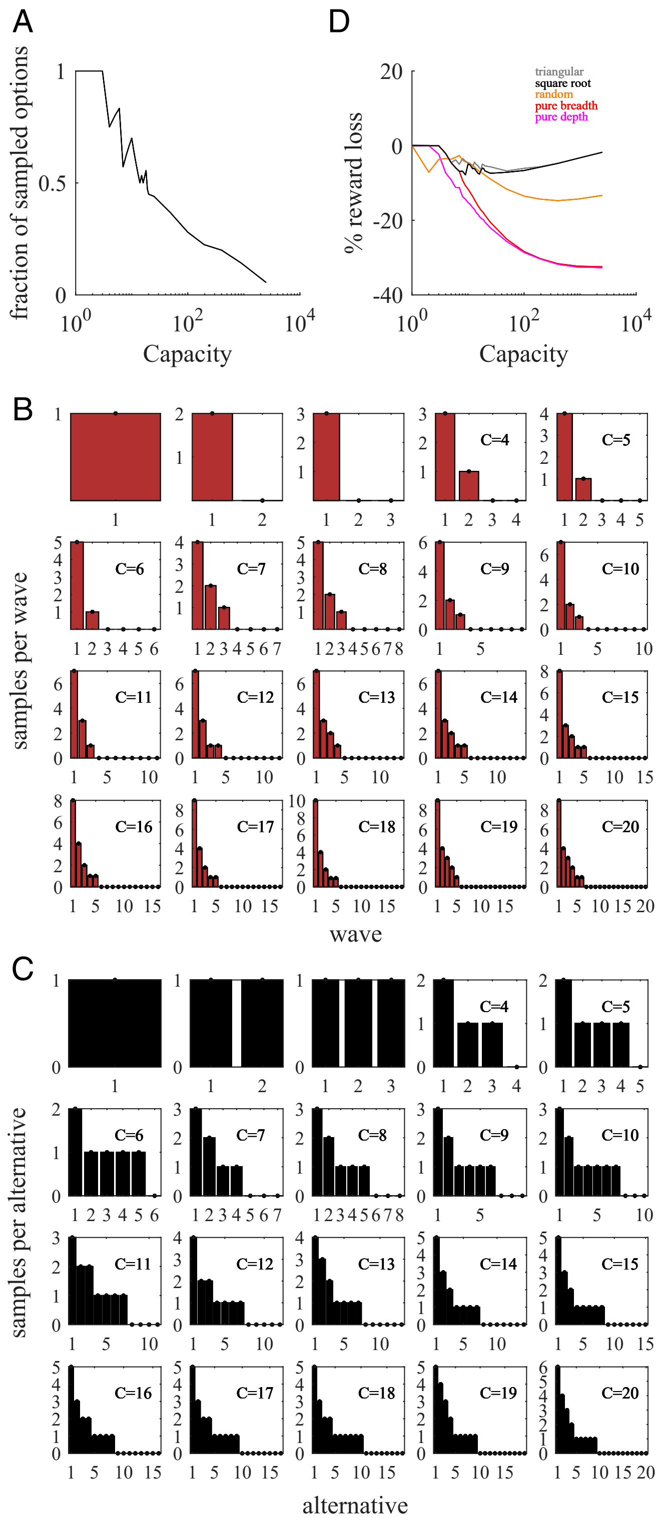

Fig. 6. Optimal dynamic sample allocations display a sharp transition at low capacity, distribute samples unevenly across alternatives, and ignore a vast number of alternatives at high capacity. (*A*) Fraction of sampled alternatives (compared to the maximum number of potentially accessible alternatives, equal to *C*) as a function of capacity *C* for the flat environment (uniform prior). The fraction is 1 for small-capacity values and decays rapidly to zero at large capacity. (*B*) Optimal sample waves, indicating the number of samples allocated in each wave. The number of samples allocated in each wave lies between 1 and *C*, and they sum up to the total available capacity *C*. The maximum allowed number of waves is *C*. (*C*) Optimal dynamic sample allocations and choice sets after the whole capacity has been allocated through the sample waves. The alternatives with largest number of successes are allocated a higher number of samples compared to static allocations (cf. Fig. 4). Many alternatives are given just one sample, typically arising from the first wave, which produces broader sample allocations compared to static allocations. (*D*) Percentage loss in averaged reward by using triangular (gray), square root sampling law (black), random (orange), pure breadth (red), and pure depth (pink) static heuristics compared to optimal dynamic allocation.

dilemma. It will be interesting to integrate the two sets of results within a general framework of multialternative sequential sampling (74–76) under limited resources.

A second possible extension of our work is reconsidering the nature of capacity. For instance, "rate distortion theory" defines a natural capacity constraint over the mutual information between the inputs and the outputs in a system (77, 78). This capacity constraint might more naturally enforce a finite capacity than fixing the total number of samples that a system can draw from (externally or internally). A third relevant direction would be extending our study to cases where the capacity is continuous rather than discrete, and to cases where the observations are continuous variables. Showing that Gaussian rather than Bernoulli outcomes yield qualitatively similar strategies is a first step in this direction. Although it remains a topic for future research, we do not expect qualitative differences in behavior in other continuous settings, as for large capacity the continuous limit approximation applies, and for low capacities the optimality of low number of sampled alternatives is expected.

While we do not know of direct tests of BD capacities in humans, indirect measurements suggest that the square root sampling law can be at work in some realistic conditions, such as chess. It has been argued that chess players can image around 100 moves before deciding their next move (79). Assuming that their capacity is 100, then the square root sampling law would predict that players should sample 10 immediate moves followed by around 10 continuations. Indeed, estimates indicate that chess players mentally contemplate roughly between 6 and 12 immediate moves followed by their continuations (79) before capacity is exhausted due to time pressure. Although decisions in trees like this surely involve other types of search heuristics beyond balancing breadth and depth, the quantitative similarity between predictions and observations is intriguing.

Finally, our work potentially opens ways to understand confirmation biases. Confirmation biases happen when people extensively sample too few alternatives, thus effectively seeking information for the same source. We have demonstrated that oversampling some alternatives and completely ignoring others is optimal in certain conditions. It remains to be seen, however, whether this is actually the optimal strategy under more general conditions or whether the oversampling strategy induces severely harmful biases in certain niches.

It is important to note that we have described the phenomenology of the BD dilemma in conditions where all alternatives are, a priori, equally good. Thus, ignoring a large fraction of options and the associated square root sampling law can only be the worst-case scenario, in the sense that if there are biases or knowledge that a subset of alternatives is initially better than the rest, then fewer number of alternatives should be sampled. This consideration reassures us in the conclusion that the number of alternatives that ought to be sampled is much smaller than sampling capacity, an observation that might turn out to be of general validity in both decision-making setups as well as in terms of brain organization for cognition.

Methods

A detailed description of the finite-capacity model, a derivation of Eqs. 1–5, and a description of the numerical methods used to generate the figures can be found in *SI Appendix*.

Data Availability. The data that support the findings of this study, as well as the codes used for analysis and to generate figures, are publicly available in GitHub at https://github.com/rmorenobote/breadth-depth-dilemma.

ACKNOWLEDGMENTS. This work is supported by the HHMI (Grant 55008742), Ministry of Economic Affairs and Digital Transformation (Spain) (Grant BFU2017-85936-P), and Catalan Institution for Research and Advanced Studies–Academia (2016) to R.M.-B.; by NIH (Grant DA037229) to B.Y.H.; and by a scholar award from the James S. McDonnell Foundation (Grant 220020462) to J.D.

1. E. Horowitz, S. Sahni, *Fundamentals of Computer Algorithms*, (Computer Science Press, 1978).
2. R. E. Korf, Depth-first iterative-deepening. *Artif. Intell.* **27**, 97–109 (1985).
3. D. P. Miller, "The depth/breadth tradeoff in hierarchical computer menus" in *Proceedings of the Human Factors Society Annual Meeting*, (The Human Factors Society, 1981), Vol. vol. 25, pp. 296–300.
4. S. Pratt, E. Mallon, D. Sumpter, N. Franks, Quorum sensing, recruitment, and collective decision-making during colony emigration by the ant *Leptothorax albipennis*. *Behav. Ecol. Sociobiol.* **52**, 117–127 (2002).
5. P. P. Balasubramani, R. Moreno-Bote, B. Y. Hayden, Using a simple neural network to delineate some principles of distributed economic choice. *Front. Comput. Neurosci.* **12**, 22 (2018).
6. B. R. Eisenreich, R. Akaishi, B. Y. Hayden, Control without controllers: Toward a distributed neuroscience of executive control. *J. Cogn. Neurosci.* **29**, 1684–1698 (2017).
7. H. Halpert, Folklore: Breadth versus depth. *J. Am. Folkl.* **71**, 97 (1958).
8. M. S. Schwartz, P. M. Sadler, G. Sonnert, R. H. Tai, Depth versus breadth: How content coverage in high school science courses relates to later success in college science coursework. *Sci. Ed.* **93**, 798–826 (2009).
9. S. F. Turner, R. A. Bettis, R. M. Burton, Exploring depth versus breadth in knowledge management strategies. *Comput. Math. Organ. Theory* **8**, 49–73 (2002).
10. J. D. Cohen, S. M. McClure, A. J. Yu, Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **362**, 933–942 (2007).
11. V. D. Costa, A. R. Mitz, B. B. Averbeck, Subcortical substrates of explore-exploit decisions in primates. *Neuron* **103**, 533–545.e5 (2019).
12. N. D. Daw, J. P. O'Doherty, P. Dayan, B. Seymour, R. J. Dolan, Cortical substrates for exploratory decisions in humans. *Nature* **441**, 876–879 (2006).
13. R. B. Ebitz, E. Albarran, T. Moore, Exploration disrupts choice-predictive signals and alters dynamics in prefrontal cortex. *Neuron* **97**, 450–461.e9 (2018).
14. R. C. Wilson, A. Geana, J. M. White, E. A. Ludvig, J. D. Cohen, Humans use directed and random exploration to solve the explore-exploit dilemma. *J. Exp. Psychol. Gen.* **143**, 2074–2081 (2014).
15. B. B. Averbeck, Theory of choice in bandit, information sampling and foraging tasks. *PLoS Comput. Biol.* **11**, e1004164 (2015).
16. W. Chen *et al*., Combinatorial multi-armed bandit with general reward functions. *Adv. Neural Inf. Process. Syst.* **29**, 1659–1667 (2016).
17. J. C. Gittins, R. Weber, K. D. Glazebrook, *Multi-Armed Bandit Allocation Indices*, (Wiley, ed. 2, 2011).
18. A. Tversky, Elimination by aspects: A theory of choice. *Psychol. Rev.* **79**, 281–299 (1972).
19. J. R. Bettman, M. F. Luce, J. W. Payne, Constructive consumer choice processes. *J. Consum. Res.* **25**, 187–217 (1998).
20. E. Brandstätter, G. Gigerenzer, R. Hertwig, The priority heuristic: Making choices without trade-offs. *Psychol. Rev.* **113**, 409–432 (2006).
21. G. Gigerenzer, W. Gaissmaier, Heuristic decision making. *Annu. Rev. Psychol.* **62**, 451–482 (2011).
22. D. Timmermans, The impact of task complexity on information use in multi-attribute decision making. *J. Behav. Decis. Making* **6**, 95–111 (1993).
23. J. R. Busemeyer, S. Gluth, J. Rieskamp, B. M. Turner, Cognitive and neural bases of multi-attribute, multi-alternative, value-based decisions. *Trends Cognit. Sci.* **23**, 251–263 (2019).
24. H. A. Simon, A behavioral model of rational choice. *Q. J. Econ.* **69**, 99 (1955).
25. T. T. Hills, P. M. Todd, R. L. Goldstone, The central executive as a search process: Priming exploration and exploitation across domains. *J. Exp. Psychol. Gen.* **139**, 590–609 (2010).
26. E. Koechlin, C. Summerfield, An information theoretical approach to prefrontal executive function. *Trends Cognit. Sci.* **11**, 229–235 (2007).
27. A. Shenhav, L. F. Barrett, M. Bar, Affective value and associative processing share a cortical substrate. *Cogn. Affect. Behav. Neurosci.* **13**, 46–59 (2013).
28. A. Shenhav *et al*., Toward a rational and mechanistic account of mental effort. *Annu. Rev. Neurosci.* **40**, 99–124 (2017).
29. N. Cowan *et al*., On the capacity of attention: Its estimation and its role in working memory and cognitive aptitudes. *Cognit. Psychol.* **51**, 42–100 (2005).
30. E. Awh, B. Barton, E. K. Vogel, Visual working memory represents a fixed number of items regardless of complexity. *Psychol. Sci.* **18**, 622–628 (2007).
31. S. J. Luck, E. K. Vogel, Visual working memory capacity: From psychophysics and neurobiology to individual differences. *Trends Cognit. Sci.* **17**, 391–400 (2013).
32. W. J. Ma, M. Husain, P. M. Bays, Changing concepts of working memory. *Nat. Neurosci.* **17**, 347–356 (2014).
33. T. T. Hills, M. N. Jones, P. M. Todd, Optimal foraging in semantic memory. *Psychol. Rev.* **119**, 431–440 (2012).
34. R. Ratcliff, B. B. Murdock, Retrieval processes in recognition memory. *Psychol. Rev.* **83**, 190–214 (1976).
35. M. N. Shadlen, D. Shohamy, Decision making and sequential sampling from memory. *Neuron* **90**, 927–939 (2016).

36. S. J. Gershman, E. J. Horvitz, J. B. Tenenbaum, Computational rationality: A converging paradigm for intelligence in brains, minds, and machines. *Science* **349**, 273–278 (2015).
37. T. L. Griffiths, F. Lieder, N. D. Goodman, Rational use of cognitive resources: Levels of analysis between the computational and the algorithmic. *Top. Cogn. Sci.* **7**, 217–229 (2015).
38. S. Russell, E. Wefald, Principles of metareasoning. *Artif. Intell.* **49**, 361–395 (1991).
39. R. E. Bechhofer, R. V. Kulkarni, Closed sequential procedures for selecting the multinomial events which have the largest probabilities. *Commun. Stat. Theory Methods* **13**, 2997–3031 (1984).
40. S. S. Gupta, T. Liang, Selecting the best binomial population: Parametric empirical Bayes approach. *J. Stat. Plan. Inference* **23**, 21–31 (1989).
41. M. Sobel, M. J. Huyett, Selecting the best one of several binomial populations. *Bell Syst. Tech. J.* **36**, 537–576 (1957).
42. K. P. Murphy, *Machine Learning: A Probabilistic Perspective*, (MIT Press, 2012).
43. G. Andrews, *The Theory of Partitions*, (Cambridge University Press, 1998).
44. J. A. Deutsch, D. Deutsch, Some theoretical considerations. *Psychol. Rev.* **70**, 80–90 (1963).
45. A. M. Treisman, Strategies and models of selective attention. *Psychol. Rev.* **76**, 282–299 (1969).
46. S. Yantis, J. C. Johnston, On the locus of visual selection: Evidence from focused attention tasks. *J. Exp. Psychol. Hum. Percept. Perform.* **16**, 135–149 (1990).
47. R. Desimone, J. Duncan, Neural mechanisms of selective visual attention. *Annu. Rev. Neurosci.* **18**, 193–222 (1995).
48. T. F. Brady, T. Konkle, G. A. Alvarez, A review of visual memory capacity: Beyond individual items and toward structured representations. *J. Vis.* **11**, 4 (2011).
49. G. A. Miller, The magical number seven plus or minus two: Some limits on our capacity for processing information. *Psychol. Rev.* **63**, 81–97 (1956).
50. C. R. Sims, Rate-distortion theory and human perception. *Cognition* **152**, 181–198 (2016).
51. D. A. Norman, T. Shallice, "Attention to action" in *Consciousness and Self-Regulation*, R. J. Davidson, G. E. Schwartz, D. Shapiro, Eds. (Springer, 1986), pp. 1–18.
52. B. J. Sleezer, M. D. Castagno, B. Y. Hayden, Rule encoding in orbitofrontal cortex and striatum guides selection. *J. Neurosci.* **36**, 11223–11237 (2016).
53. P. Cisek, J. F. Kalaska, Neural mechanisms for interacting with a world full of action choices. *Annu. Rev. Neurosci.* **33**, 269–298 (2010).
54. B. Y. Hayden, R. Moreno-Bote, A neuronal theory of sequential economic choice. *Brain Neurosci. Adv.* **2**, 2398212818766675 (2018).
55. I. Krajbich, C. Armel, A. Rangel, Visual fixations and the computation and comparison of value in simple choice. *Nat. Neurosci.* **13**, 1292–1298 (2010).
56. S.-L. Lim, J. P. O'Doherty, A. Rangel, The decision value computations in the vmPFC and striatum use a relative value code that is guided by visual attention. *J. Neurosci.* **31**, 13214–13223 (2011).
57. A. D. Redish, Vicarious trial and error. *Nat. Rev. Neurosci.* **17**, 147–159 (2016).
58. E. L. Rich, J. D. Wallis, Decoding subjective decisions from orbitofrontal cortex. *Nat. Neurosci.* **19**, 973–980 (2016).
59. K. Diehl, C. Poynor, Great expectations?! Assortment size, expectations, and satisfaction. *J. Mark. Res.* **47**, 312–322 (2010).
60. S. S. Iyengar, M. R. Lepper, When choice is demotivating: Can one desire too much of a good thing? *J. Pers. Soc. Psychol.* **79**, 995–1006 (2000).
61. M. F. Rushworth, M. P. Noonan, E. D. Boorman, M. E. Walton, T. E. Behrens, Frontal cortex and reward-guided learning and decision-making. *Neuron* **70**, 1054–1069 (2011).
62. M. Siegel, T. J. Buschman, E. K. Miller, Cortical information flow during flexible sensorimotor decisions. *Science* **348**, 1352–1355 (2015).
63. T. J. Vickery, M. M. Chun, D. Lee, Ubiquity and specificity of reinforcement signals throughout the human brain. *Neuron* **72**, 166–177 (2011).
64. S. B. M. Yoo, B. Y. Hayden, Economic choice as an untangling of options into actions. *Neuron* **99**, 434–447 (2018).
65. J. R. Hauser, B. Wernerfelt, An evaluation cost model of consideration sets. *J. Consum. Res.* **16**, 393 (1990).
66. G. J. Stigler, The economics of information. *J. Polit. Econ.* **69**, 213–225 (1961).
67. B. Scheibehenne, R. Greifeneder, P. M. Todd, Can there ever be too many options? A meta-analytic review of choice overload. *J. Consum. Res.* **37**, 409–425 (2010).
68. N. Mehta, S. Rajiv, K. Srinivasan, Price uncertainty and consumer search: A structural model of consideration set formation. *Mark. Sci.* **22**, 58–84 (2003).
69. J. H. Roberts, J. M. Lattin, Development and testing of a model of consideration set composition. *J. Mark. Res.* **28**, 429–440 (1991).
70. B. D. los Santos, A. Hortaçsu, M. R. Wildenbeest, Testing models of consumer search using data on web browsing and purchasing behavior. *Am. Econ. Rev.* **102**, 2955–2980 (2012).
71. P. Sepulveda *et al*., Visual attention modulates the integration of goal-relevant evidence and not value. *Neuroscience*, 10.1101/2020.04.14.031971 (2020).
72. P. Morgan, R. Manning, Optimal search. *Econometrica* **53**, 923 (1985).
73. W. E. Hick, On the rate of gain of information. *Q. J. Exp. Psychol.* **4**, 11–26 (1952).
74. R. M. Roe, J. R. Busemeyer, J. T. Townsend, Multialternative decision field theory: A dynamic connectionist model of decision making. *Psychol. Rev.* **108**, 370–392 (2001).
75. S. Tajima, J. Drugowitsch, A. Pouget, Optimal policy for value-based decision-making. *Nat. Commun.* **7**, 12400 (2016).
76. M. Usher, J. L. McClelland, Loss aversion and inhibition in dynamical models of multialternative choice. *Psychol. Rev.* **111**, 757–769 (2004).
77. C. J. Bates, R. A. Lerch, C. R. Sims, R. A. Jacobs, Adaptive allocation of human visual working memory capacity during statistical and categorical learning. *J. Vis.* **19**, 11 (2019).
78. C. A. Sims, Implications of rational inattention. *J. Monet. Econ.* **50**, 665–690 (2003).
79. H. A. Simon, "Theories of bounded rationality" in *Decision and Organization*, C. B. McGuire, R. Radner (North-Holland Publishing Company, 1972), chap. 8.

Heuristics and optimal solutions to the breadth-depth dilemma

Rubén Moreno-Bote, Jorge Ramírez-Ruiz, Jan Drugowitsch and Benjamin Y. Hayden

Supplementary Information

1 Finite capacity model

We consider a two-stage decision process in a multi-alternative decision-making problem modeled as a partially observable Markov decision process. There are N alternatives, defined each by a Bernoulli random processes, whose trial by trial ($t = 1, ...$) outcomes follow $s_i^t \sim \text{Bern}(p_i)$, $s_i^t \in \{0, 1\} = \{\text{failure}, \text{success}\}$, $i = 1, ..., N$. The outcomes are independently distributed for all trials t and across alternatives. The values of the success probabilities are unknown to the decision-maker, and follow a prior distribution $p_i \sim \text{Beta}(\alpha, \beta)$ i.i.d. for all alternatives, with known hyperparameters (α, β). Allowed actions follow a two-stage decision process. In the first stage, the decision-maker can draw a total of $C = N$ samples at once, namely, a one-go decision is considered [1, 2, 3]. We consider the case where the total number of alternatives N exhausts sampling capacity C, but the results are equivalent if the number of alternatives is larger than capacity, with the addition of more rejected or non-sampled alternatives. The action space is $\mathsf{A_L} = \{\vec{L} : L_i \geq 0 \ \forall i, \sum_i L_i = C\}$, where $\vec{L} = (L_1, ..., L_N)$ is the number of samples drawn from each of the alternatives, with the constraint $\sum_i L_i = C$ (we often refer to the vector $\vec{L}$ as sample allocation). Note that the decision-maker can decide to sample the same alternative several times (i.e.,

www.pnas.org/cgi/doi/10.1073/pnas.2004929117

$L_i > 1$ for some i), and also decide not to sample from several alternatives (i.e., $L_i = 0$ for other i). In general, $M \leq C = N$ alternatives are sampled. If just a few alternatives are sampled ($M \sim 1$), many samples can be allocated to each. If C alternatives are sampled, only one sample could be allocated to each of them. Outcomes of the samples from the sampled alternatives are revealed all at once, not sequentially. In the second stage of the decision-making process, after outcomes are observed, the decision-maker should decide what alternative to choose. We initially assume that it is possible to choose only among the sampled alternatives. Thus, the action space in the second stage is defined by the set $\mathsf{A_C} = \{c : L_c > 0\}$ of size M, ordered as $\{c_1, ..., c_M\}$. The sufficient statistics of the outcomes of the Bernoulli processes to infer the success probabilities are the counts of successes for each of the M sampled alternatives, $\vec{n} = (n_{c_1}, ..., n_{c_M})$, with $n_j = \sum_{t=1}^{L_{c_j}} s_{c_j}^t$, and thus the decision of what option to choose should be a function of those counts and on the sample allocation vector $\vec{L}$, which together constitute the information state of the decision process. The counts, conditioned on the success probabilities, follow $n_i \sim \mathrm{Bin}(p_i, L_i)$. Note that the dimension of the vector $\vec{n}$ depends on the number of sampled alternatives (those satisfying $L_i > 0$) and thus the consideration set changes size depending on the first stage decision.

We define the utility of a choice $i \in \mathsf{A_C}$ as the hidden value of the success probability of the corresponding Bernoulli variable, $U_i = p_i$. We assume that the decision-maker maximizes expected utility. This involves determining the optimal allocation of samples $\vec{L}^*$ to be used in the first stage followed by defining an optimal decision rule that selects one of the sampled alternatives based on $\vec{n}$. A decision rules maps an observation $\vec{n}$, given the allocation vector $\vec{L}$, into an element of the action space $\mathsf{A_C}$. By considering all possible decision rules, $\delta = \{\delta(\vec{n}, \vec{L}) : (\vec{n}, \vec{L}) \to \mathsf{A_C}\}$, we show in Sec. (5) that the optimal decision rule, $\delta^*(\vec{n}, \vec{L})$, is the one that selects always, for any sample allocation $\vec{L}$, the alternative with the maximum posterior mean success probability $\mathbb{E}(p_i|n_i, L_i) = \frac{n_i+\alpha}{L_i+\alpha+\beta}$, $i \in A_C$, or chooses any of the maximum ones if there are ties. Therefore, the expected utility for a given sample allocation $\vec{L}$ following the optimal decision rule is

$$U(\vec{L}) = \sum_{\vec{n}} p\left(\vec{n}|\vec{L}, \alpha, \beta\right) \max_{i \in \mathsf{A_C}} \left(\frac{n_i + \alpha}{L_i + \alpha + \beta} \right), \tag{1}$$

where the joint posterior over $\vec{n}$ factorizes into beta-binomial distributions as $p\left(\vec{n}|\vec{L}, \alpha, \beta\right) = \prod_{i \in \mathsf{A_C}} \mathrm{Bb}(n_i|L_i, \alpha, \beta)$. Then, the optimal sample allocation $\vec{L}^*$ equals

$$\vec{L}^* = \arg\max_{\vec{L} \in \mathsf{A_L}} U(\vec{L}) = \max_{\vec{L} \in \mathsf{A_L}} \sum_{\vec{n}} p\left(\vec{n}|\vec{L}, \alpha, \beta\right) \max_{i \in \mathsf{A_C}} \left(\frac{n_i + \alpha}{L_i + \alpha + \beta} \right), \tag{2}$$

and the corresponding maximum expected utility becomes

$$U^* = \max_{\vec{L} \in \mathsf{A_L}} U(\vec{L}). \tag{3}$$

Finding the optimal solution in Eq. (2) is hard because of the large number of sample allocations that it is possible to form out of C samples. The number of unique partitions of C samples equals the number of integer partitions of C (not to be confused with the Bell number), for which we are not aware of simple exact expressions. We should only consider unique partitions because all the alternatives are initially (before sampling) indistinguishable. Therefore, without loss of generality, we can always assume that we sample the alternatives by using the sample allocation $\vec{L} \in \mathsf{A_L}$ where we impose the additional constraint that $L_i \geq L_{i+1}$ for $i = 1, ..., N-1$. That is, we sample the first alternative with more or the same number of samples as the second alternative, the second alternative with more or the same number of samples as the third one, and so forth. We describe a stochastic hill climbing algorithm bellow in Sec. (4) to find the optimal sample allocation exactly for small capacity C and approximately for large capacity. To find useful analytical expressions for Eqs. (2, 3), we restrict ourselves further by first looking for optimal even sample allocations, that is, allocation of samples across $M \leq C$ options with the same number of samples L per alternative. Optimal even sample allocation across alternatives is discussed in Sec. (2).

2 Analytical expressions for optimal even sample allocation

Because the space of actions $\mathsf{A_L} = \{\vec{L} : L_i \geq 0 \ \forall i, \sum_i L_i = C\}$ is very large, we restrict ourselves to a subset of possible actions, consisting in dividing the capacity C into M alternatives equally sampled with L samples each. Without loss of generality, we assume that we sample the first M alternatives and we ignore the rest of $C - M$ alternatives. Even splitting of the capacity is only possible if $C = M \times L$ holds exactly, so we will only examine the pairs (M, L) that satisfy that condition. The advantage of working in this subset of actions is that it is possible to obtain useful, exact analytical expression that will reveal non-trivial properties of the decision process. Methods for finding globally optimal sample allocation strategies are provided in Sec. (4). In the main results we also show that optimal sample allocations are not greatly better than the optimal even ones, so that even sample allocation is close-to-optimal. For an even capacity split, the optimal L^* under the constraint $C = ML$ can be obtained by specializing Eq. (2) to this case as

$$L^* = \arg\max_L \sum_{\vec{n}} \prod_{j=1}^{M} p\left(n_j|L,\alpha,\beta\right) \max_i \left(\frac{n_i+\alpha}{L+\alpha+\beta}\right), \tag{4}$$

where $i \in \{1,...,M\}$ and $p\left(n_j|L,\alpha,\beta\right) = \mathrm{Bb}(n_j|L,\alpha,\beta)$. Naturally, the optimal number of alternatives to be sampled is $M^* = C/L^*$

A particularly simple expression results from the case $\alpha = \beta = 1$, corresponding to a uniform prior over the success probabilities of the Bernoulli variables. This is because $p\left(n_j|L,1,1\right) = \mathrm{Bb}(n_j|L,1,1) = \frac{1}{L+1}$, thus becoming a discrete uniform distribution over $n_j \in \{0,...,L\}$, independent of n_j. Then, replacing this expression in Eq. (4), the optimal even sample allocation simplifies to

$$\begin{aligned}
L^* &= \arg\max_L U(L), \\
U(L) &= \frac{1}{(L+1)^M} \sum_{n_1,...,n_M=0}^{L} \max_i \left(\frac{n_i+1}{L+2}\right) && (5)\\
&= \frac{1}{(L+1)^M(L+2)} \left((L+1)^M + \sum_{n_1,...,n_M=0}^{L} \max(n_1,...,n_M)\right) \\
&= \frac{1}{(L+1)^M(L+2)} \left((L+1)^M + \sum_{s=0}^{L} \left((s+1)^M - s^M\right) s\right) && (6)\\
&= 1 - \frac{\sum_{s=0}^{L}(s+1)^M}{(L+1)^M(L+2)}, && (7)
\end{aligned}$$

with $M = C/L$. Eq. (6) in the derivation results from realizing that the sum over $\max_i(n_i)$ contains exactly $1^M - 0$ zeros, $2^M - 1$ ones, $3^M - 2^M$ twos, etc. The sum in Eq. (7) is the sum of the $M-th$ powers of the first $L+1$ integers, and it can be computed using Faulhaber's formula. Eq. (7) confirms the intuition that the expected utility $U(L)$ for any L is smaller than one. Finally, the optimal number of evenly allocated samples (over the sampled options) can be written as

$$L^* = \arg\min_L \frac{\sum_{s=0}^{L}(s+1)^M}{(L+1)^M(L+2)} \tag{8}$$

It is interesting to examine some limits of Eq. (7) by relaxing the constraint $C = M \times L$. For large M and $L = 1$, the expected utility in Eq. (7) becomes $\lim_{M\to\infty} U \to \frac{2}{3}$. This observation is not surprising, as when a very large number of alternatives is sampled with just one sample, it is very likely that at least one of them will have a successful outcome. Therefore, the expected utility of that alternative under the uniform prior will be $\frac{2}{3}$. This limit is visible in the rightmost point of Fig. 2a. In the opposite scenario, when only one alternative is sampled, $M = 1$, then the expected utility is $\frac{1}{2}$ for all L. That is, if just one alternative is sampled, then the expected probability of success of the sampled alternative is $\frac{1}{2}$, which equals the prior mean. This limit is visible in the leftmost point of Fig. 2a.

A more general way of performing the integrals involved in Eq. (4) is by using cumulative distribution function of the beta-binomial distributions, $F(n|L,\alpha,\beta) = \sum_{m\leq n} \mathrm{Bb}(m|L,\alpha,\beta)$. By noting that the extreme value distribution has probability mass function $F^M(n) - F^M(n-1)$ (where M denotes exponent and we have dropped conditioning to avoid cluttered notation), we can write the optimal even sample allocation in Eq. (4) as

$$L^* = \arg\max_L \sum_{n=0}^{L} \left[F^M(n|L,\alpha,\beta) - F^M(n-1|L,\alpha,\beta)\right]\left(\frac{n+\alpha}{L+\alpha+\beta}\right), \tag{9}$$

Note that the extreme value distribution $F^M(n_{max}) - F^M(n_{max}-1)$ is the distribution of $n_{max} = \max(n_1, ..., n_M)$ where $\vec{n}$ follows the above factorized beta-binomial distribution. In other words, the extreme value distribution for n_{max} is the probability that no alternative has more than n_{max} successful samples (hence the first term $F^M(n_{max})$) but removing the cases where there is no alternative with more than $n_{max}-1$ successful samples (hence the second negative term $F^M(n_{max}-1)$). For the uniform prior case, $\alpha = \beta = 1$, we recover Eq. (8), for which the cumulative can be exactly computed. For arbitrary values of α and β, Eq. (9) is solved numerically. These solutions are used in Fig. 2d.

The general Eq. (2) valid for any allocation of samples, and the specific Eq. (9) valid for even sample allocations, assume that a choice is made from the sampled alternatives, while non-sampled alternatives are excluded from the choice set. However, if none of the sampled alternatives turns to be good ones (e.g., because $n_i \ll L_i$ for $i \in \mathsf{A_C}$), then it would be better to choose randomly from any of the non-sampled alternatives. This is particularly so if the expected utility of any of the sampled alternatives, $\frac{n_i+\alpha}{L_i+\alpha+\beta}$, is smaller than $\frac{\alpha}{\alpha+\beta}$, which is the default expected utility of the non-sampled alternatives given that the success probabilities are drawn from a $\mathrm{B}(\alpha, \beta)$. It is straightforward to generalize these results by adding a default alternative, assumed to have utility p_0. In this case, the optimal even allocation of samples obeys

$$L^* = \arg\max_L \sum_{n=0}^{L} \left(F^M\left(n|L,\alpha,\beta\right) - F^M\left(n-1|L,\alpha,\beta\right)\right) \max\left(\frac{n+\alpha}{L+\alpha+\beta},\ p_0\right). \tag{10}$$

3 Asymptotic behavior for large capacity: the square root sampling law

It is possible to derive an approximation for the limiting behavior of the optimal number of sampled alternatives M^* and their associated optimal number of samples per alternative L^* by using Eq. (5) for large capacity C in the case of the uniform prior distribution. For large capacity C, we assume that L^* grows to infinity. This assumption is confirmed later, when the asymptotic optimal L^* is derived. If L is large, then Eq. (5) can be approximated by

$$\begin{aligned} U(L) &= \frac{1}{(L+1)^M} \sum_{n_1,...,n_M=0}^{L} \max_i \left(\frac{n_i+1}{L+2} \right) \\ &= \frac{1}{(L+2)} \left(1 + \frac{1}{(L+1)^M} \sum_{n_1,...,n_M=0}^{L} \max(n_1,...,n_M) \right) \\ &\approx \frac{1}{(L+2)} \left(1 + L \int_0^1 dx_1 ... \int_0^1 dx_M \max(x_1,...,x_M) \right), \end{aligned} \tag{11}$$

where the sum in the second equation has been approximated in the third equation by an integral in the interval $[0,1]^M$ over a uniform distribution by using the transformation $n_i = Lx_i$ for $i = 1,...,M$. The continuous approximation is valid when L is large, as assumed, since then the transformation delivers values of x_i that are dense in the unit interval. The integral can be rewritten as

$$\int_0^1 dx_1 ... \int_0^1 dx_M \max(x_1,...,x_M) = \int_0^1 dx_{max} x_{max} f(x_{max}),$$

where we have defined the extreme value $x_{max} = \max(x_1,...,x_M)$. The extreme value follows the extreme value distribution $f(x_{max}) = (F(x_{max})^M)' =$

Mx_{max}^{M-1}, where we have used that $F(x) = x$ is the cumulative of the continuous uniform distribution in $[0, 1]$. Therefore,

$$
\begin{aligned}
U(L) &\approx \frac{1}{(L+2)} \left(1 + L \int_0^1 dx_{max} M x_{max}^M\right) \\
&= \frac{1}{(L+2)} \left(1 + \frac{ML}{M+1}\right).
\end{aligned} \tag{12}
$$

Finally, by maximizing $U(L)$ as a function of L with the constraint $C = ML$ we obtain the asymptotic optimal number of sampled alternatives M^* and optimal number of samples per sampled alternative L^*

$$
\lim_{C\to\infty} M^* = \sqrt{C}, \quad \lim_{C\to\infty} L^* = \sqrt{C},
$$

which corresponds to the square root sampling law.

In the above derivation we have assumed that L^* grows with C. To see that this corresponds to the only valid assumption to obtain L^*, let us assume now that L^* does not grow with C, that is, it is a constant or decreases with C. For any fixed value L, using Eq. (7) we see that $U(L) \leq 1 - 1/(L+2)$. This utility is smaller than the one obtained by using the square root law, which converges to 1, as can be easily derived from Eq. (12). Therefore, the square root law delivers the highest utility.

4 Optimal sample allocation

For low capacity $C \leq 7$ we found the globally optimal sample allocation strategy by exhaustive search over all possible sample allocations. For larger capacity, we searched the optimal sample allocation by using stochastic hill climbing. With this method, we confirmed that for values up to $C \leq 20$ the globally optimal sample allocations were correct up to a precision in expected utility of 10^{-4}.

We started the algorithm by using even sample allocation using the square root law heuristic: if $C \leq 7$ all options were sampled with one sample, and if capacity was larger we used the square root law by sampling $\sqrt{C}$ alternatives $\sqrt{C}$ times each. We considered the possibility that the resulting square root was not an integer, and thus we allocated the residual number of samples to a randomly chosen additional alternative; we call this allocation scheme 'even allocation'. At every iteration, we computed the expected utility of the current best sample allocation $\vec{L}$ through a Monte Carlo simulation of the Bernoulli variables and averaging utility over 4×10^5 repetitions for $C \leq 20$ and 5×10^4 for larger capacity values. A perturbed sample allocation was proposed by randomly selecting two alternatives. One sample was removed from the first alternative and added to the second one, but only if the first alternative had already assigned at least one sample. To exploit symmetry, we only consider changes of one sample from one alternative i to another $j > i$ if $L_{j-1} \geq L_j$ and $L_i \geq L_j$. If $j < i$, there were not restrictions.

With the proposed perturbed sample allocation, we computed the expected utility using the same Monte Carlo method. If the new expected value was larger than the previous one, then the proposed perturbed sample allocation became the current best sample allocation. This process was iterated 2×10^4 times for $C \leq 20$ and 3×10^3 for larger capacity values. Because at each iteration we reevaluate the expected value of the current best sample allocation, we avoid the possibility of getting stuck in a random fluctuation leading to a spuriously large expected value. The stochastic hill climbing method found optimal sample allocations that were identical to those found with the exhaustive search for low capacity $C \leq 7$. Although we do not know whether the found optimal sample allocation corresponds to a global maximum when capacity is larger, we confirmed that the optimal sample allocations found were stable against different random number seeds and initial conditions. Figs. 4 and 5 use the above method. Percentage reward gain in Fig. 5b is computed as $100 \times (U^* - U_{even})/U_{even}$, where U^* is the utility estimate of the globally optimal allocation and U_{even} is the estimate of the initial even allocation. Percentage reward loss in Fig. 5c is computed as $100 \times (U_{heuristic} - U^*)/U^*$, where U^* is the utility estimate of the globally optimal allocation and $U_{heuristic}$ is the utility estimate from triangular, square root sampling law, pure breadth or pure depth heuristics.

We also employed another version of stochastic hill climbing that avoided using extensive sampling of the Bernoulli variables to estimate expected utility. This method was used to confirm robustness of the previous results. We define the optimal utility as

$$U^* = \max_{\vec{L}} \sum_{\vec{n}} p(\vec{n}|\vec{L}, \alpha, \beta) \max_{i} \left(\frac{n_i + \alpha}{L_i + \alpha + \beta} \right). \tag{13}$$

We thus can design a Markov Chain Monte Carlo method to sample from the probability distribution

$$p(\vec{n}|\vec{L}, \alpha, \beta) = \prod_{j} Bb(n_j|\vec{L}, \alpha, \beta)$$

appearing in the sum of Eq. (13) as follows (these samples can be then used to approximate the sum). Detailed balance imposes that the probability of transitioning from a state with $\vec{n}$ to $\vec{n}'$ is the same as the converse,

$$P_{\vec{n},\vec{n}'}\, p(\vec{n}|\vec{L}, \alpha, \beta) = P_{\vec{n}',\vec{n}}\, p(\vec{n}'|\vec{L}, \alpha, \beta).$$

By proposing a change to a single alternative $n'_j = n_j \pm 1$, we can get a simple expression for the acceptance rate $r(\vec{n} \to \vec{n}')$. If $n'_j = n_j + 1$ the acceptance rate is

$$r(\vec{n} \to \vec{n}') = \min\left(1, \frac{(n_j + \alpha)(L_j - n_j)}{(L_j - n_j + \beta + 1)(n_j + 1)}\right),$$

while if $n'_j = n_j - 1$, it becomes

$$r(\vec{n} \to \vec{n}') = \min\left(1, \frac{(L_j - n_j + \beta)n_j}{(n_j + \alpha - 1)(L_j - n_j)}\right),$$

where we have made use of the Metropolis-Hastings algorithm. These two changes are proposed with equal probability and randomly across all the options. Utilities are estimated using 10^6 samples. The search over $\vec{L}$ is made using $50 \times C$ iterations. Results in Fig. 4 were reproduced by this method.

For the optimal dynamic allocations described in Fig. 6, we employed again a stochastic hill climbing method identical to the one described at the start of this section by using the vector of numbers of allocated samples per wave, M_i, $i = 1, 2, ...C$, instead of the number of samples per alternative, L_i. The method proceeded by proposing a new vector of waves $\vec{M}$ by adding a sample to a randomly chosen wave and removing a sample from another randomly chosen wave. This was done only if the second wave had at least one allocated sample to it and if the resulting proposed perturbed allocation satisfied the constraint $M_{i+1} \leq M_i$ for all i. The number of iterations and samples for Monte Carlo utility estimates are the same as above. Optimal dynamic allocations found are correct up to a precision of 10^{-4} in the utility estimates. Very similar results to those described in Fig. 6 are found when options to be sampled in each wave are selected based on their current posterior mean probabilities instead of their current number of total successes. Percentage reward loss of static heuristics compared to optimal dynamic allocations described in Fig. 6d are computed as in Fig. 5c.

5 Consistency

Perhaps intuitively, but wrongly, we might assume that by always opting for the alternative with larger number of successful outcomes (larger n_i in Eq. (2)), this would result in 'cherry picking', that is, in selecting a spuriously good option. This, in turn, would mean that we would obtain a reward that is lower than the expected utility in Eq. (3). Here we show, however, that the decision rule of choosing always the alternative with the highest posterior mean is both optimal and delivers on average a reward that is equal to the expected utility. This is a well-known result in statistical decision theory [4, 5, 6]. Here we show the derivation for completeness.

Consider any possible decision rule $\vec{d} = \delta(\vec{n})$ that assigns the counts of successes for the M sampled alternatives, $\vec{n}$, to a decision $\vec{d} \equiv \vec{d}(\vec{n}) = (d_{c_1}(\vec{n}), ..., d_{c_M}(\vec{n}))$, encoded as a one-hot vector of length M (i.e., $d_{c_i} = 1$ if alternative c_i is chosen, and $d_{c_i} = 0$ otherwise; we omit the potential dependence of the decision rule on $\vec{L}$ to avoid cluttered notation). If the success probabilities of the sampled alternatives, $\vec{p}$, are known, then by using the decision rule δ the decision-maker would have an expected utility

$$U(\vec{p}, \vec{L}, \delta) = \sum_{\vec{n}} \prod_{i \in \mathsf{A_C}} \mathrm{Bin}(n_i | L_i, p_i) \; p_i^{d_i(\vec{n})},$$

where $\vec{L}$ is the allocated number of samples over the alternatives. Note that the expected utility is an average over the values of the chosen p_i given the decision rule averaged across all possible outcomes given the allocated number of samples over alternatives. As probabilities are unknown, they are marginalized out with their prior beta distributions, resulting in the overall expected utility

$$U(\vec{L},\delta) = \sum_{\vec{n}} \prod_{i\in \mathsf{A_C}} \frac{\Gamma(L_i+1)}{\Gamma(n_i+1)\Gamma(L_i-n_i+1)} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \times \frac{\Gamma(n_i+\alpha+d_i)\Gamma(L_i-n_i+\beta)}{\Gamma(L_i+\alpha+\beta+d_i)}. \tag{14}$$

We note that for each term in the sum over $\vec{n}$, there is only one value of i for which $d_i = 1$ in the product, while $d_j = 0$ for $j \neq i$. The term i in the product with $d_i = 1$ gives an extra factor $\frac{n_i+\alpha}{L_i+\alpha+\beta}$ (by expanding the gamma functions just one step) that is not present in the product terms with $d_j = 0$. Therefore, the product is maximized iff $d_i = 1$ for the alternative i with maximum $\frac{n_i+\alpha}{L_i+\alpha+\beta}$ (if the maximum is not unique, any alternative with the maximum value will give exactly the same result). This result proves that the optimal decision rule δ^* is the one that chooses always the alternative with the highest posterior expected utility given $\vec{n}$.

Now, we can show that for the optimal decision rule δ^*, the expected utility is the same as that in Eq. (3). We can rewrite Eq. (14) as

$$U(\vec{L},\delta^*) = \sum_{\vec{n}} \max_{i\in \mathsf{A_C}} \left(\frac{n_i+\alpha}{L+\alpha+\beta}\right) \prod_{i\in \mathsf{A_C}} \frac{\Gamma(L_i+1)}{\Gamma(n_i+1)\Gamma(L_i-n_i+1)} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \times \frac{\Gamma(n_i+\alpha)\Gamma(L_i-n_i+\beta)}{\Gamma(L_i+\alpha+\beta)},$$

which is identical to the maximum expected utility $U(\vec{L})$ in Eq. (3), that is, $U(\vec{L},\delta^*) = U(\vec{L})$. This shows that 'cherry picking' is optimal.

6 Supplementary Figure

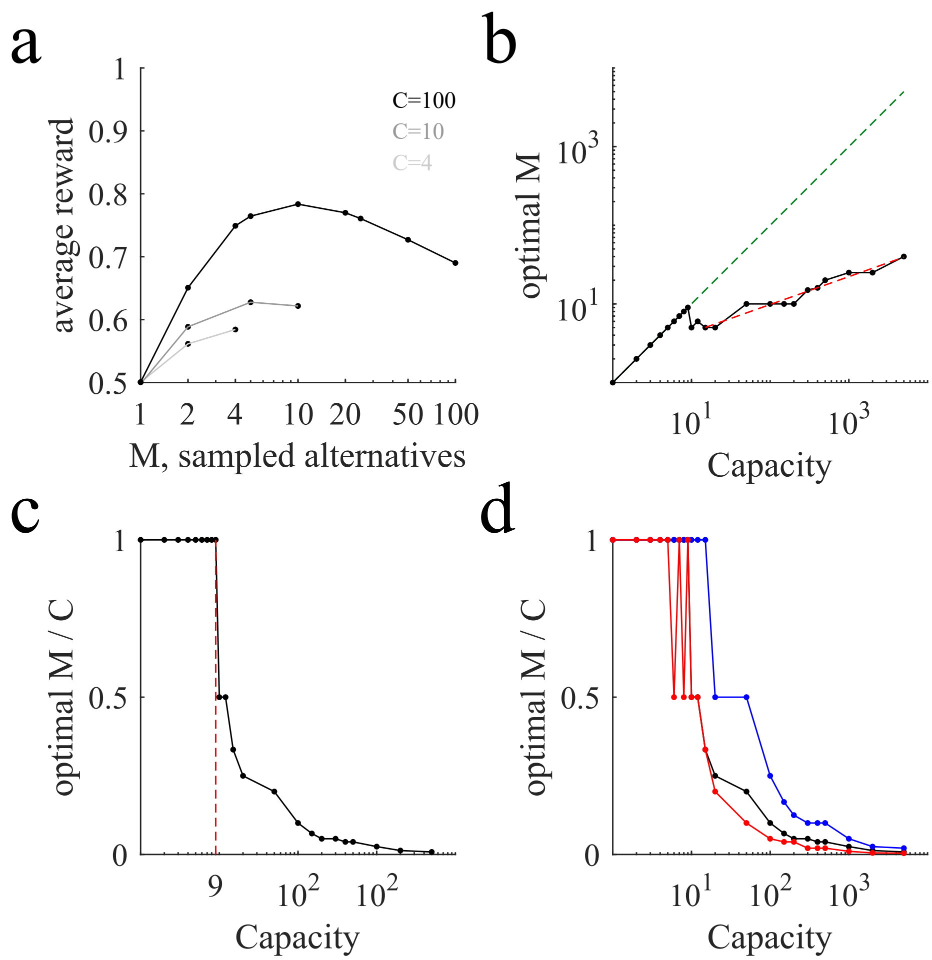

Figure SI 1: Sharp transitions in optimal number of sampled alternatives at low capacity and power law behavior at high capacity in a breadth-depth (BD) model with Gaussian outcomes. Each option is modelled as a Gaussian with known variance σ^2 and unknown mean reward μ drawn independently for each option from a uniform prior over $[0, 1]$. The goal is to optimize the even allocation of C independent samples over at most C options to maximize the posterior mean reward of the best option. The less options samples are allocated to, the better the estimates of the underlying means of the sampled options. The optimal even allocations observed qualitatively match those found in Figure

1 for Bernoulli observations with unknown success probabilities. (a) Average reward (points and lines, simulations) as a function of the number of sampled alternatives M for three different capacities ($C = 4, 10, 100$; light, intermediate and dark lines respectively) for variance $\sigma^2 = 1$, comparable to prior's width. The maximum occurs at the right extreme for low capacity but at a relatively low values for large capacities. Note log horizontal scale. (b) Optimal number of sampled alternatives as a function of capacity. When capacity is smaller than around 9, a linear trend of unit slope is observed (dashed green line), but when capacity is above 9, the behavior becomes sublinear (dashed red line corresponds to the best power law fit, with exponent close to 1/3; power law fit, exponent = 0.35, 95% CI = $[0.30, 0.41]$). The transition between these two regimes is sharp. (c) The sharp transition is clearer when plotting the optimal number of sampled alternatives to capacity ratio as a function of capacity. For low capacity, the ratio is one, but for large capacity the ratio decreases very rapidly. The last point below which the optimal ratio is always one (critical capacity) corresponds to capacity equal to 9 (indicated by the vertical red line). (d) Number of sampled alternatives to capacity ratios for different variances $\sigma^2 = 0.1, 1, 10$ (blue, black, red lines, respectively), corresponding to reliable, standard, and unreliable Gaussian samples. All points and lines correspond to simulations. When samples are reliable (blue line), breadth search is favored, as can be seen from the increase of the critical capacity and the slower decay of the optimal ratio M/C. In contrast, when samples are unreliable (red line), depth search is favored.

References

[1] Bechhofer, R. E., Kulkarni, R. V. Closed sequential procedures for selecting the multinomial events which have the largest probabilities. Communications

in Statistics - Theory and Methods 13, 2997–3031 (1984).

[2] Gupta, S. S., Liang, T. Selecting the best binomial population: parametric empirical Bayes approach. Journal of Statistical Planning and Inference 23, 21–31 (1989).

[3] Sobel, M., Huyett, M. J. Selecting the Best One of Several Binomial Populations. Bell System Technical Journal 36, 537–576 (1957).

[4] Moreno-Bote, R. Decision Confidence and Uncertainty in Diffusion Models with Partially Correlated Neuronal Integrators. Neural Computation 22, 1786–1811 (2010).

[5] Beck, J. M. et al. Probabilistic population codes for Bayesian decision making. Neuron 60, 1142–52 (2008).

[6] Berger, J. O. Statistical Decision Theory and Bayesian Analysis. (Springer New York, 1985).