Efficient sampling and noisy decisions

Authors

Joseph A. Heng[1], Michael Woodford[2] & Rafael Polanía[1*]

[1]Decision Neuroscience Lab
Department of Health Sciences and Technology, ETH Zurich, Switzerland

[2]Department of Economics, Columbia University, New York, USA

* Corresponding author: Rafael Polanía, rafael.polania@hest.ethz.ch

Abstract

Noise and information sampling are ubiquitous components determining the precision in our ability to discriminate between decision alternatives, however, the source and nature of these imprecisions is unclear. Moreover, how the nervous system simultaneously considers regularities in the environment, goals of the organism, and capacity constraints to guide goal-directed behavior, remains unknown. To address these issues, we elaborate a biologically- and cognitive-relevant efficient coding mechanism for discrimination, which assumes a limited number of units available to process incoming information, where in turn, each of these units can generate a set of few discrete samples to encode input values. We formally demonstrate that in order to be efficient, noise in the operation of this system is a necessary requirement in order to fulfill the organism's goal, for instance, maximize the amount of correct responses or maximize rewards tied to input stimuli. When applying this theory to numerosity discrimination, we found that humans employ a more flexible but less optimal sampling mechanism that relies on limited samples drawn from memory of encountered stimuli, irrespective of incentivized goals. These findings provide a general mechanistic framework for understanding decision behavior while accounting for biological restrictions of information coding.

Main text

It has been suggested that the rules guiding behavior are not arbitrary, but should follow fundamental principles of acquiring information from environmental regularities in order to make the best decisions. Moreover, these principles should incorporate strategies of information coding that take into consideration the cost of making decisions as well as biological constraints of information acquisition. Therefore, the elaboration and understanding of such unifying theories could help to establish frameworks that allow comparing behavior not only across different tasks and goals, but also across different levels of description, for instance, from single cell activity to observed behavior.

Here, we adopt such an approach to understand what determines the degree of precision of people's discrimination between alternative stimuli presented to them, and in particular how this varies over the range of possible stimuli. In a number of sensory modalities, it has been argued that principles of efficient coding provide an answer: that the degree to which a particular stimulus can be distinguished from relatively similar stimuli follows from a principle of efficient allocation of scarce neural resources[1–6]. However, the cases in which this principle has been shown to successfully explain measured discrimination thresholds are typically ones in which the stimulus feature identifies "which one" of a large range of possibilities is present (for instance, which direction of motion or orientation, which frequency of pitch), rather than "how much" of something is present, for instance, in the case when physical stimuli are tied to reward values. Population codes in which many different neurons have preferred stimuli at different points in the spectrum of possibilities are known to be used in early stages of processing of sensory features of the former kind[1,3], but are not obviously relevant in the latter case. Yet, it is also important to understand the precision of discrimination of stimulus features of the latter kind; in particular, the values that must be represented in the case of reward-based decision making[2,4,7], which could in turn be beneficial to develop strategies for reward or fitness maximization. Can a unifying theory of efficient coding be developed that reflects constraints on feasible internal representations that is relevant to both cases?

Another hallmark in neural systems is the well-known observation that neural activity and behavior are considerably variable trial-to-trial, even when input stimuli is maintained as constant as possible[8–10]. However, it remains unclear what actually causes behavioral variability, which is usually blamed on distinct sources of noise that apparently corrupt ideal behavior[11,12]. Based on the fact that organisms have limited set of available resources to process information, noise has been usually attributed to these information processing constraints. However, the role of this assumed

unavoidable noise[13] and the extent to which it positively or negatively contributes to goal-directed behavior remains unclear.

Another issue that is not addressed by the existing literature on efficient coding is the time scale over which the adaptation of the coding scheme to the frequency distribution of stimuli encountered in one's environment occurs, and the mechanism through which such adaptation occurs. Typical treatments derive an optimal coding scheme for an environmental frequency distribution[14,15] (e.g., extracted from some database of naturally occurring stimuli) that is assumed to be invariant over time, and compares the predictions to measured discrimination thresholds that are also assumed to be invariant over time. But the frequency distribution of stimulus features that will be encountered varies as a person moves from one context to another; and this is especially likely to be important in the case of higher-level cognitive constructs such reward-based decisions[2,16]. Thus if one seeks to apply efficient coding principles to imprecision in decision making, it is necessary to determine how quickly the way in which particular magnitudes are encoded can adapt to new contexts with changes in the distribution of magnitudes that one encounters.

Here we seek to address these issues, using the case of discrimination of numerosity as an example of the type of perceptual discrimination that is likely to provide a better model for the kind of internal representation that is not only used to understand perception, but is also used in the case of more complex cognitive constructs such as reward-based decisions.

Theoretical results: A theory of efficient coding requires a particular assumption about the kind of internal representations that are feasible. Here we elaborate an efficient coding theory for the case when a system receives as input a continuous random variable, but can only encode information via ones and zeros. The relevance of developing and understanding such a theory is that the brain appears to encode information using such coding scheme (e.g., via the generation or absence of action potentials[17] or representation of information via binary states[18]) in order to guide decision behavior.

Suppose that the subjective perception of an environmental variable with value v is determined by n independent samples of a binary random variable, i.e. outcomes are either "one" or "zero" readings, where we define the noisy percept k as the number of times (out of a limited available set of n samples) that a "one" reading was generated. The coding rule can then be represented by a function $\theta(v)$, indicating the probability of a "one" reading when the input stimulus magnitude is v which we assume to be drawn from a prior distribution $f(v)$ (Fig. 1a). Below we will show that

this abstract formulation corresponds to an algorithm that could be implemented by a minimalistic neural architecture (Fig. 2), but in the first part of this work we will focus on developing the algorithmic formulation.

The nature of the resource-constrained encoder to be optimized depends on the performance measure that we use to determine the usefulness of a given encoding scheme. A common assumption in the literature on efficient coding has been that the encoding scheme maximizes the mutual information between the true stimulus magnitude and its internal representation[1–3]. We start by characterizing the optimal $\theta(v)$ for a given prior distribution $f(v)$, according to this criterion. It can be shown that for large n, the mutual information between θ and k (hence the mutual information between v and k) is maximized if the prior distribution over θ is Jeffreys' prior[19]

$$\mathrm{Beta}(\theta; 0.5, 0.5) = \frac{1}{\pi\sqrt{\theta(1-\theta)}}, \qquad (1)$$

also known as the arcsine distribution. Hence, the mapping $\theta(v)$ induces a prior distribution over θ given by the arcsine distribution (Fig. 1a, right panel). Based on this result, it can be shown that the optimal encoding rule $\theta(v)$ that guarantees maximization of mutual information between the random variable v and the noisy encoded percept k is given by (see Supplementary Notes 1 and 2)

$$\theta(v) = \left[\sin\left(\frac{\pi}{2}F(v)\right)\right]^2, \qquad (2)$$

where $F(\theta)$ is the CDF of the prior distribution $f(v)$.

So far, we have derived the optimal encoding rule to maximize mutual information when the transmission channel receives a continuous random variable as input, but can only encode information with ones and zeros codes (Eq. 2). However, one may ask what the implications are of such a theory for discrimination performance. This is important to investigate given that achieving channel capacity does not necessarily imply that the goals of the organism are also optimized[20]. Independent of information maximization assumptions, here we start from scratch and investigate what are the necessary conditions for minimizing discrimination errors when a system is restricted to represent information based on ones and zeros with limited number of samples. Here we solve this problem for the case of two alternative forced choice tasks, where the average probability of error is given by (see Supplementary Note 3)

$$\mathrm{E[error]} = \iint \mathrm{P_{error}}[\theta(v_1), \theta(v_2)] f(\theta_1) f(\theta_2) d\theta_1 d\theta_2, \qquad (3)$$

where $P_{error}[]$ represents the probability of erroneously choosing the alternative with the lowest value v given a noisy percept k (assuming that the goal of the organism in any given trial is to choose the alternative with the highest value). Here, we want to find the density function $f(\theta)$ that guarantees the smallest average error (Eq. 3).

Interestingly, the solution to this problem is (see Supplementary Note 3)

$$f(\theta) = \frac{1}{\pi\sqrt{\theta(1-\theta)}}, \qquad (4)$$

which is exactly the same prior density function over θ that maximizes mutual information (Eq. 1 and Fig. 1a). Crucially, please note that we have obtained this expression based on minimizing the frequency of erroneous choices and not the maximization of mutual information as a goal in itself. This provides a further (and normative) justification for why maximizing mutual information under this coding scheme is beneficial when the goal of the agent is to minimize discriminations errors (i.e., maximize accuracy).

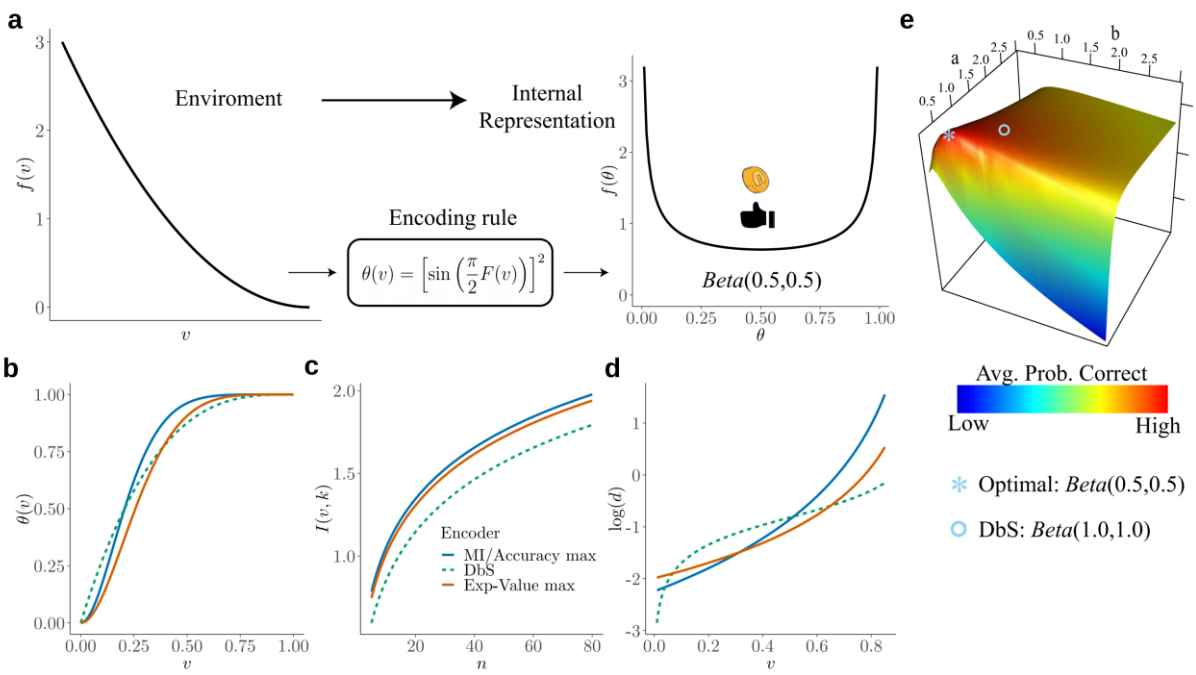


Figure 1. a) Schematic representation of our theory. Left: example prior distribution $f(v)$ of values v encountered in the environment. Right: Prior distribution in the encoder space (Eq. 1) due to optimal encoding (Eq. 2). This optimal mapping determines the probability θ of generating a "high" or "low" reading. The ex-ante distribution over θ that guarantees maximization of mutual information is given by the arcsine distribution (Eq. 1). **b)** Encoding rules $\theta(v)$ for different decision strategies under binary sampling coding: accuracy maximization (blue), reward maximization (red), DbS (green dashed). **c)** Mutual information

$I(v,k)$ for the different encoding rules (see Supplementary Note 2) as a function of the number of samples n. As expected $I(v,k)$ increases with n, however the rule that results in the highest loss of information is DbS. **d)** Discriminability thresholds d (log-scaled for better visualization) for the different encoding rules as a function of the input values v for the prior $f(v)$ given in panel a. **e)** Graphical representation of the perceptual accuracy optimization landscape. We plot the average probability of correct responses (Eq. 4) using as benchmark a *Beta* distribution with parameters a and b. The blue star shows the average error probability assuming that $f(\theta)$ is the arcsine distribution (Eq. 1), which is the optimal solution. The blue open circle shows the average error probability based on the encoding rule assumed in DbS, which is located near the optimal solution, but still a suboptimal strategy under the binomial coding scheme. Please note that when formally solving this optimization problem (Supplementary Note 3), we did not assume a priori that the solution is related to the beta distribution. We use the beta distribution in this figure just as a benchmark for visualization.

The results presented above imply that this encoding framework imposes limitations on the ability of capacity-limited systems to discriminate between different values of the encoded variables. Moreover, we have shown that error minimization in discrimination tasks imply a particular shape of the prior distribution of the encoder (Eq. 4) that is exactly the prior density that maximizes mutual information between the input v and the encoded noisy readings k (Eq. 1, Fig. 1a right panel). Does this imply a relation between prior and discriminability over the space of the encoded variable? Intuitively, following the efficient coding hypothesis, the relation should be that lower discrimination thresholds should occur for ranges of stimuli that occur more frequently in the environment or context.

Recently, it was shown that using an efficiency principle for encoding sensory variables (e.g., with a heterogeneous population of noisy neurons[21]) it is possible to obtain an explicit relationship between the statistical properties of the environment and perceptual discriminability[21,22]. The theoretical relation states that discriminability thresholds d should be inversely proportional to the density of the prior distribution $f(v)$. Here, we investigated whether this particular relation also emerges in the efficient coding scheme that we propose in this study.

Remarkably, we obtain the following relation between discriminability thresholds, prior distribution of input variables, and the number of limited samples n (derivation in Supplementary Fig. 1):

$$d = \frac{1}{\sqrt{n}\pi f(v)} \propto \frac{1}{f(v)} \quad . \qquad (5)$$

Interestingly, this relationship between prior distribution and discriminability thresholds holds empirically across several sensory modalities (Supplementary Fig. 1), thus once again demonstrating that the efficient coding framework that we propose here seems to incorporate the right kind of constraints to explain observed perceptual phenomena as consequences of optimal allocation of finite capacity for internal representation.

Until now, we have elaborated an efficient noisy sampling theory for the case when the goal of the organism is to minimize the number of mistakes in discrimination tasks. However, it is important to consider the case when the goal of the organism is to maximize fitness or expected reward[23]. For example, when spending the day foraging fruit, one must make successive decisions about which tree has more fruits. Fitness depends on the number of fruit collected which is not a linear function of the number of accurate decisions, as each choice yields a different amount of fruit.

Therefore, in the case of reward maximization, we are interested in minimizing reward loss which is given by the following expression

$$\mathrm{E}[\mathrm{loss}] = \iint |v_1 - v_2| \cdot \mathrm{P}_{\mathrm{error}}[\theta(v_1), \theta(v_2)] f(\theta_1) f(\theta_2) d\theta_1 d\theta_2 . \qquad (6)$$

Note that this expression is similar for the case of average error rate (Eq. 3), with the difference that now we want to take into consideration how much reward is lost every time an erroneous decision is made (when making a decision between input values v_1 and v_2, the decision maker gives up $|v_1 - v_2|$ for every erroneous decision). Thus, the goal is to find the encoding rule $\theta(v)$ which guarantees that the amount of reward loss is as small as possible given our proposed coding framework.

Here we show that the optimal encoding rule $\theta(v)$ that guarantees maximization of expected value is given by

$$\theta(v) = \sin\left[\frac{\pi}{2} \cdot c \int_{-\infty}^{v} f(\tilde{v})^{2/3} d\tilde{v}\right]^2 , \qquad (7)$$

where c is a normalizing constant which guarantees that the expression within the integral is a probability density function (Supplementary Note 5). The first observation based on this result is that the encoding rule for maximizing fitness is different from the encoding rule that maximizes accuracy (compare Eqs. 2 and 7), which leads to a slight loss of information transmission (Fig. 1c). Additionally, one can also obtain discriminability threshold predictions for this new encoding rule. Assuming a right-skewed prior distribution, which is often the case for various natural priors in the environment[22] (e.g., like the one shown in Fig. 1a), we find that discriminability for

small input values is lower for reward maximization compared to perceptual maximization, however this pattern inverts for higher values (Fig. 1d). In other words, when we intend to maximize reward (given the shape of our assumed prior, Fig. 1a), the agent should allocate more resources to higher values (compared to the perceptual case), however without completely giving up sensitivity for lower values, as these values are still encountered more often.

So far, we have proposed a *computational* problem and elaborated an *algorithmic* solution to our resource-constrained framework. However, it would be desirable to propose an *implementational* representation that also solves our resource-constrained problem[24,25]. Here we show that an advantage of our proposed coding scheme is that it allows a minimalistic and parsimonious implementational architecture. The number of samples n available to code any input value v can be seen as an array of n threshold non-linearities (e.g., neurons), each of which is subject to independently sampled random additive noise drawn from a noise distribution f_η (Fig. 2). Here, we assume that each neuron produces a "high" reading if and only if the signal v plus noise (drawn from f_η) is greater than a firing threshold τ. Using this simplified implementational scheme, it can be shown that channel capacity can be optimized by finding the optimal noise distribution f_η (see Supplementary Note 4)[26]

$$f_\eta(v) = \frac{\pi}{2}\sin[\pi(1 - F(\tau - v))]f(\tau - v). \quad (8)$$

Remarkably, this result is independent of the number of samples (resources) available to encode the input variable, and generalizes to any prior distribution f (recall that F is defined as its cumulative density function).

This result reveals three important aspects of neural function and decision behavior: First, it makes explicit why a system that evolved to code information using our proposed coding scheme (i.e., ones and zeros codes or binary-state states[18,27]) must be necessarily noisy. That is, we do not attribute the randomness of peoples' responses to a particular set of stimuli or decision problem to unavoidable randomness of the hardware used to process the information. Instead, the relevant constraints are assumed to be the limited set of output states for each neuron, the limited number of neurons, and the requirement that the neurons operate in parallel (so that each one's output state must be statistically independent of the others, conditional on the input stimulus). Given these constraints, we show that it is efficient for the operation of the neurons to be random. Second, it shows how the nervous system may take advantage of these noisy properties by reshaping its noise structure to optimize decision behavior.

Third, it shows that the noise structure can remain unchanged irrespective of the amount of resources available to guide behavior.

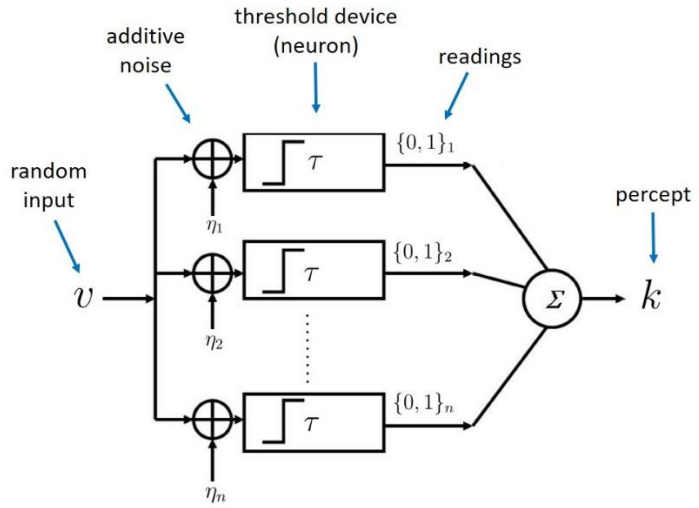


Figure 2. Schematic of a minimalistic implementational architecture. The system consists of an array of n threshold non-linearities (neurons). For each sample, neurons receive noisy versions of v where the noisy signals are i.i.d. additive random signals independent of v. The output of the neuron for each sample is "high" reading (one) if $(v + \eta_i) > \tau$ and zero otherwise. The noisy percept of the input is simply the sum of the outputs of each sample given by k.

In the elaboration of the optimal discrimination encoding rules described above, we suppose that there has been full learning of (and optimal adjustment to) the distribution of values encountered in a given environment or context over a long periods of time. However, for cases of more rapidly changing environments, a question arises as to whether the person's brain can really have enough information soon enough about the current frequency distribution for fully optimal adaptation of the encoding rule at each point in time to be possible. This leads us to consider an architecture that is more flexible in the way that it allows the encoding rule to vary quickly with changes in the distribution of magnitudes that have recently been encountered. This, at the cost of being less optimal in the way that the encoding rule would be adapted to a single frequency distribution of stimuli that was maintained forever.

In this alternative architecture, instead of assuming a distribution from which noise η is drawn that has been programmed at some earlier date, and optimized for some training data set that was available at that earlier time (Fig. 2), we allow the value of η to be drawn from a different distribution at each date, that depends on the

distribution of magnitudes that have been encountered before that date. We furthermore assume that the algorithm economizes on the number of samples that must be drawn from previous experience, by drawing a single exemplar from the distribution of recently encountered past magnitudes on each such occasion, and basing the value of η for that occasion on the magnitude of the exemplar that is drawn. We assume that the function that determines the value of η as a function of the exemplar that is drawn on any given occasion is fixed for all time. According to this model, the effective encoding rule $\theta(v)$ varies depending on the frequency distribution from which v is drawn in a given context, but only because the distribution of values for η will depend on the distribution of exemplars that will be drawn from memory in different contexts.

In this case, the encoding rule (the probability that a "high" reading is generated) approximates:

$$\theta(v) = F(v). \qquad (9)$$

The influential decision by sampling (DbS) model[28] is an example of a sampling-based encoding rule of this kind. We consider its performance in a given environment, from the standpoint of each of the possible performance criteria considered above (i.e., maximize correct responses or fitness), and show that it differs from the optimal encoding rule under any of those criteria (Fig. 1b-e). In particular, here we show that using the encoding rule employed in DbS results in considerable loss of information compared to the optimal solutions (Fig. 1c). An additional interesting observation is that for the strategy employed in DbS, the agent is more sensitive for extreme input values but gives up precision for intermediate values (Fig 1d). In other words, DbS agents are more sensitive to saliency. Thus, our theory illustrates the tradeoffs that emerge between accuracy and reward maximization – and how they compare to DbS – when considering the statistics of the environment, as well as capacity and neural coding constraints. Our goal now is to compare back-to-back these resource-limited coding frameworks in a fundamental cognitive function for human behavior: numerosity perception.

Empirical Results: Numerosity perception appears to be a well-conserved "sense" across species with assumed evolutionary advantages, such as facilitating choice of areas with more food and/or more conspecies[29]. However, to the best of our knowledge, it has not been formally tested whether humans adapt their non-symbolic number sense to maximize discrimination accuracy or to maximize fitness, or perhaps rely on less optimal but more flexible resource-limited sampling strategies. The

development of our theory allows the possibility to investigate this fundamental aspect of behaviour.

In Experiment 1, healthy volunteers (n=7) took part in a two-alternative forced choice numerosity task, where each participant completed ~2,400 trials across four consecutive days (methods). On each trial, they were simultaneously presented with two clouds of dots and asked which one contained more dots, and were given feedback on their reward and opportunity losses on each trial (Fig. 3a). Participants were either rewarded for their accuracy (*perceptual* condition, where maximizing the amount of correct responses is the optimal strategy) or the number of dots they selected (*value* condition, where maximizing reward is the optimal strategy). Each condition was tested for two consecutive days with the starting condition randomized across participants. Crucially, we imposed a prior distribution $f(v)$ with a right-skewed quadratic shape (Fig. 3b), whose parametrization allowed tractable analytical solutions of the encoding rules $\theta_A(v)$, $\theta_R(v)$ and $\theta_D(v)$, that correspond to the encoding rules for Accuracy maximization, Reward maximization, and DbS, respectively (Fig. 3e and methods). Qualitative predictions of behavioral performance indicate that the accuracy maximization model is the most accurate for trials with lower numerosities (the most frequent ones), while the reward-maximization model outperforms the others for trials with larger numerosities (trials where the difference in the number of dots in the clouds, and thus the potential reward, is the largest, Figs. 1d and 3f). In contrast, the DbS strategy appears to be a compromise between the other two models, however with lower levels of accuracy and expected reward.

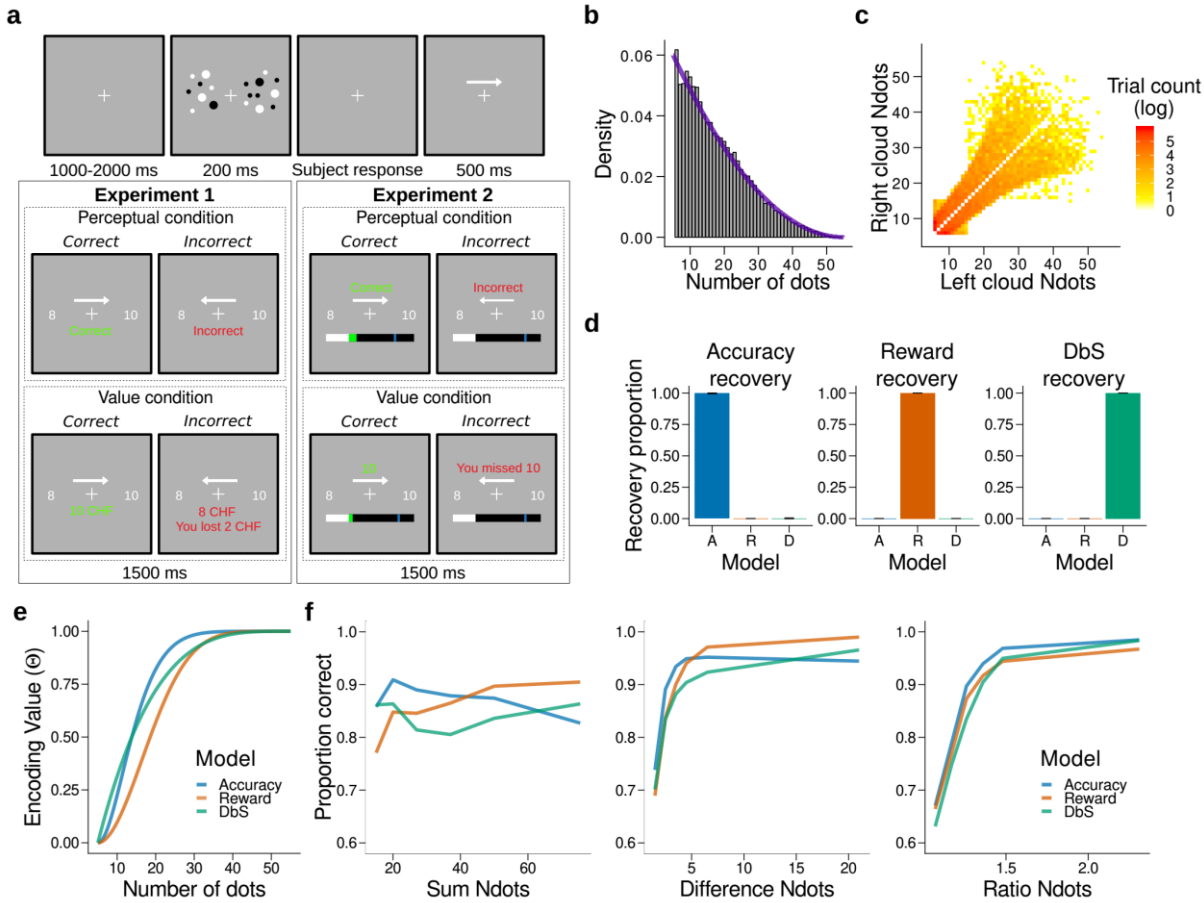


Figure 3. Experimental design, model simulations and recovery. **a**) Schematic task design of Experiments 1 and 2. After a fixation period (1-2s) participants were presented two clouds of dots (200ms) and had to indicate which cloud contained the most dots. Participants were rewarded for being accurate (Perceptual condition) or for the number of dots they selected (Value condition) and were given feedback. In Experiment 2 participants collected on correctly answered trials a number of points equal to a fixed amount (Perception condition) or a number equal to the dots in the cloud they selected (Value condition) and had to reach a threshold of points on each run. **b)** Empirical (grey bars) and theoretical (purple line) distribution of the number of dots in the clouds of dots presented across Experiments 1 and 2. **c)** Distribution of the numerosity pairs selected per trial. **d)** Synthetic data preserving the trial set statistics and number of trials per participant used in Experiment 1 was generated for each encoding rule (Accuracy (left), Reward (middle), and DbS (right)) and then the latent-mixture model was fit to each generated dataset. The figures show that it is theoretically possible to recover each generated encoding rule. **e)** Encoding function $\theta(v)$ for the different sampling strategies as a function of the input values v (i.e., the number of dots). **f)** Qualitative predictions of the three models (blue: Accuracy, red: Reward, green: Decision by Sampling) on trials from experiment 1 with $n = 25$. Performance of each model as a function of the sum of the number of dots in both clouds (left), the absolute difference between the number of dots in both clouds (middle) and the ratio of the number of dots in the most numerous cloud over the less numerous cloud (right).

In our modelling specification, the choice structure is identical for the three different sampling models, differing only in the encoding rule $\theta(v)$ (methods). Therefore, answering the question of which encoding rule is the most favored for each participant can be parsimoniously addressed using a latent-mixture model, where each subject uses $\theta_A(v)$, $\theta_R(v)$ or $\theta_D(v)$ to guide their decisions (methods). Before fitting this model to the empirical data, we confirmed the validity of our model selection approach through a validation procedure using synthetic choice data (Fig. 3d, Supplementary Fig. 2 and methods). After we confirmed that that we can reliably differentiate between our competing encoding rules, the latent-mixture model was initially fit to each condition (perception or value) using a hierarchical Bayesian approach (methods). Surprisingly, we found that at the population level, participants did not follow the accuracy or reward optimization strategy in the respective experimental condition, but favored the DbS strategy (probability that DbS deemed best in the perceptual $p_{DbS_favored}$ = 0.86 and value $p_{DbS_favored}$ = 0.93 conditions, Fig. 4a). Importantly, this population-level result also holds at the individual level: DbS was strongly favored in 6 out of 7 participants in the perceptual condition, and 7 out of 7 in the value condition (Supplementary Fig. 3). These results are not likely to be affected by changes in performance over time, as performance was stable across the four consecutive days (Supplementary Fig. 4). Additionally, we investigated whether biases induced by choice history effects may have influenced our results[30–32]. Therefore, we incorporated both choice- and correctness-dependence history biases in our models and fitted the models once again (methods). We found that inclusion of these choice history biases did not affect our results both at the population and individual levels. Population probability that DbS deemed best in the perceptual $p_{DbS_favored}$ = 0.87 and value $p_{DbS_favored}$ = 0.93 conditions (Fig. 4c). At the individual level, DbS was again strongly favored in 6 out of 7 participants in the perceptual condition, and 7 out of 7 in the value condition (Supplementary Fig. 3).

In order to investigate further the robustness of this effect, we introduced a slight variation in the behavioral paradigm. In this new experiment (Experiment 2), participants were given points on each trial and had to reach a certain threshold in each run for it to be eligible for reward (Fig. 3a and methods). This class of behavioral task is thought to be in some cases more ecologically valid than trial-independent choice paradigms[33]. In this new experiment, either a fixed amount of points for a correct trial was given (perceptual condition) or an amount equal to the number of dots in the chosen cloud if the response was correct (value condition). We recruited a new set of participants (n=6), who were tested on these two conditions, each for two consecutive days with the starting condition randomized across participants (each participant completed ~2,400 trials). The quantitative results revealed once again that

participants did not change their encoding strategy depending on the goals of the task, with DbS being strongly favored for both perceptual and value conditions. At the population level, DbS deemed best in the perceptual ($p_{DbS_favored}$ = 0.999) and value ($p_{DbS_favored}$ = 0.91) conditions (Fig. 4a) and at the individual level, DbS was strongly favored in 6 out of 6 participants in both the perceptual and value conditions (Supplementary Fig. 3). Once again, we found that inclusion of choice history biases in this experiment did not significantly affect our results both at the population and individual levels. Population probability that DbS deemed best in the perceptual ($p_{DbS_favored}$ = 0.999) and value ($p_{DbS_favored}$ = 0.90) conditions (Supplementary Fig. 3), and at the individual level DbS was strongly favored in 6 out of 6 participants in the perceptual condition and 5 of 6 in the value condition (Supplementary Fig. 3). Thus, experiments 1 and 2 strongly suggest that our results are not driven by specific instructions or characteristics of the behavioral task.

As a further robustness check, for each participant we grouped the data in different ways across experiments (Experiments 1 and 2) and experimental conditions (perceptual or value) and investigated which sampling model was favored. We found that irrespective of how the data was grouped, DbS was the model that clearly deemed best at the population (Fig. 4a) and individual level (Supplementary Fig. 5). Additionally, we investigated whether these quantitative results specifically depended on our choice of using a latent-mixture model. Therefore, we also fitted each model independently and compared the quality of the model fits based on out-of-sample cross-validation metrics (methods). Once again, we found that the DbS was favored independently of experiment and conditions (Fig. 4b). Additionally, we investigated whether the DbS model makes more accurate predictions than the widely used logarithmic model of numerosity discrimination tasks[34]. We found that DbS still made better out of sample predictions than the log-model (Fig. 4b, Fig. 5f,g, and Supplementary Fig. 6). Moreover, these results continued to hold after taking into account possible choice history biases (Supplementary Fig. 7). In addition to these quantitative results, qualitatively we also found that behavior closely matched the predictions of the DbS model remarkably well (Fig. 4c), based on virtually only 1 free parameter, namely, the number of samples (resources) n. Together, these results provide compelling evidence that DbS is the most likely resource-constrained sampling strategy used by participants.

Here it is important to emphasize that all sampling models and the logarithmic model of numerosity have the same degrees of freedom (performance is determined by n in the sampling models and Weber's fraction σ in the log model, methods). Therefore, qualitative and quantitative differences favoring the DbS model cannot be explained by differences in model complexity.

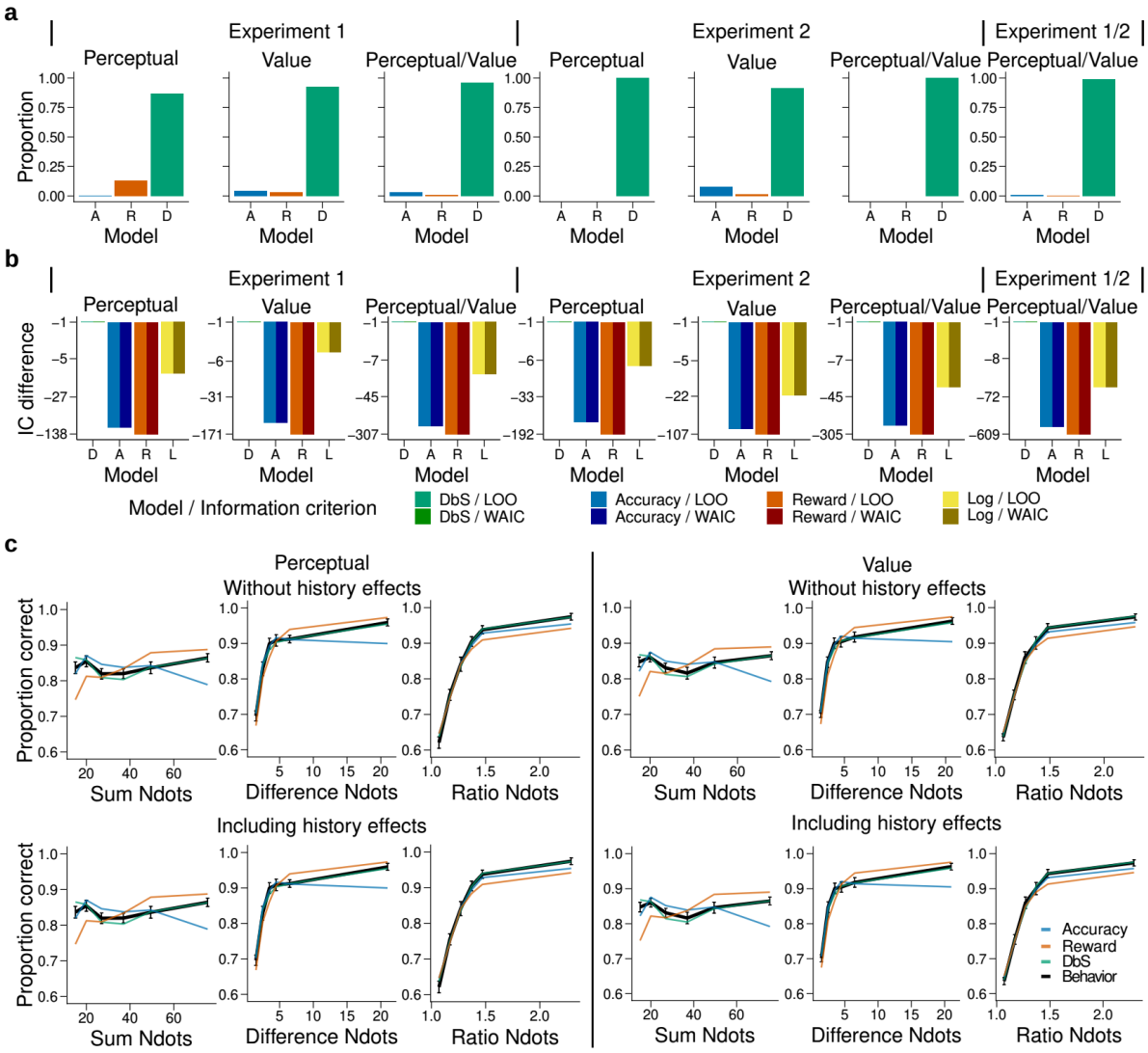
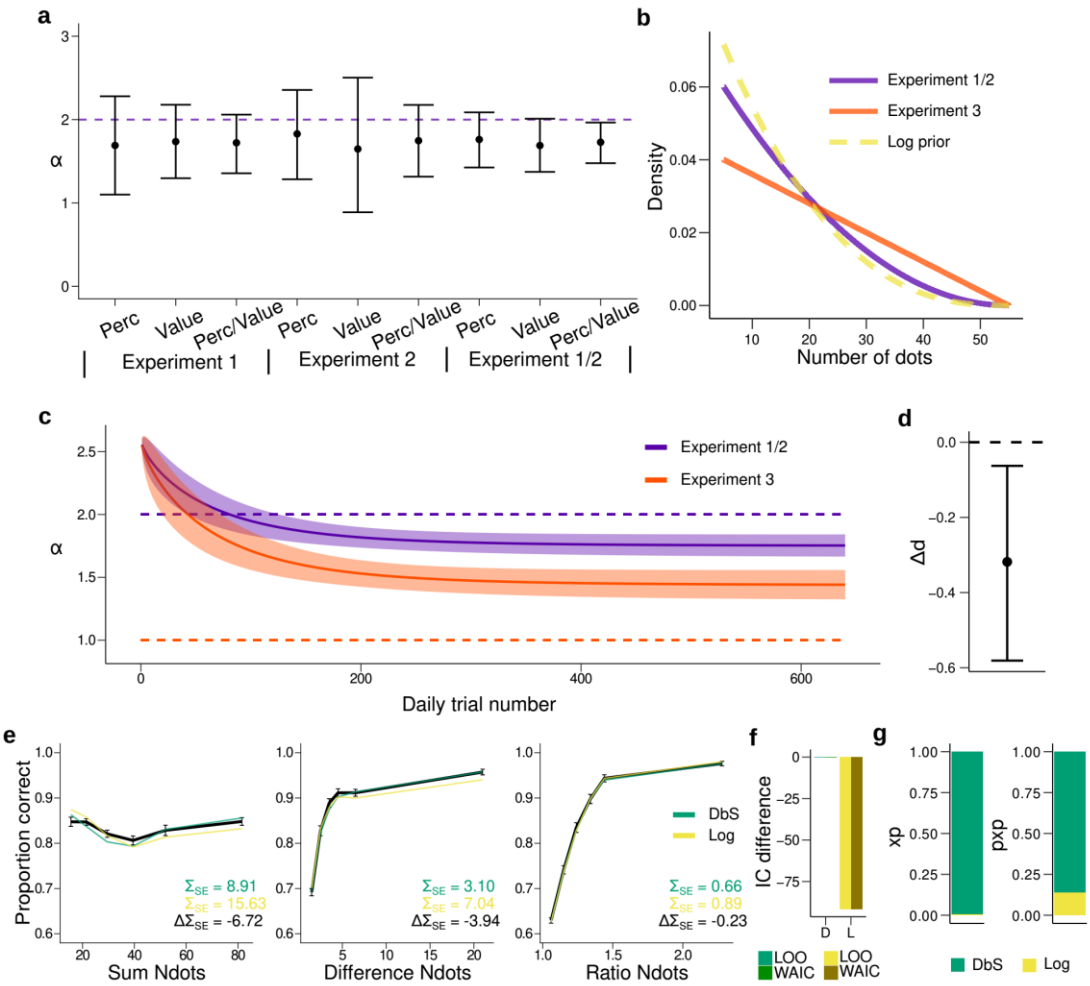


Figure 4. Behavioral results. **a)** Bars represent proportion of times an encoding rule (Accuracy (A, blue), Reward (R, red), DbS (D, green)) was selected by the Bayesian latent-mixture model based on the posterior estimates across participants. Each panel shows the data grouped for each and across experimental conditions and experiments (see titles on top of each panel). The results show that DbS was clearly the favored encoding rule. Individual-level results are presented in Supplementary Fig. 5. **b)** Difference in LOO and WAIC between the best model (DbS (D) in all cases) and the competing models: Accuracy (A), Reward (R) and Logarithmic (L) models. Each panel shows the data grouped for each and across experimental conditions and experiments (see titles on top of each panel). **c)** Behavioral data (black, error bars represent s.e.m. across participants) and model predictions based on fits to the empirical data. Data and model predictions are presented for both the perceptual (left panels) or value (right panels) conditions, and excluding (top panels) or including (bottom) choice history effects. Performance of data model predictions are presented as function of the sum of the number of dots in both clouds (left), the absolute difference between the number of dots in both clouds (middle) and the ratio of the number of dots in the most numerous cloud over the less numerous cloud (right). Results reveal a remarkable overlap of the behavioral data and predictions by DbS, thus confirming the quantitative results presented in panels a and b.

Up to now, fits and comparison across models have been done under the assumption that the participants learned the prior distribution $f(v)$ imposed in our task. If participants are employing DbS, it is important to understand the dynamical nature of adaptation in our task. Note that the shape of the prior distribution (Figs. 1b, 3b and 5b), is determined by parameter α (Eq. 10, methods). First, we made sure based on model recovery analyses that the DbS model could jointly and accurately recover both the shape parameter α and the resource parameter n based on synthetic data (Supplementary Fig. 8). Then we fitted this model to the empirical data and found that the recovered value of the shape parameter α closely followed the value of the empirical prior with a slight underestimation (Fig. 5a). Next, we investigated the dynamics of prior adaptation. To this end, we ran a new experiment (Experiment 3, n=7 new participants) where we set the shape parameter of the prior to a lower value compared to Experiments 1-2 (Fig. 5b, methods). We investigated the change of α over time by allowing this parameter to change with trial experience (Eq. 18, methods) and compared the evolution of α for Experiments 1 and 2 (empirical $\alpha = 2$) with Experiment 3 (empirical $\alpha = 1$, Fig. 5b). If participants show prior adaption in our numerosity discrimination task, we hypothesized that the asymptotic value of α should be lower for Experiments 1-2 than for Experiment 3. First, we found that for Experiments 1-2, the value of α quickly reached an asymptotic value close to the target value (Fig. 5c). On the other hand, for Experiment 3 the value of α continued to decrease during the experimental session, but slowly approaching its target value. This seemingly slower adaptation to the shape of the prior in Experiment 3 might be explained by the following observation. The prior parametrized with $\alpha = 1$ in Experiment 3, is further away from an agent hypothesized to have a natural numerosity discrimination based on a log scale ($\alpha = 2.58$, Fig. 5b and methods), which is closer in value to the shape of the prior in Experiments 1 and 2 ($\alpha = 2$). This result is in line with previous DbS studies showing that adaptation to "unnatural" priors by sensory systems is slower[35] and may require many more trials or training experience than it is possible in laboratory experiments. Irrespective of these considerations, the key result to confirm our adaptation hypothesis is that the asymptotic value of α is lower for Experiment 3 compared to Experiments 1 and 2 ($P_{MCMC} = 0.006$; Fig. 5c,d). Additionally, we found that this DbS model again provides more accurate qualitative and quantitative out of sample predictions than the log model (Fig. 5e,f).



494

495 ***Figure 5.*** Prior adaptation analyses. **a)** Estimation of the shape parameter α for the DbS model
496 by grouping the data for each and across experimental conditions and experiments. Error bars
497 represent the 95% highest density interval of the posterior estimate of α at the population
498 level. The dashed line shows the theoretical value of α. **b)** Theoretical prior distribution $f(v)$
499 in Experiments 1 and 2 ($\alpha = 2$, purple) and 3 ($\alpha = 1$, orange). The dashed line represents the
500 value of α of our prior parametrization that approximates the DbS and log discriminability
501 models. **c)** Posterior estimation of the α_t adaptation model for α_t (Eq. 18) as a function of the
502 number of trials t in each daily session for Experiments 1 and 2 (purple) and Experiment 3
503 (orange). The results reveal that, as expected, α_t reaches a lower asymptotic value δ. Error
504 bars represent ± 1 s.d. of 3,000 simulated α_t values drawn from the posterior estimates of the
505 HBM (see methods). **d)** Difference in the δ parameter between Experiments 1-2 and
506 Experiment 3 based on the posterior parameter estimates of the HBM. This analysis reveals a
507 significant difference ($P_{MCMC} = 0.006$). Error bars represent the 95% highest density interval of
508 the posterior differences in the HBM. **e)** Behavioral data (black) and model fit predictions of
509 the DbS (green) and Log (yellow) models. Performance of each model as a function of the sum
510 of the number of dots in both clouds (left), the absolute difference between the number of
511 dots in both clouds (middle) and the ratio of the number of dots in the most numerous cloud
512 over the less numerous cloud (right). Error bars represent s.e.m. **f)** Difference in LOO and WAIC
513 between the best fitting DbS (D) and Logarithmic (Log) model. **g)** Population exceedance
514 probabilities (xp, left) and protected exceedance probabilities (pxp, right) for DbS (green) vs
515 Log (yellow) of a Bayesian model selection analysis[36]: xp_{DbS}=0.99, pxp_{DbS}=0.87. These results
516 provide a clear indication that the adaptive DbS explains the data better than the Log model.

517 ***Discussion:*** The brain is a metabolically expensive inference machine[37,38]. Therefore it
518 has been suggested that evolutionary pressure has driven it to make productive use of
519 its limited resources by exploiting statistical regularities[2,22,28,39]. Here, we incorporate
520 this important – often ignored – aspect in models of behavior by introducing a general
521 framework of decision-making under the constraints that the system: (i) encodes
522 information based on binary codes, and (ii) has limited number of samples available to
523 encode information. Crucially, this framework takes into consideration the statistics of
524 the environment, and we provide a general solution for the specific goals that
525 systems/organism may require to guide optimal discriminations (e.g. maximize
526 perceptual accuracy or expected reward) given the above-mentioned constraints.
527 When the organism must rely on internal "noisy" percepts to guide behavior, we show
528 that the encoding rule that maximizes mutual information is the same rule that
529 maximizes decision accuracy in two-alternative decision tasks. This happens to be the
530 case for accuracy maximization, however, note that there is nothing privileged about
531 maximizing mutual information, as it does not mean that the goals of the organism are
532 necessarily achieved[20]. In fact, we show that if the goal of the organism is instead to
533 maximize expected reward, the system should not rely on maximizing information
534 transmission to fulfill this goal and must give up precision in information coding.

535 Another key implication of our work is that it provides an alternative explanation
536 to the usual conception of noise as the main cause of behavioral performance
537 degradation, where noise is usually artificially added to models of decision behavior to
538 generate the desired variability[40,41]. On the contrary, our work makes it formally
539 explicit why a system that evolved to encode information based on binary codes must
540 be necessarily noisy, also revealing how the system could take advantage of its
541 unavoidable noisy properties[12] to optimize decision behavior[13]. While the idea of noise
542 as a facilitator of perceptual processes is not new[13,26,42] (e.g. stochastic resonance
543 phenomena), our formal analysis provides a unique view of the necessity of noise to
544 optimize decision behavior in capacity-limited systems that rely on information coding
545 based on binary samples. Interestingly, this notion appears to be consistent with the
546 recent controversial finding that dynamics of LIP neurons likely reflect binary (discrete)
547 coding states to guide decision behavior[18,27]. Based on this potential link between our
548 and their work, our theoretical framework generates testable predictions that could
549 be investigated in future neurophysiological work. For instance, noise distribution in
550 neural circuits should dynamically adapt according to the prior distribution of inputs
551 and goals of the organism. Consequently, the rate of "step-like" coding in single
552 neurons[27] should also be dynamically adjusted (perhaps optimally) to statistical
553 regularities and behavioral goals.

Our theory is closely related to DbS, which is an influential account of decision behavior derived from principles of retrieval and memory comparison by taking into account the regularities of the environment[28]. We show that DbS represents a rule that is similar to (though not exactly like) the optimal encoding rule for each of a variety of different possible frequency distributions encountered in the recent past, and that it automatically adjusts the encoding in response to changes in the frequency distribution from which exemplars are drawn in approximately the right way, while providing a simple answer to the question of how such adaptation of the encoding rule to a changing frequency distribution occurs. Crucially, here we demonstrate that DbS does not guarantee maximization of decision accuracy or fitness maximization. In other words, DbS appears to be – in principle – a "suboptimal" strategy. In line with this result, a recent work demonstrated that noise added after having "efficiently" encoded input values (based on the DbS encoding function) improves information transmission and helps to account for context sensitivity adjustments[39]. Our framework provides a formal explanation for this result, however with a different neural implementational perspective: the prior is directly embedded in an optimal noise distribution, which "corrupts" the input of each neuron that in turn generates the noisy but optimal percept. Irrespective of these considerations, here it is important to emphasize that all novel contributions made by DbS to explain the shape of ubiquitous psycho-economic functions (e.g. utility concavity, loss aversion, and risk probability curves[28]) continue to hold under our framework, with the important added value that we provide explicit optimal solutions to maximize accuracy and reward expectation.

We tested these resource-limited coding frameworks in non-sysmbolic numerosity discrimination, a fundamental cognitive function for behavior in humans and other animals, which may have emerged during evolution to support fitness maximization. Here, we find that the way in which the precision of numerosity discrimination varies with the size of the numbers being compared is consistent with the hypothesis that the internal representations on the basis of which comparisons are made are sample-based. In particular, we find that the encoding rule varies depending on the frequency distribution of values encountered in a given environment, and that this adaptation occurs fairly quickly once the frequency distribution changes.

This adaptive character of the encoding rule differs, for example, from the common hypothesis of a logarithmic encoding rule (independent of context), which we show fits our data less well. Nonetheless, we can reject the hypothesis of full optimality of the encoding rule for each distribution of values used in our experiments, even after subjects have had extensive experience with a given distribution. Thus, a possible explanation of why DbS is the favored model in our numerosity task is that accuracy and reward maximization requires optimal adaptation of the noise distribution based

592 on our imposed prior, requiring complex neuroplastic changes to be implemented,
593 which are in turn metabolically costly[43]. Relying on samples from memory might be less
594 metabolically costly as these systems are plastic in short time scales, and therefore a
595 relatively simpler heuristic to implement allowing faster adaptation.

596 Interestingly, it was recently shown that in a reward learning task, a model that
597 estimates values based on memory samples from recent past experiences can explain
598 the data better than canonical incremental learning models[44]. Based on their and our
599 findings, we conclude that sampling from memory is a potentially powerful mechanism
600 for guiding choice behavior, as it allows quick learning and generalization of
601 environmental contexts based on recent experience. However, it should be noted that
602 relying on such mechanisms alone might be suboptimal from a performance- and goal-
603 based point of view, where neural calibration of optimal strategies may require
604 extensive experience, possibly via direct interactions between memory and sensory
605 systems[45,46].

606 Taken together, our findings emphasize the need of studying optimal models,
607 which serve as anchors to understand the brain's computational goals without ignoring
608 the fact that biological systems are limited in their capacity to process information. We
609 addressed this by proposing a computational problem, elaborating an algorithmic
610 solution, and proposing a minimalistic implementational architecture that solves the
611 resource-constrained problem. This is essential, as it helps to establish frameworks
612 that allow comparing behavior not only across different tasks and goals, but also across
613 different levels of description, for instance, from single cell operation to observed
614 behavior[24]. We argue that this approach is fundamental to provide benchmarks for
615 human performance that can lead to the discovery of alternative heuristics[47] that could
616 also be suboptimal under resource-constrained frameworks – as we demonstrate here
617 for DbS – but that might be in turn the optimal strategy to implement if one considers
618 costs of optimal adaptation[25,48]. We conclude that the understanding of brain function
619 under a principled research agenda, which takes into account decision mechanisms
620 that are biologically feasible, will be essential to accelerate the elucidation of the
621 mechanisms underlying human cognition.

METHODS

Participants. The study tested young healthy volunteers with normal or corrected-to-normal vision (total n = 20, age 19–36 years, 9 females: n = 7 in experiment 1, 2 females; n = 6 new participants in experiment 2, 3 females; n = 7 new participants in experiment 3, 4 females). Participants were randomly assigned to each experiment and no participant was excluded from the analyses. Participants were instructed about all aspects of the experiment and gave written informed consent. None of the participants suffered from any neurological or psychological disorder or took medication that interfered with participation in our study. Participants received monetary compensation for their participation in the experiment partially related to behavioral performance (see below). The experiments conformed to the Declaration of Helsinki and the experimental protocol was approved by the Ethics Committee of the Canton of Zurich.

Experiment 1: Participants (n=7) carried out a numerosity discrimination task for four consecutive days for approximately one hour per day. Each daily session consisted of a training run followed by 8 runs of 75 trials each. Thus, each participant completed ~2,400 trials across the four days of experiment.

After a fixation period (1-1.5s jittered), two clouds of dots (left and right) were presented on the screen for 200ms. Participants were asked to indicate the side of the screen where they perceived more dots. Their response was kept on the screen for 1 s followed by feedback consisting of the symbolic number of dots in each cloud as well as the monetary gains and opportunity losses of the trial depending on the experimental condition (Fig. 3a). In the *value* condition, participants were explicitly informed that each dot in a cloud of dots corresponded to 1 Swiss Franc (CHF). Participants were informed that they would receive the amount in CHF corresponding to the total number of dots on the chosen side (Fig. 3a). At the end of the experiment a random trial was selected and they received the corresponding amount. In the *accuracy* condition, participants were explicitly informed that they could receive a fixed reward (15 Swiss Francs (CHF)) for each correct trial. This fixed amount was selected such that it approximately matched the expected reward received in the value condition (as tested in pilot experiments). At the end of the experiment, a random trial was selected and they would receive this fixed amount if they chose the cloud with more dots (i.e. the correct side). Each condition lasted for two consecutive days with the starting condition randomized across participants. Only after completing all four experiment days, participants were compensated for their time with 20 CHF per hour, in addition to the money obtained based on their decisions on each experimental day.

Experiment 2: Participants (n=6) carried out a numerosity discrimination task where each of four daily sessions consisted of 16 runs of 40 trials each, thus each participant completed ~2,560 trials. A key difference with respect to Experiment 1 is that participants had to accumulate points based on their decisions and had to reach a predetermined threshold on each run. The rules of point accumulation depended on the experimental condition. In the *perceptual* condition, a fixed amount of points was awarded if the participants chose the cloud with more dots. In this condition, participants were instructed to accumulate a number of points and reach a threshold given a limited number of trials. Based on the results obtained in Experiment 1, the threshold corresponded to 85% of correct trials in a given run, however the participants were unaware of this. If the participants reached this threshold, they were eligible for a fixed reward (20 CHF) as described in Experiment 1. In the value condition, the number of points received was equal to the number of dots in the cloud, however, contrary to experiment 1, points were only awarded if the participant chose the cloud with the most dots. Participants had to reach a threshold that was matched in the expected collection of points of the perceptual condition. As in Experiment 1, each condition lasted for two consecutive days with the starting condition randomized across participants. Only after completing all the four days of experiment, participants were compensated for their time with 20 CHF per hour, in addition to the money obtained based on their decisions on each experimental day.

Experiment 3: The design of Experiment 3 was similar to the value condition of Experiment 2 (n=7 participants) and was carried out over three consecutive days. The key difference between Experiment 3 and Experiments 1-2 was the shape of the prior distribution $f(v)$ that was used to draw the number of dots for each cloud in each trial (see below).

Stimuli statistics and trial selection: For all experiments, we used the following parametric form of the prior distribution

$$f(v) = c(1 - v)^{\alpha}, \qquad \text{(Eq. 10)}$$

initially defined in the interval [0,1] for mathematical tractability in the analytical solution of the encoding rules $\theta(v)$ (see below), with $\alpha > 0$ determining the shape of the distribution, and c is a normalizing constant. For Experiments 1 and 2 the shape parameter was set to $\alpha = 2$, and for Experiment 3 was set to $\alpha = 1$. i.i.d. samples drawn from this distribution where then multiplied by 50, added an offset of 5, and finally were rounded to the closest integer (i.e., the numerosity values in our experiment ranged from $v_{\min} = 5$ to $v_{\max} = 55$). The pairs of dots on each trial were determined by sampling from a uniform density window in the CDF space (Eq. 10 is its corresponding PDF). The pairs of dots in each trial were selected with the conditions

that, first, their distance in the CDF space was less than a constant (0.25, 0.28 and 0,23 for Experiments 1, 2 and 3 respectively), and second, the number of dots in both clouds was different. Figure 3c illustrates the probability that a pair of choice alternatives was selected for a given trial in Experiments 1 and 2.

Power analyses and model recovery: Given that adaptation dynamics in sensory systems often require long-term experience with novel prior distributions, we opted for maximizing the number of trials for a relatively small number of participants per experiment, as it is commonly done for this type of psychophysical experiments[49–51]. Note that based on the power analyses described below, we collected in total ~42,000 trials across the three Experiments, which is above the average number of trials typically collected in human studies.

In order to maximize statistical power in the differentiation of the competing encoding rules, we generated 10,000 sets of experimental trials for each encoding rule and selected the sets of trials with the highest discrimination power (i.e. largest differences in -Log-Likelihood) between the encoding models. In these power analyses, we also investigated what was the minimum number of trials that would allow accurate generative model selection at the individual level. We found that ~1,000 trials per participant in each experimental condition would be sufficient to predict accurately ($P>0.95$) the true generative model. Based on these analyses, we decided to collect at least 1,200 trials per participant and condition (perceptual and value) in each of the three experiments. Model recovery analyses presented in Figure 3d illustrate the result of our power analyses (see also Supplementary Figure 2).

Apparatus: Eyetracking (EyeLink 1000 Plus) was used to check the participants fixation during stimulus presentation. When participants blinked or move their gaze (more than 2° of visual angle) away from the fixation cross during the stimulus presentation the trial was canceled (only 212 out of 42,000 trials were aborted, i.e., < 0.006% of the trials). Participants were informed when a trial was aborted and were encouraged not to do so as they would not receive any reward for this trial. A chinrest was used to keep the distance between the participants and the screen constant (55cm). The task was run using Psychtoolbox Version 3.0.14 on Matlab 2018a. The diameter of the dots varied between 0.42° and 1.45° of visual angle. The center of each cloud was positioned 12.6° of visual angle horizontally from the fixation cross and had a maximum diameter of 19.6° of visual angle. Following previous numerosity experiments[52,53], either the average dot size or the total area covered by the dots was maintained constant in both clouds for each trial. The color of each dot (white or black) was randomly selected for each dot. Stimuli set were different for each participant but identical between the two conditions.

Encoding rules and model fits: The parametrization of the prior $f(v)$ (Eq. 10) allows tractable analytical solutions of the encoding rules $\theta_P(v)$, $\theta_R(v)$ and $\theta_D(v)$, that correspond to Accuracy maximization, Reward maximization, and DbS, respectively:

- Accuracy maximization: $$\theta_{\mathrm{A}}(v) = \sin\left[\frac{\pi}{2}(1-(1-v)^{\alpha+1}\right]^2 \quad \text{(Eq. 11)}$$
- Reward maximization: $$\theta_{\mathrm{R}}(v) = \sin\left[\frac{\pi}{2}(1+(v-1)((1-v)^{\alpha})^{2/3}\right]^2 \quad \text{(Eq. 12)}$$
- DbS: $$\theta_{\mathrm{D}}(v) = 1-(1-v)^{\alpha+1} \quad \text{(Eq. 13)}$$

Graphical representation of the respective encoding rules is shown in Fig. 3e for Experiments 1 and 2. Given an encoding rule $\theta(v)$, we now define the decision rule. The goal of the decision maker in our task is always to decide which of two input values v_1 and v_2 is larger. Therefore, the agent choses v_1 if and only if the internal readings $k_1 > k_2$. Following the definitions of expected value and variance of binomial variables, and approximating for large n (see Supplementary Note 3), the probability of choosing v_1 is given by

$$\mathrm{P}_{\text{choose } v_1} \approx \Phi\left(\frac{\theta_1-\theta_2}{\sqrt{\frac{\theta_1(1-\theta_1)+\theta_2(1-\theta_2)}{n}}}\right), \quad \text{(Eq. 14)}$$

where $\Phi(\cdot)$ is the standard CDF, and θ_1 and θ_2 are the encoding rules for the input values v_1 and v_2, respectively. Thus, the choice structure is the same for all models, only differing in their encoding rule. The three models generate different qualitative performance predictions for a given number of samples n (Fig. 3f).

Crucially, this probability decision rule (Eq. 14) can be parsimoniously extended to include potential side biases independent of the encoding process as follows

$$\mathrm{P}_{\text{choose } v_1} \approx \Phi\left(\frac{\theta_1-\theta_2}{\sqrt{\frac{\theta_1(1-\theta_1)+\theta_2(1-\theta_2)}{n}}}+\beta_0\right), \quad \text{(Eq. 15)}$$

where β_0 is the bias term. This is the base model used in our work. We were also interested in studying whether choice history effects[31,32] may have influence in our task, thus possibly affecting the conclusions that can be drawn from the base model. Therefore, we extended this model to incorporate the effect of decision learning and choices from the previous trial

$$\mathrm{P}_{\text{choose } v_1} \approx \Phi\left(\frac{\theta_1 - \theta_2}{\sqrt{\frac{\theta_1(1-\theta_1)+\theta_2(1-\theta_2)}{n}}} + \beta_0 + \beta^{\mathrm{L}} a_{t-1} r_{t-1} + \beta^{\mathrm{Ch}} a_{t-1}\right), \quad \text{(Eq. 16)}$$

where a_{t-1} is the choice made on the previous trial (+1 for left choice and -1 for right choice) and r_{t-1} is the "outcome learning" on the previous trial (+1 for correct choice and -1 for incorrect choice). β^{L} and β^{Ch} capture the effect of decision learning and choice in the previous trial, respectively.

Given that the choice structure is the same for all three sampling models considered here, we can naturally address the question of what decision rule the participants favor via a latent-mixture model. We implemented this model based on a hierarchical Bayesian modelling (HBM) approach. The base-rate probabilities for the three different encoding rules at the population level are represented by the vector $\boldsymbol{\pi}$, so that π_m is the probability of selecting encoding rule model m. We initialize the model with an uninformative prior given by

$$\boldsymbol{\pi} \sim \mathrm{Dirichlet}(1_{m=1}, 1_{m=2}, 1_{m=3}).$$

This base-rate is updated based on the empirical data, where we allow each participant s to draw from each model categorically based on the updated base-rate

$$m_s \sim \mathrm{Categorical}(\boldsymbol{\pi}),$$

where the encoding rule θ for model m is given by

$$\theta_{m,s} = \begin{cases} \theta_A \text{ (Eq. 11)}, & m = 1 \\ \theta_R \text{ (Eq. 12)}, & m = 2 \\ \theta_D \text{ (Eq. 13)}, & m = 3 \end{cases}$$

The selected rule was then fed into equations 15 or 16 to determine the probability of selecting a cloud of dots. The number of samples n was also estimated within the same HBM with population mean μ and standard deviation σ initialized based on uninformative priors with plausible ranges

$$\mu_n \sim \mathrm{Uniform}(1,1000)$$
$$\sigma_n \sim \mathrm{Uniform}(0.01,100)$$

allowing each participant s to draw from this population prior assuming that n is normally distributed at the population level

$$n_s \sim \mathrm{Normal}(\mu_n, \sigma_n),$$

Similarly, the latent variables β in equations 15 and 16 were estimated by setting population mean μ_β and standard deviation σ_β initialized based on uninformative priors

$$\mu_\beta \sim \mathrm{Uniform}(-10,10)$$

$$\sigma_\beta \sim \text{Uniform}(0.01,100)$$

allowing each participant s to draw from this population prior assuming that β is normally distributed at the population level

$$\beta_s \sim \text{Normal}(\mu_\beta, \sigma_\beta).$$

In all the results reported in Figures 3 and 4, the value of the shape parameter of the prior was set to its true value $\alpha = 2$. The estimation of α in Figure 5a was investigated with a similar hierarchical approach, allowing each participant to sample from the normal population distribution with uninformative priors over the population mean and standard deviation

$$\mu_\alpha \sim \text{Uniform}(0,01,20)$$
$$\sigma_\alpha \sim \text{Uniform}(0.0001,100).$$

The choice rule of the standard logarithmic model of numerosity discrimination is given by:

$$\text{P}_{\text{choose } v_1} = \Phi\left(\frac{\log(v_1) - \log(v_2)}{\sigma\sqrt{2}}\right), \quad \text{(Eq. 17)}$$

where σ is the internal noise in the logarithmic space. This model was extended to incorporate bias and choice history effects in the same way as implemented in the sampling models. Here we emphasize that all sampling and log models have the same degrees of freedom, where performance is mainly determined by n in the sampling models and Weber's fraction σ in the log model, and biases are determined by parameters β. For all above-mentioned models, the trial-by-trial likelihood of the observed choice (i.e. the data) given probability of a decision was based on a Bernoulli process

$$y_{t,s} \sim \text{Bernoulli}(P_{\text{choose } v_1}),$$

where $y_{t,s} \in \{0,1\}$ is the decision of each participant s in each trial t. In order to allow for prior adaptation, the model fits presented in Figures 3 and 4 were fit starting after a fourth of the daily trials (corresponding to 150 trials for experiment 1 and 160 trials for experiment 2) to allow for prior adaptation and fixing the shape parameter to its true generative value $\alpha = 2$.

The dynamics of adaptation (Fig. 5) were studied by allowing the shape parameter α to evolve through trial experience using all trials collected on each experiment day. This was studied using the following function

$$\alpha_t = \delta + \eta e^{-t/\tau}, \qquad \text{(Eq. 18)}$$

where δ represents a possible target adaptation value of α, t is the trial number, and η, τ determine the shape of the adaptation. Therefore, the encoding rule of the DbS model also changed trial-to-trial

$$\theta_{\mathrm{D}}^{t}(v) = 1 - (1 - v)^{\alpha_t + 1}. \qquad \text{(Eq. 19)}$$

Adaptation was tested based on the hypothesis that participants initially use a logarithmic discrimination rule (Eq. 17) (this strategy also allowed improving identification of the adaptation dynamics). Therefore, Eq. 18 was parametrized such that the initial value of the shape parameter ($\alpha_{t=0}$) guaranteed that discriminability between the DbS and the logarithmic rule was as close as possible. This was achieved by finding the value of α in the DbS encoding rule (θ_D, Eq 13) that minimizes the following expression

$$\sum_{t=1}^{T}\left[\left(\frac{\theta_D(v_{1,t}) - \theta_D(v_{2,t})}{\sqrt{\theta_D(v_{1,t})(1-\theta_D(v_{1,t})) + \theta_D(v_{2,t})(1-\theta_D(v_{2,t}))}}\right) - (\log(v_{1,t}) - \log(v_{2,t}))\right]^2, \text{(Eq. 20)}$$

where $v_{1,t}$ and $v_{2,t}$ are the numerosity inputs for each trial t. This expression was minimized based on all trials generated in Experiments 1-3 (note that minimizing this expression does not require knowledge of the sensitivity levels σ and n for the log and DbS models, respectively). We found that the shape parameter value that minimizes Eq. 20 is $\alpha = 2.58$. Based on our prior $f(v)$ parametrization (Eq. 10), this suggests that the initial prior is more skewed than the priors used in Experiments 1-3 (Fig. 5b). This is an expected result given that log-normal priors – typically assumed in numerosity tasks – are also highly skewed. We fitted the δ parameter independently for Experiments 1-2 and Experiments 3 but kept the τ parameter shared across all experiments. If adaptation is taking place, we hypothesized that the asymptotic value δ of the shape parameter α should be larger for Experiments 1-2 compared to Experiment 3.

Posterior inference of the parameters in all the hierarchical models described above was performed via the Gibbs sampler using the Markov Chain Monte Carlo (MCMC) technique implemented in JAGS. For each model, a total of 50,000 samples were drawn from an initial burn-in step and subsequently a total of new 50,000 samples were drawn for each of three chains (samples for each chain were generated based on a different random number generator engine, and each with a different seed). We applied a thinning of 50 to this final sample, thus resulting in a final set of 1,000 samples for each chain (for a total of 3,000 pooling all 3 chains). We conducted Gelman–Rubin tests for each parameter to confirm convergence of the chains. All latent variables in our Bayesian models had $\hat{R} < 1.05$, which suggests that all three chains converged to a target posterior distribution. We checked via visual inspection

that the posterior population level distributions of the final MCMC chains converged to our assumed parametrizations. When evaluating different models, we are interested in the model's predictive accuracy for unobserved data, thus it is important to choose a metric for model comparison that considers this predictive aspect. Therefore, in order to perform model comparison, we used a method for approximating leave-one-out cross-validation (LOO) that uses samples from the full posterior[54]. These analyses were repeated using an alternative Bayesian metric: the WAIC[54].

Data availability

The data that support the findings of this study will be made available via github.

Code availability

Code that support the findings of this study will be made available via github.

Acknowledgements

This work was supported by an ERC starting grant (ENTRAINER) to R.P and by a grant of the U.S. National Science Foundation to M.W. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. 758604).

Author contributions

R.P. and M.W. developed the theoretical framework. J.H. and R.P. designed the experiments. J.H. collected and analyzed the data. All authors interpreted the results and wrote the manuscript.

Competing Interests Statement

The authors declare no competing financial interests.

References

1. Wei, X.-X. & Stocker, A. A. A Bayesian observer model constrained by efficient coding can explain'anti-Bayesian'percepts. *Nat. Neurosci.* **18,** 1509–1517 (2015).

2. Polanía, R., Woodford, M. & Ruff, C. C. Efficient coding of subjective value. *Nat. Neurosci.* **22,** 134–142 (2019).

3. Ganguli, D. & Simoncelli, E. P. Efficient sensory encoding and Bayesian inference with heterogeneous neural populations. *Neural Comput.* **26,** 2103–2134 (2014).

4. Louie, K. & Glimcher, P. W. Efficient coding and the neural representation of value. *Ann. N. Y. Acad. Sci.* **1251,** 13–32 (2012).

5. Sims, C. R. Efficient coding explains the universal law of generalization in human perception. *Science* **360,** 652–656 (2018).

6. Rustichini, A., Conen, K. E., Cai, X. & Padoa-Schioppa, C. Optimal coding and neuronal adaptation in economic decisions. *Nat. Commun.* **8,** 1208 (2017).

7. Summerfield, C. & Tsetsos, K. Do humans make good decisions? *Trends Cogn. Sci.* **19,** 27–34 (2014).

8. Mainen, Z. F. & Sejnowski, T. J. Reliability of spike timing in neocortical neurons. *Science* **268,** 1503–6 (1995).

9. Stringer, C., Pachitariu, M., Steinmetz, N., Carandini, M. & Harris, K. D. High-dimensional geometry of population responses in visual cortex. *Nature* 1 (2019). doi:10.1038/s41586-019-1346-5

10. Drugowitsch, J., Wyart, V., Devauchelle, A.-D. & Koechlin, E. Computational Precision of Mental Inference as Critical Source of Human Choice Suboptimality. *Neuron* **92,** 1398–1411 (2016).

11. Beck, J. M., Ma, W. J., Pitkow, X., Latham, P. E. & Pouget, A. Not noisy, just wrong: the role of suboptimal inference in behavioral variability. *Neuron* **74,** 30–9 (2012).

12. Faisal, A. A., Selen, L. P. J. & Wolpert, D. M. Noise in the nervous system. *Nat. Rev. Neurosci.* **9,** 292–303 (2008).

13. Tsetsos, K. *et al.* Economic irrationality is optimal during noisy decision making. *Proc. Natl. Acad. Sci.* **113,** 3102–3107 (2016).

14. Girshick, A. R., Landy, M. S. & Simoncelli, E. P. Cardinal rules: visual orientation perception reflects knowledge of environmental statistics. *Nat. Neurosci.* **14,** 926–932 (2011).

15. Wang, Z., Stocker, A. A. & Lee, D. D. Efficient Neural Codes That Minimize L p Reconstruction Error. *Neural Comput.* **28,** 2656–2686 (2016).

16. Khaw, M. W., Glimcher, P. W. & Louie, K. Normalized value coding explains dynamic adaptation in the human valuation process. *Proc. Natl. Acad. Sci. U. S. A.* **114,** 12696–12701 (2017).

17. Schreiber, S., Machens, C. K., Herz, A. V. M. & Laughlin, S. B. Energy-Efficient Coding with Discrete Stochastic Events. *Neural Comput.* **14,** 1323–1346 (2002).

18. Latimer, K. W., Yates, J. L., Meister, M. L. R., Huk, A. C. & Pillow, J. W. Single-trial spike trains in parietal cortex reveal discrete steps during decision-making. *Science* **349,** 184–7 (2015).

19. Clarke, B. S. & Barron, A. R. Jeffreys' prior is asymptotically least favorable under entropy risk. *J. Stat. Plan. Inference* **41,** 37–60 (1994).

20. Park, I. M. & Pillow, J. W. Bayesian Efficient Coding. *bioRxiv* 178418 (2017). doi:10.1101/178418

21. Ganguli, D. & Simoncelli, E. P. Neural and perceptual signatures of efficient sensory coding. (2016).

22. Wei, X.-X. & Stocker, A. A. Lawful relation between perceptual bias and discriminability. *Proc. Natl. Acad. Sci. U. S. A.* **114,** 10244–10249 (2017).

23. Pirrone, A., Stafford, T. & Marshall, J. A. R. When natural selection should optimize speed-accuracy trade-offs. *Front. Neurosci.* **8,** 73 (2014).

24. Marr, D. Vision: A computational investigation into the human representation and processing of visual information. (1982).

25. Lieder, F. & Griffiths, T. L. Resource-rational analysis: understanding human cognition as the optimal use of limited computational resources. *Behav. Brain Sci.* 1–85 (2019). doi:10.1017/S0140525X1900061X

26. McDonnell, M. D., Stocks, N. G. & Abbott, D. Optimal stimulus and noise distributions for information transmission via suprathreshold stochastic resonance. *Phys. Rev. E* **75,** 061105 (2007).

27. Zoltowski, D. M., Latimer, K. W., Yates, J. L., Huk, A. C. & Pillow, J. W. Discrete Stepping and Nonlinear Ramping Dynamics Underlie Spiking Responses of LIP Neurons during Decision-Making. *Neuron* **102,** 1249–1258.e10 (2019).

28. Stewart, N., Chater, N. & Brown, G. D. A. Decision by sampling. *Cogn. Psychol.* **53,** 1–26 (2006).

29. Butterworth, B., Gallistel, C. R. & Vallortigara, G. Introduction: The origins of numerical abilities. *Philos. Trans. R. Soc. B Biol. Sci.* **373,** 20160507 (2018).

958 30. Keung, W., Hagen, T. A. & Wilson, R. C. Regulation of evidence accumulation by
959 pupil-linked arousal processes. *Nat. Hum. Behav.* 1 (2019). doi:10.1038/s41562-
960 019-0551-4

961 31. Abrahamyan, A., Silva, L. L., Dakin, S. C., Carandini, M. & Gardner, J. L.
962 Adaptable history biases in human perceptual decisions. *Proc. Natl. Acad. Sci.*
963 **113,** E3548–E3557 (2016).

964 32. Talluri, B. C., Urai, A. E., Tsetsos, K., Usher, M. & Donner, T. H. Confirmation Bias
965 through Selective Overweighting of Choice-Consistent Evidence. *Curr. Biol.* **28,**
966 3128–3135.e8 (2018).

967 33. Kolling, N., Wittmann, M. & Rushworth, M. F. S. F. S. Multiple Neural
968 Mechanisms of Decision Making and Their Competition under Changing Risk
969 Pressure. *Neuron* **81,** 1190–202 (2014).

970 34. Dehaene, S. The neural basis of the Weber-Fechner law: a logarithmic mental
971 number line. *Trends Cogn. Sci.* **7,** 145–147 (2003).

972 35. Stewart, N., Reimers, S. & Harris, A. J. L. On the Origin of Utility, Weighting, and
973 Discounting Functions: How They Get Their Shapes and How to Change Their
974 Shapes. *Manage. Sci.* **61,** 687–705 (2015).

975 36. Stephan, K. E., Penny, W. D., Daunizeau, J., Moran, R. J. & Friston, K. J. Bayesian
976 model selection for group studies. *Neuroimage* **46,** 1004–1017 (2009).

977 37. Hawkes, K. *et al.* Grandmothering, menopause, and the evolution of human life
978 histories. *Proc. Natl. Acad. Sci. U. S. A.* **95,** 1336–9 (1998).

979 38. Navarrete, A., van Schaik, C. P. & Isler, K. Energetics and the evolution of
980 human brain size. *Nature* **480,** 91–93 (2011).

981 39. Bhui, R. & Gershman, S. J. Decision by sampling implements efficient coding of
982 psychoeconomic functions. *Psychol. Rev.* **125,** 985–1001 (2018).

983 40. Wang, X. J. Probabilistic decision making by slow reverberation in cortical
984 circuits. *Neuron* **36,** 955–968 (2002).

985 41. Ratcliff, R. & Rouder, J. N. Modeling Response Times for Two-Choice Decisions.
986 *Psychol. Sci.* **9,** 347–356 (1998).

987 42. McDonnell, M. D. & Ward, L. M. The benefits of noise in neural systems:
988 bridging theory and experiment. *Nat. Rev. Neurosci.* **12,** 415–426 (2011).

989 43. Buchanan, K. L., Grindstaff, J. L. & Pravosudov, V. V. Condition dependence,
990 developmental plasticity, and cognition: implications for ecology and evolution.
991 *Trends Ecol. Evol.* **28,** 290–296 (2013).

992 44. Bornstein, A. M., Khaw, M. W., Shohamy, D. & Daw, N. D. Reminders of past

993 choices bias decisions for reward in humans. *Nat. Commun.* **8,** 15958 (2017).

994 45. Saleem, A. B., Diamanti, E. M., Fournier, J., Harris, K. D. & Carandini, M.
995 Coherent encoding of subjective spatial position in visual cortex and
996 hippocampus. *Nature* **562,** 124–127 (2018).

997 46. Gluth, S., Sommer, T., Rieskamp, J. & Büchel, C. Effective Connectivity between
998 Hippocampus and Ventromedial Prefrontal Cortex Controls Preferential Choices
999 from Memory. *Neuron* **86,** 1078–1090 (2015).

1000 47. Qamar, A. T. *et al.* Trial-to-trial, uncertainty-based adjustment of decision
1001 boundaries in visual categorization. *Proc. Natl. Acad. Sci.* **110,** 20332–20337
1002 (2013).

1003 48. Gardner, J. L. Optimality and heuristics in perceptual neuroscience. *Nat.*
1004 *Neurosci.* 1 (2019). doi:10.1038/s41593-019-0340-4

1005 49. Stocker, A. A. & Simoncelli, E. P. Noise characteristics and prior expectations in
1006 human visual speed perception. *Nat. Neurosci.* **9,** 578–85 (2006).

1007 50. Zylberberg, A., Wolpert, D. M. & Shadlen, M. N. Counterfactual Reasoning
1008 Underlies the Learning of Priors in Decision Making. *Neuron* **0,** (2018).

1009 51. Brunton, B. W., Botvinick, M. M. & Brody, C. D. Rats and Humans Can Optimally
1010 Accumulate Evidence for Decision-Making. *Science (80-.).* **340,** 95–98 (2013).

1011 52. Izard, V. & Dehaene, S. Calibrating the mental number line. *Cognition* **106,**
1012 1221–1247 (2008).

1013 53. van den Berg, R., Lindskog, M., Poom, L. & Winman, A. Recent is more: A
1014 negative time-order effect in nonsymbolic numerical judgment. *J. Exp. Psychol.*
1015 *Hum. Percept. Perform.* **43,** 1084–1097 (2017).

1016 54. Vehtari, A., Gelman, A. & Gabry, J. Practical Bayesian model evaluation using
1017 leave-one-out cross-validation and WAIC. *Stat. Comput.* 1–20 (2016).

1018