

Rationalizing spatial exploration patterns of wild animals and humans through a temporal discounting framework

Vijay Mohan K. Namboodiri^{a,b,1}, Joshua M. Levy^{c,1}, Stefan Mihalas^d, David W. Sims^{e,f,g}, and Marshall G. Hussain Shuler^{c,2}

^aDepartment of Psychiatry, University of North Carolina, Chapel Hill, NC 27599; ^bNeuroscience Center, University of North Carolina, Chapel Hill, NC 27599; ^cDepartment of Neuroscience, Johns Hopkins University, Baltimore, MD 21205; ^dAllen Institute for Brain Science, Seattle, WA 98109; ^eMarine Biological Association of the United Kingdom, The Laboratory, Plymouth PL1 2PB, United Kingdom; ^fOcean and Earth Science, National Oceanography Centre Southampton, University of Southampton, Southampton SO14 3ZH, United Kingdom; and ^gCentre for Biological Sciences, University of Southampton, Southampton SO17 1BJ, United Kingdom

Edited by H. Eugene Stanley, Boston University, Boston, MA, and approved May 26, 2016 (received for review February 1, 2016)

Understanding the exploration patterns of foragers in the wild provides fundamental insight into animal behavior. Recent experimental evidence has demonstrated that path lengths (distances between consecutive turns) taken by foragers are well fitted by a power law distribution. Numerous theoretical contributions have posited that “Lévy random walks”—which can produce power law path length distributions—are optimal for memoryless agents searching a sparse reward landscape. It is unclear, however, whether such a strategy is efficient for cognitively complex agents, from wild animals to humans. Here, we developed a model to explain the emergence of apparent power law path length distributions in animals that can learn about their environments. In our model, the agent’s goal during search is to build an internal model of the distribution of rewards in space that takes into account the cost of time to reach distant locations (i.e., temporally discounting rewards). For an agent with such a goal, we find that an optimal model of exploration in fact produces hyperbolic path lengths, which are well approximated by power laws. We then provide support for our model by showing that humans in a laboratory spatial exploration task search space systematically and modify their search patterns under a cost of time. In addition, we find that path length distributions in a large dataset obtained from free-ranging marine vertebrates are well described by our hyperbolic model. Thus, we provide a general theoretical framework for understanding spatial exploration patterns of cognitively complex foragers.

Lévy walks | temporal discounting | optimal search | decision making | foraging theory

Lévy walks are a special kind of random walk whose path lengths form a power law distribution at their asymptotic limit (x): $p(x) \propto x^{-\mu}$; $1 < \mu < 3$; $x > x_{\min}$ (1–4). Numerous recent papers have demonstrated that foraging animals in the wild or under controlled conditions show path lengths consistent with power laws (5–11), which are proposed to arise from an underlying Lévy walk process. Theoretical models have demonstrated that such a process can be optimal for memoryless agents searching for randomly distributed rewards across space under certain conditions (1, 2, 12). Together, these findings have led to the Lévy flight foraging hypothesis, which states that such search patterns have arisen due to their evolutionary advantage (2, 3). However, because many animals, including humans, are cognitively complex and can learn from their environments, it is important to address whether such heavy-tailed path lengths are optimal even for cognitively complex agents (Fig. S1). The question of how memory influences foraging patterns has been approached in some contexts (13–16) but has not yet been sufficiently addressed (17–20).

Because power law path lengths have been observed in sparse and dynamic environments (e.g., open ocean), in which foragers rarely revisited previously rewarded locations (8, 10, 21, 22), it is

reasonable to assume, as foundational models have, that there is little advantage to learning about the reward distributions at any given spatial location. Hence, under this assumption, prior studies constrained the class of models studied to random searches in the absence of learning. However, despite the fact that remembering spatial locations in environments such as open oceans may not be advantageous, it is widely believed that many ecological parameters, including prey distributions, show high degrees of spatial autocorrelation (23, 24). Moreover, it has been found that these distributions can exist as hierarchies, wherein large, global spatial structures comprise smaller, more local structures, and that predators potentially learn these mean scales in the spatial distribution of prey (23, 24). Therefore, given the existence of patterns in the distribution of prey in relative space, it may be advantageous for predators to build representations, or models, of these patterns as opposed to performing purely random searches (Fig. 1A and B). In this paper, we show that to optimally learn the reward distribution across relative spatial scales in the service of future reward rate maximization, foragers would produce approximate power law path lengths, resembling Lévy walks.

How might foragers build a model of the relative spacing between food items? Consider the foraging behavior of albatrosses,

Significance

Understanding the movement patterns of wild animals is a fundamental question in ecology with implications for wildlife conservation. It has recently been hypothesized that random search patterns known as Lévy walks are optimal and underlie the observed power law movement patterns of wild foragers. However, as Lévy walk models assume that foragers do not learn from experience, they may not apply generally to cognitive animals. Here, we present a decision-theoretic framework wherein animals attempting to optimally learn from their environments will show near power law-distributed path lengths due to temporal discounting. We then provide experimental support for this framework in human exploration in a controlled laboratory setting and in animals foraging in the wild.

Author contributions: V.M.K.N., J.M.L., S.M., and M.G.H.S. designed research; V.M.K.N. and J.M.L. performed research; V.M.K.N. and J.M.L. analyzed data; V.M.K.N., J.M.L., D.W.S., and M.G.H.S. wrote the paper; S.M. contributed to mathematical aspects of theory; and D.W.S. contributed data.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

¹V.M.K.N. and J.M.L. contributed equally to this work.

²To whom correspondence should be addressed. Email: shuler@jhmi.edu.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1601664113/-DCSupplemental.

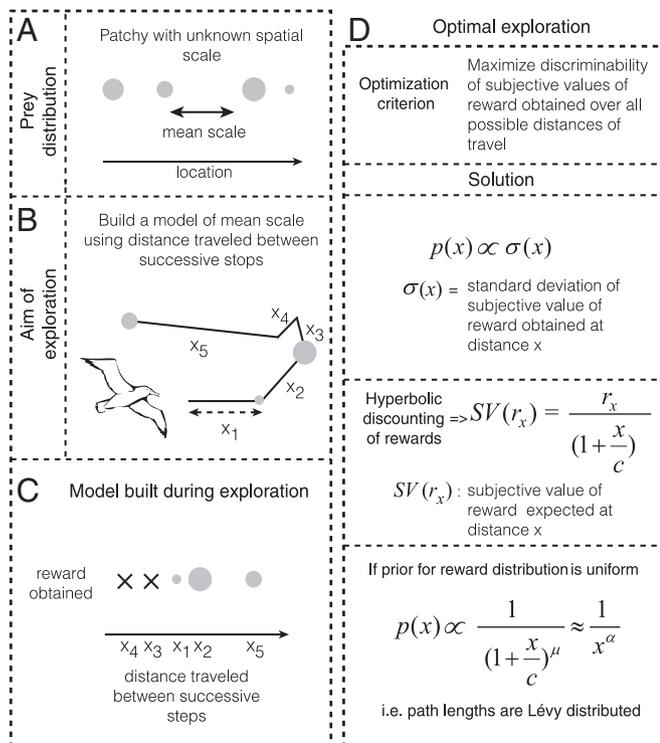


Fig. 1. Model for adaptive benefits of apparent Lévy walks. (A) The environments of foragers often show spatial autocorrelation with a mean spatial scale. (B and C) Thus, it is likely that foragers attempt to build a model of the mean spatial scale by building a model of rewards obtained for different flight distances (main text). (D) An optimal model of exploration that maximizes discriminability requires foragers to sample different flight distances (or durations) in proportion to the uncertainty in subjective value of rewards predicted at those distances. If the prior expectation of rewards is uniform, the sampling of different flight distances will produce a hyperbolic-like distribution (*SI Results, 2.1 Optimal Exploration of Reward Distributions Across Relative Space*)—due to hyperbolic discounting—that can appear to be power law distributed (Fig. 2).

for instance. A straightforward solution is to build a model of rewards obtained as a function of distance flown on each step (Fig. 1C). Because the speed of movement during searching is often nearly constant for many foragers (e.g., refs. 8 and 16), this model can also be built with respect to the time flown on each step. The question faced by the forager then becomes how to sample different step lengths to maximize the ability to detect differences in reward distributions associated with each step length. However, foragers do not treat the same reward available at different delays equally: The later the receipt of a reward is, the lower its subjective value (26, 27). In other words, the subjective value of a reward expected after a long flight is smaller than that of the same reward obtained immediately. Many behavioral experiments have shown that the cost of time associated with a delayed reward takes a specific functional form: that of a hyperbolic (for $\mu = 1$) or hyperbolic-like (for $\mu \neq 1$) function (26–33), shown as

$$SV(r, t) = rD(r, t) = \frac{r}{(1 + t/c)^\mu}, \quad [1]$$

where $SV(r, t)$ and $D(r, t)$ represent the subjective value and discounting function, respectively, of a reward of magnitude r delayed by a time t . c and μ represent parameters that measure the rate at which the value of a delayed reward is discounted. In light of such a cost of time, exploration of a given flight time should be done under consideration of its utility for future exploitation. Thus, the foragers must explore to maximize their

ability to discriminate subjective value (not reward) distributions associated with different step lengths.

Here, we show that to maximize the ability to discriminate the subjective value distributions associated with different step lengths, the forager has to sample each step length in proportion to the uncertainty in subjective value associated with that step length (*SI Results, 2.1 Optimal Exploration of Reward Distributions Across Relative Space* and Fig. 1D). This strategy makes intuitive sense because the higher the uncertainty associated with an option, the more it must be sampled to learn its properties. Such a strategy of exploring in proportion to uncertainty has previously been assumed to be an exploration heuristic (34). However, we show that it is in fact optimal for maximizing discriminability (*SI Results, 2.1 Optimal Exploration of Reward Distributions Across Relative Space*). For a forager that initiates exploration under a uniform prior (i.e., no a priori assumption regarding the distribution of rewards), sampling in proportion to uncertainty in subjective value equates to sampling in proportion to the discounting function associated with a flight time. As previously mentioned, the discounting function over flight time is hyperbolic. Therefore, for constant speed, the discounting function for a path length is also hyperbolic. Thus, we predict that the path length distribution of a forager attempting to explore the reward distribution across relative space will be

$$p(x) \propto \frac{1}{(1 + x/c)^\mu}; \quad \text{or rearranging by a constant, } p(x) \propto \frac{1}{(c + x)^\mu}. \quad [2]$$

This is similar in appearance to a power law, but differs due to the presence of an additive constant c . Consequently, whereas it decays asymptotically as a power law, it predicts a constant probability at low values. Further, note that any distribution that is consistent with a strict power law [e.g., prior observations of foraging patterns (8–10)] will necessarily be consistent with the above distribution because the power law distribution is a special case of Eq. 2 in which $c = 0$. Because foragers might limit the range of their exploration to a bounded interval of step lengths (e.g., due to behavioral/environmental constraints), the above probability distribution would be expected to hold only in a truncated domain (between x_{\min} and x_{\max}) under exploration. Further, in reward dense environments, we propose that observed path lengths would reflect not just the intended path lengths shown in Eq. 2, but also the truncation due to prey encounter, thus resulting in exponential path lengths [as has previously been shown (7)] (*SI Results, 2.4 Truncation Due to Prey Encounter* and Fig. S2).

The model described above makes several predictions about search behavior of a cognitively complex agent that we could test with humans in a controlled laboratory setting. Specifically, we sought to determine (i) whether humans search space in a random search pattern as expected from the Lévy walk model or in a systematic and deterministic way and (ii) whether spatial search patterns are sensitive to the cost of time incurred in traversing the space. To test this, we designed a spatial exploration task for humans with and without the cost of time (*Methods*, Fig. 2A, and *SI Discussion, 1.1 Human Exploratory Task*). In phase 1 (with a time cost) of this task, subjects could stop an image of a flying albatross to reveal a fish at a given spatial location. The goal of the subjects was to build a model of the distribution of fish as the knowledge acquired during this exploration phase could then be used on one exploitation trial to collect the largest fish possible. Crucially, flying longer distances across the screen took proportionally more time (the longest distance corresponded to waiting 10 s). Unknown to the subjects was that the distribution of fish sizes at any given location was uniform between fixed bounds. To test path lengths in the absence of a time cost, we removed the distance–time relationship in phase 2 and allowed the subjects to explore by merely clicking a given location with a computer mouse. In other words, they no longer had to wait for the albatross to fly to that location.

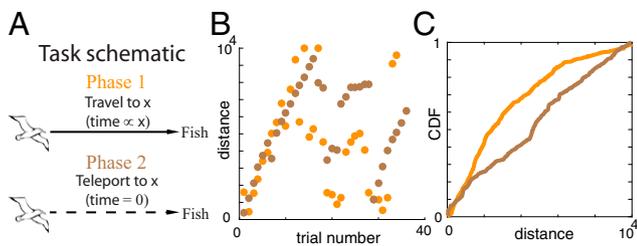


Fig. 2. Human spatial exploration task with and without temporal costs. (A) Schematic of the computer task (*Methods* and *SI Methods*). In phase 1, an albatross flies across the screen from a nest at a constant speed. In phase 2, subjects can make the albatross teleport (i.e., flight time is negligible). (B) Data from example subjects showing systematic search behavior across space. (C) Raw cumulative distribution function (CDF) of the population data across subjects for phase 1 and phase 2, showing sensitivity of exploration to the cost of time.

We found that in both phases, the pattern of search was nonrandom. Subjects systematically explored the space by, for instance, undertaking longer and longer path lengths or undertaking shorter and shorter path lengths from the end of the screen (see Fig. 2B for two example subjects and Fig. S3 for all subjects). Statistically, the probability of finding bouts of positive path length differences or negative path length differences between consecutive paths was higher than chance in every subject for phase 2 ($P < 0.05$, runs test with Benjamini–Hochberg correction for multiple comparisons, $n = 12$). For phase 1, the nonrandomness in the search was statistically significant in 10 of 12 subjects ($P < 0.05$, runs test with Benjamini–Hochberg correction for multiple comparisons). Thus, human spatial search in this random environment is not random. This conclusion is also bolstered by prior studies demonstrating that numerous animals remember spatial locations to produce nonrandom spatial search patterns in the wild (20, 35–41). We also found that the distribution of path lengths in phase 1 was significantly different from that in phase 2 (Fig. 2C; $P < 0.001$, two-tailed two-sample Kolmogorov–Smirnov test) due to the cost of time, as predicted by our temporal discounting model. Thus, human data support two key predictions of our model, i.e., that spatial search by cognitively complex agents is systematic and nonrandom and that temporal discounting plays a fundamental role in the shaping of such a search. These datasets are relatively small, however—as it is difficult to encourage human subjects to explore for long periods in a laboratory setting—and, therefore, are insufficient for model comparisons (although our model is consistent with the data). Therefore, to perform model comparisons, where considerable amounts of data are required (Fig. 3A), we turned to foraging data in the wild where, in some instances, thousands of path lengths have been recorded from individual animals.

Given the preponderance of evidence that foraging path lengths are well fitted by the power law distribution (5–9), the immediate question to be addressed is whether the hyperbolic distribution of path lengths expected from Eq. 2 can be well described by a power law. Because the above distribution (and a power law) is defined in a bounded domain, we tested against a truncated power law (see *Methods* for details). We found that for random numbers generated using Eq. 2, Akaike information criterion weights ($wAIC$) overwhelmingly support a truncated power law compared with a truncated exponential ($wAIC_{tp} = 1.000$ and $wAIC_{exp} = 0.000$) for all parameters tested (Fig. 3A and Fig. S4).

At this point, we wondered whether previously analyzed foraging data (8) may be well explained by our model. For this analysis, we compared a hyperbolic model to a power law model, as the power law distribution was found to provide a good fit to the data (8) and is generally compared against the exponential distribution to assert the presence of Lévy walks (5–7, 9–12). To

be clear, we compared against the power law distribution, not the family of distributions that have power law distributions at their asymptotic limits—to which a hyperbolic and a power law model both belong. One important point to note here is that whereas a strict power law distribution is a special case of the hyperbolic distribution, it is not necessarily true that in a direct comparison between the two models the data will be better fitted by a hyperbolic distribution. If the path length distributions decay as steeply or more steeply than a power law, the best fit hyperbolic model will reduce to the best fit power law and, hence, model selection (by $wAIC$) would favor the model with fewer parameters (i.e., the pure power law model).

Because differentiating between two highly similar distributions requires considerable statistical power, we limited our test to eight individual marine animals, comprising four blue sharks *Prionace glauca* (PG) and four basking sharks *Cetorhinus maximus* (CM), for which a substantial number of path lengths ($>10,000$) were recorded. The results for two individuals, blue sharks PG2 and PG4, are shown in Fig. 3B and C, respectively. In both cases, the hyperbolic fit (cyan) provides an excellent fit to the data. Notably, the truncated power law fit is visually compelling for PG2 (Fig. 3B) but not for PG4 (Fig. 3C). Indeed, individual PG2 represents a typical case where the fits are not easily distinguishable visually (as in the simulation in Fig. 3A), but where the $wAIC$ overwhelmingly favors a hyperbolic model. Examining all individuals, we found that the hyperbolic model provided a superior explanation of the data compared with power law and exponential models in all but one individual (Table 1). In this individual (CM3), the exponential model provided the best fit, potentially due to prey encounter-related truncation (7). In all other cases, the hyperbolic model was overwhelmingly favored

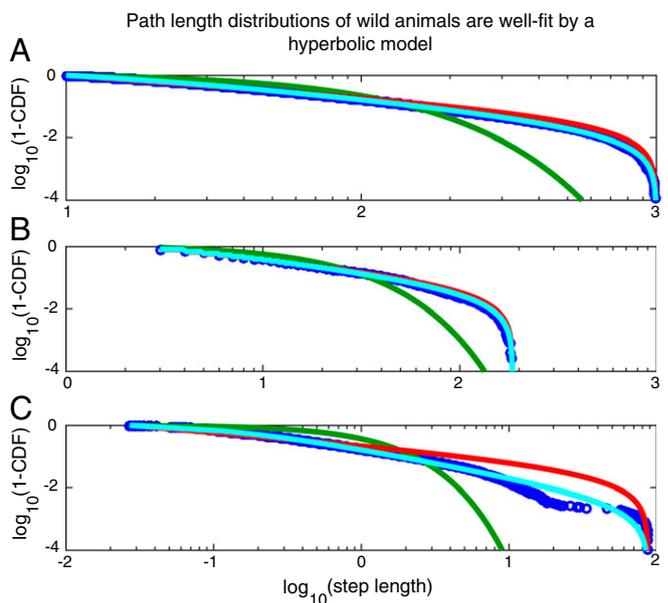


Fig. 3. Previously collected data from wild animals are better fitted by a hyperbolic model than by a power law model. (A) Random data generated from a hyperbolic distribution (blue circles) can be well approximated by a power law distribution (red), but not by an exponential distribution (green): 25,000 random numbers were generated from a hyperbolic distribution (Eq. 2) (*Methods*) with truncation set to be between 10 and 1,000 (see Fig. S4 for more parameters). The best fit truncated power law describes the data significantly better than the best fit exponential ($wAIC_{tp} = 1.000$ and $wAIC_{exp} = 0.000$). (B) Previously collected data (blue circles) that are well fitted by a hyperbolic model (cyan) and a power law model (red), but not by an exponential model (green) (see Fig. S5 for fits across many subjects). (C) Data from a subject in which the hyperbolic model is considerably preferred to any alternative model.

Table 1. Data from model fits to step length distributions of eight sharks

Individual	Data type	x_{\min}	x_{\max}	μ_{pl}	μ_{hyp}	c_{hyp}	KS D statistic			wAIC		
							pl	hyp	exp	pl	hyp	exp
CM1	Discrete	1	322	1.19	1.41	1.19	0.061	0.026	0.356	$<10^{-6}$	1.000	$<10^{-6}$
CM2	Discrete	6	91	2.07	3.03	7.20	0.056	0.083	0.127	$<10^{-6}$	1.000	$<10^{-6}$
CM3	Discrete	7	68	1.49	8	88.64	0.077	0.051	0.059	1×10^{-91}	1×10^{-11}	1.000
CM4	Discrete	7	63	1.44	1.79	4.25	0.015	0.007	0.061	2×10^{-4}	1.000	$<10^{-6}$
PG1	Discrete	2	185	2.10	2.24	0.36	0.013	0.003	0.248	$<10^{-6}$	1.000	$<10^{-6}$
PG2	Discrete	3	185	1.63	1.82	1.23	0.022	0.007	0.219	$<10^{-6}$	1.000	$<10^{-6}$
PG3	Discrete	4	218	1.70	1.77	0.57	0.025	0.018	0.222	0.292	0.708	$<10^{-6}$
PG4	Continuous	0.027	87.32	1.38	1.90	0.13	0.150	0.031	0.376	$<10^{-6}$	1.000	$<10^{-6}$

The first seven animals had quantized data that could be considered as resulting from a discrete probability distribution. For these animals, we divided the data by their common factor to get the discrete data (e.g., the first animal had unique observations 1, 2, 3, ..., 322). x_{\min} and x_{\max} represent the best fit truncation across all three distributions (Methods). The maximum step length and percentile of x_{\max} for each animal were 322 (100th percentile), 91 (100th percentile), 78 (92nd percentile), 80 (80th percentile), 965 (84th percentile), 910 (80th percentile), 958 (80th percentile), and 87 (100th percentile), respectively. The best fit parameters for the power law and hyperbolic distributions are shown. The goodness-of-fit was quantified by the KS D statistic and the relative quality of fit was quantified by wAIC.

(wAIC = 1.000), except in PG3 where support was not as clear-cut (wAIC = 0.708). Hence, our theoretical model provides a superior fit to previously collected foraging data.

In our model for foraging animals in the wild, we previously assumed perfect ability to estimate the time flown in a given step. Because we know that the error in estimating longer intervals is larger than that in estimating shorter ones (42), we derived the optimal exploration model for the biologically realistic case in which time perception is subjective and noisy (Fig. 4). In this model, the sampling per bin of path length (or equivalently, real time) for maximizing discriminability of rewards associated with that path length is determined by the degree of nonlinearity in time perception as different bins in subjective time are scaled differently, depending on the nonlinearity (Fig. 4C). Theoretically, it has been proposed previously that the degree of nonlinearity in time perception is directly related to the discounting function in subjective time (43) (Fig. 4D, Left). Consequently, we show—based on our prior theory of temporal perception and decision making (43–46)—that the optimal path length distribution would be

$$p(x) \propto \frac{1}{(vT_{\text{time}} + x)^3}. \quad [3]$$

Here, v is the speed of the animal. The term T_{time} is the interval over which the past reward rate experienced by an animal is estimated to make appropriate intertemporal decisions that maximize reward rates (43–45). Importantly, this term governs the nonlinearity of time perception and the steepness of temporal discounting (43–46) (SI Results, 2.1.5 Exploration under noisy temporal estimation). Thus, the power law that best approximates the above distribution would have an exponent determined by the nonlinearity of time perception (Fig. 4D; SI Results, 2.1.5 Exploration under noisy temporal estimation; and Fig. S6).

It is important to note that the derivations mentioned above necessarily simplify the foraging problem faced by animals in the wild. For instance, one factor that we did not yet take into account is that animals might account for other sources of risk such as that resulting from competition in their exploratory model. In such a case, we show in SI Results, 2.1.6 Modeling risk due to competition that the probability distribution of path length durations can be calculated as

$$p(t) \propto \frac{1}{(T_{\text{time}} + t)^3 (1 + k\alpha r^{-\alpha t})^{1/\alpha}}, \quad [4]$$

where k and α represent the magnitude of competition such that an increase in their values represents more competition and

hence, shorter path lengths. r in Eq. 4 can be thought of as the mean reward expected in an environment; the larger the mean reward expected is, the larger the competition and the shorter the path lengths. The asymptotic limit of Eq. 4, for positive α , will have a power law exponent greater than 3 and hence will be outside the Lévy range of exponents. However, in cases where the asymptotic limit cannot be reached, as is often the case in biology where path lengths are often truncated either by the physical world or potentially by some internal limit set by the forager, a best fit truncated power law will appear to have a lower exponent than the real generative process, with the apparent exponent lying between 0 and $3 + 1/\alpha$.

An even more complete model of animal movements would involve additional factors, some of which are mentioned in SI Discussion, 1.2 Potential Predicted Deviations from the Simplified

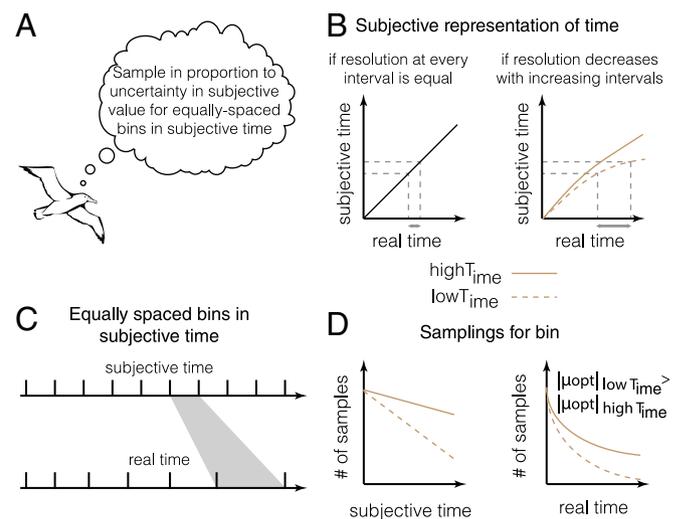


Fig. 4. Optimal exploration when temporal representations are noisy. (A) Optimal algorithm for exploration. (B) If temporal resolution is constant at every interval, subjective representation of time can be represented as a linear function of real time, with constant noise. However, it is known that errors in timing increase with the interval being timed (42). In this case, subjective representation of time can be represented as a nonlinear function with the nonlinearity controlled by the parameter T_{time} (43–45). (C) When subjective representation of time is nonlinear (concave), equal bins in subjective time correspond to bins of increasing width in real time. (D) A theory of reward rate maximization (43–45) predicts linear sampling for optimal exploration in subjective time, with the slope determined by T_{time} . In real time, this sampling becomes hyperbolic (Eq. 3) with its decay controlled by T_{time} (SI Results, 2.1.5 Exploration under noisy temporal estimation).

Model Presented Here. Nevertheless, the simplified model presented here demonstrates that the path lengths of foragers with spatial memory who build a map of their environment for future exploitation can be heavy tailed and nearly power law distributed.

To conclude, we argue that if foragers seek to learn about their reward landscape, an optimal model of exploration would require sampling in proportion to the uncertainty in subjective value associated with a reward obtained after a given path length. Because it is often observed that the subjective value of rewards obtained after a delay is discounted hyperbolically with respect to the delay, we showed that the resultant path lengths would be hyperbolically distributed. In support of our model, we found that humans engaged in a laboratory exploration task searched space systematically and account for the cost of time in traversing space. Next, we showed that data generated from a hyperbolic distribution are better fitted by a power law distribution than by an exponential distribution and that previously collected data from foraging animals in the wild can be better explained by a hyperbolic model than by a power law model. Additionally, we extended our model to show that for foragers in the wild with noisy temporal perception, the exponent of the best fit power law is governed by the nonlinearity of time perception and the amount of competition faced from other foragers. Thus, we contribute to the ongoing discussion regarding the mechanistic origins of power law path lengths in foragers (47–58) by arguing that search patterns are unlikely to be purely random when cognitive modeling is advantageous and that approximate power law path lengths emerge due to the temporal discounting of farther away rewards. A deeper understanding of the neurobiological basis of spatial search (59–63) may further enrich this model and provide greater insight into the movements of wild animals and humans.

Methods

Description of Experiment.

Subjects. The 12 subjects that participated in this experiment were healthy individuals aged 22–35 y recruited from Johns Hopkins University. All procedures were approved by the Institutional Review Board at Johns Hopkins School of Medicine under application identification NA_00075036. All subjects gave oral consent before the start of the experiment.

Exploration task. We developed an exploration task for human subjects. Our task was divided into two phases. At the start of each trial in phase 1, an image of an albatross would begin flying (i.e., moving) from its “nest” left to right across the “sky” (i.e., light blue patch of screen). Subjects were instructed to stop the albatross at any point by clicking a mouse. When the albatross was stopped, it would instantaneously dive into the “water” (i.e., dark blue patch of screen) and an image of a fish would be revealed. Unknown to the subject, the size of the fish was drawn from a uniform distribution with five discrete outcomes. After the fish was displayed for 1 s, the albatross returned to its nest and immediately began to fly in the next trial.

In phase 1, it would take the albatross 10 s to fly the entire length of the screen. The speed of the albatross was constant and was 109 pixels per second. If the subject waited for the albatross to traverse the entire screen, the albatross automatically dove into the water and a fish was revealed. Before the start of phase 1, subjects were informed that they would have exactly 3 min to explore the region and “discover where the biggest fish swim.” They were also informed that at the end of the phase, they would be given just one chance to catch the largest fish they could and that their payout would be “determined exclusively by the size of the fish on this one trial” and not by the fish caught during exploration. In this way, the subjects were incentivized to explore the region.

Following phase 1, phase 2 began. In phase 2, the subjects were instructed that the albatross was flying over a different region of the ocean and, thus, they had to explore again to know where the biggest fish swim. The albatross remained in its nest until the subject clicked on a region of space (indicated by a gray region that ran the length of the screen) to which it instantaneously teleported. Therefore, there was no time cost in exploring farther regions of space, as there was in phase 1. The instructions for phase 2 were similar to those for phase 1 except that subjects were informed that they had 1 min to explore the region. This limit was imposed so that subjects would complete a similar number of trials in phase 1 and phase 2 (because trials in phase 2

are shorter as the albatross teleports rather than flies). Aesthetic changes (background color, fish image, and fish size) were made between phases to encourage exploration by underscoring the instruction that the environments in phase 1 and 2 were distinct.

Procedure. Subjects were placed in a quiet room in front of a 13-inch MacBook Pro. On-screen instructions were read aloud by the experimenter to ensure that the subject understood them. At the end of the experiment, subjects answered a questionnaire administered by the experimenter and were monetarily compensated for their participation.

Display. The experiment was controlled by custom-made code written in Java (JDK 6.0.65). The display was 1,220 pixels wide and 730 pixels high. The area of the fish image (i.e., the size of the fish) was a random integer value between 1 and 5 scaled by a constant factor. *Movie S1* shows a sample of the experiment.

Animal Foraging Data. Blue ($n = 4$) and basking sharks ($n = 4$) were each fitted with a pressure-sensitive data logger that recorded an individual time series of depth measurements as the fish swam through the water column, as described previously (8). Raw depth measurements from loggers were converted into move step lengths by calculating the vertical movement step size between successive vertical changes in direction (from down to up and vice versa), as described previously (8).

Procedure to Fit Data. The general approach used here to test the appropriateness of different models is to (i) estimate the respective parameters using maximum-likelihood estimation (MLE) for the same set of possible truncations across all models, (ii) set the best truncation to that resulting in the lowest Kolmogorov–Smirnov (KS) D statistic across all models and all truncations, and (iii) quantify relative likelihoods of models, using the AIC.

The truncated hyperbolic-like distribution, shown in Eq. 2, needed to be statistically characterized. We do so in *SI Results, 2.2 Statistical Characterization of the Hyperbolic Distribution in Eq. S1*. The truncated power law distribution is a particular example of a truncated hyperbolic distribution for $c = 0$. To test whether data generated from this distribution can be mistaken for a power law distribution (Fig. 3A), we generated data using the following random number generator (derived in *SI Results, 2.2 Statistical Characterization of the Hyperbolic Distribution in Eq. S1*),

$$t = \left[(t_{\min} + c)^{1-\mu} + \left\{ (t_{\max} + c)^{1-\mu} - (t_{\min} + c)^{1-\mu} \right\} u \right]^{1/(1-\mu)} - c,$$

where t is the random variate following the truncated hyperbolic-like distribution, t_{\min} and t_{\max} are the minimum and maximum truncation limits, c and μ are the parameters of the distribution, and u is a uniform random variate. For the purpose of Fig. 2, t_{\min} was set at 10 and t_{\max} at 1,000. More parameters are shown in Fig. S4. The procedure for fitting and testing of power law, hyperbolic, and exponential models is explained below.

Prior data were tested against three models: (i) exponential, (ii) truncated power law, and (iii) truncated hyperbolic. The same truncation parameters were fitted across all models. To this end, we fitted each model at different sets of truncation and then found the truncation that resulted in the lowest KS D statistic across all models. Thus, truncation limits are not free-fitting parameters for each distribution that add cost to the AIC. The benefit of using this approach is that the exact same data are used for comparison across all of the models, thus avoiding different domains of the probability distribution functions for the different distributions tested. The procedure for MLE is derived and explained in detail in *SI Methods, 3.1 Model Fits*.

To calculate relative likelihood of the different models (i.e., assessing which model minimizes the estimated Kullback–Liebler divergence between data and the model), the AIC was calculated for every model as shown below (explained in *SI Methods, 3.2 Model Comparisons*):

$$AIC_{tp} = 2 - 2 \log(\mathcal{L}_{tp}) + \frac{4}{n-2}$$

$$AIC_{hyp} = 4 - 2 \log(\mathcal{L}_{hyp}) + \frac{12}{n-3}$$

$$AIC_{exp} = 2 - 2 \log(\mathcal{L}_{exp}) + \frac{4}{n-2}$$

\mathcal{L} is the likelihood of the data given each model. We used the small-sample correction for the AIC.

ACKNOWLEDGMENTS. We thank Prof. David Foster and Prof. Ernst Niebur for comments and discussion. This work was funded by a National Institute of Mental Health Grant R01MH093665.

1. Reynolds AM, Rhodes CJ (2009) The Lévy flight paradigm: Random search patterns and mechanisms. *Ecology* 90(4):877–887.
2. Viswanathan GM, Raposo EP, da Luz MGE (2008) Lévy flights and superdiffusion in the context of biological encounters and random searches. *Phys Life Rev* 5(3):133–150.
3. Viswanathan GM, da Luz MGE, Raposo EP, Stanley HE (2011) *The Physics of Foraging* (Cambridge Univ Press, Cambridge, UK).
4. Mendez V, Campos D, Bartumeus F (2014) *Stochastic Foundations in Movement Ecology: Anomalous Diffusion, Front Propagation and Random Searches* (Springer, Berlin).
5. Sims DW, et al. (2014) Hierarchical random walks in trace fossils and the origin of optimal search behavior. *Proc Natl Acad Sci USA* 111(30):11073–11078.
6. de Jager M, Weissing FJ, Herman PMJ, Nolet BA, van de Koppel J (2011) Lévy walks evolve through interaction between movement and environmental complexity. *Science* 332(6037):1551–1553.
7. De Jager M, et al. (2014) How superdiffusion gets arrested: Ecological encounters explain shift from Lévy to Brownian movement. *Proc Biol Sci* 281:20132605.
8. Humphries NE, et al. (2010) Environmental context explains Lévy and Brownian movement patterns of marine predators. *Nature* 465(7301):1066–1069.
9. Sims DW, et al. (2008) Scaling laws of marine predator search behaviour. *Nature* 451(7182):1098–1102.
10. Humphries NE, Weimerskirch H, Queiroz N, Southall EJ, Sims DW (2012) Foraging success of biological Lévy flights recorded in situ. *Proc Natl Acad Sci USA* 109(19):7169–7174.
11. Raichlen DA, et al. (2014) Evidence of Lévy walk foraging patterns in human hunter-gatherers. *Proc Natl Acad Sci USA* 111(2):728–733.
12. Viswanathan GM, et al. (1999) Optimizing the success of random searches. *Nature* 401(6756):911–914.
13. Ferreira AS, Raposo EP, Viswanathan GM, Da Luz MGE (2012) The influence of the environment on Lévy random search efficiency: Fractality and memory effects. *Phys A Stat Mech Appl* 391(11):3234–3246.
14. Van Moorter B, et al. (2009) Memory keeps you at home: A mechanistic model for home range emergence. *Oikos* 118(5):641–652.
15. Tabone M, Ermentrout B, Doiron B (2010) Balancing organization and flexibility in foraging dynamics. *J Theor Biol* 266(3):391–400.
16. Boyer D, Walsh PD (2010) Modelling the mobility of living organisms in heterogeneous landscapes: Does memory improve foraging success? *Philos Trans A Math Phys Eng Sci* 368(1933):5645–5659.
17. James A, Plank MJ, Edwards AM (2011) Assessing Lévy walks as models of animal foraging. *J R Soc Interface* 8(62):1233–1247.
18. Gaustad AO, Mysterud A (2013) The Lévy flight foraging hypothesis: Forgetting about memory may lead to false verification of Brownian motion. *Mov Ecol* 1(1):9.
19. Gaustad AO (2011) Memory matters: Influence from a cognitive map on animal space use. *J Theor Biol* 287(0):26–36.
20. Fagan WF, et al. (2013) Spatial memory and animal movement. *Ecol Lett* 16(10):1316–1329.
21. Weimerskirch H, Pinaud D, Pawlowski F, Bost C-A (2007) Does prey capture induce area-restricted search? A fine-scale study using GPS in a marine predator, the wandering albatross. *Am Nat* 170(5):734–743.
22. Weimerskirch H, Gault A, Cherel Y (2005) Prey distribution and patchiness: Factors in foraging success and efficiency of wandering albatrosses. *Ecology* 86(10):2611–2622.
23. Fauchald P, Erikstad KE, Skarsfjord H (2000) Scale-dependent predator-prey interactions: The hierarchical spatial distribution of seabirds and prey. *Ecology* 81(3):773–783.
24. Fauchald P, Erikstad KE (2002) Scale-dependent predator-prey interactions: The aggregative response of seabirds to prey under variable prey abundance and patchiness. *Mar Ecol Prog Ser* 231:279–291.
25. Sims DW, Quayle VA (1998) Selective foraging behaviour of basking sharks on zooplankton in a small-scale front. *Nature* 393(June):460–464.
26. Green L, Myerson J (2004) A discounting framework for choice with delayed and probabilistic rewards. *Psychol Bull* 130(5):769–792.
27. Stephens DW, Krebs JR (1986) *Foraging Theory* (Princeton Univ Press, Princeton, NJ).
28. Kirby KN, Marakovic NN (1995) Modeling myopic decisions: Evidence for hyperbolic delay-discounting within subjects and amounts. *Organ Behav Hum Decis Process* 64(1):22–30.
29. Green L, Myerson J (1996) Exponential versus hyperbolic discounting of delayed outcomes: Risk and waiting time. *Am Zool* 36(4):496–505.
30. Vuchinich RE, Simpson CA (1998) Hyperbolic temporal discounting in social drinkers and problem drinkers. *Exp Clin Psychopharmacol* 6(3):292–305.
31. Johnson MW, Bickel WK (2002) Within-subject comparison of real and hypothetical money rewards in delay discounting. *J Exp Anal Behav* 77(2):129–146.
32. Murphy JG, Vuchinich RE, Simpson CA (2001) Delayed reward and cost discounting. *Psychol Rec* 51:571–588.
33. Simpson CA, Vuchinich RE (2000) Reliability of a measure of temporal discounting. *Psychol Rec* 50:3–16.
34. Badre D, Doll BB, Long NM, Frank MJ (2012) Rostrolateral prefrontal cortex and individual differences in uncertainty-driven exploration. *Neuron* 73(3):595–607.
35. Muller M, Wehner R (1994) The hidden spiral: Systematic search and path integration in desert ants, *Cataglyphis fortis*. *J Comp Physiol A Neuroethol Sens Neural Behav Physiol* 175:525–530.
36. Cramer AE, Gallistel CR (1997) Vervet monkeys as travelling salesmen. *Nature* 387(6632):464.
37. Janson CH (1998) Experimental evidence for spatial memory in foraging wild capuchin monkeys, *Cebus apella*. *Anim Behav* 55(5):1229–1243.
38. Valero A, Byrne RW (2007) Spider monkey ranging patterns in Mexican subtropical forest: Do travel routes reflect planning? *Anim Cogn* 10(3):305–315.
39. Asensio N, Brockelman WY, Malaiivijitnon S, Reichard UH (2011) Gibbon travel paths are goal oriented. *Anim Cogn* 14(3):395–405.
40. Taffe MA, Taffe WJ (2011) Rhesus monkeys employ a procedural strategy to reduce working memory load in a self-ordered spatial search task. *Brain Res* 1413:43–50.
41. Dyer F (1996) Spatial memory and navigation by honeybees on the scale of the foraging range. *J Exp Biol* 199(Pt 1):147–154.
42. Buhusi CV, Meck WH (2005) What makes us tick? Functional and neural mechanisms of interval timing. *Nat Rev Neurosci* 6(10):755–765.
43. Namboodiri VMK, Mihalas S, Marton TM, Hussain Shuler MG (2014) A general theory of intertemporal decision-making and the perception of time. *Front Behav Neurosci* 8(February):61.
44. Namboodiri VMK, Mihalas S, Hussain Shuler MG (2014) A temporal basis for Weber's law in value perception. *Front Integr Neurosci* 8(October):79.
45. Namboodiri VMK, Mihalas S, Hussain Shuler MG (2014) Rationalizing decision-making: Understanding the cost and perception of time. *Timing Time Percept Rev* 1(4):1–40.
46. Namboodiri VMK, Mihalas S, Hussain Shuler MG (2016) Analytical calculation of errors in time and value perception due to a subjective time accumulator: A mechanistic model and the generation of Weber's law. *Neural Comput* 28(1):89–117.
47. Reynolds A (2015) Liberating Lévy walk research from the shackles of optimal foraging. *Phys Life Rev* 14:59–83.
48. Bartumeus F (2015) Behavioural ecology cannot turn its back on Lévy walk research: Comment on "Liberating Lévy walk research from the shackles of optimal foraging" by A.M. Reynolds. *Phys Life Rev* 14:84–86.
49. Boyer D (2015) What future for Lévy walks in animal movement research?: Comment on "Liberating Lévy walk research from the shackles of optimal foraging", by A.M. Reynolds. *Phys Life Rev* 14:87–89.
50. Cheng K (2015) Answer (in part) blowing in the wind: Comment on "Liberating Lévy walk research from the shackles of optimal foraging" by A. Reynolds. *Phys Life Rev* 14:90–93.
51. da Luz MGE, Raposo EP, Viswanathan GM (2015) And yet it optimizes: Comment on "Liberating Lévy walk research from the shackles of optimal foraging" by A.M. Reynolds. *Phys Life Rev* 14:94–98.
52. Focardi S (2015) Do the albatross Lévy flights below the spandrels of St Mark?: Comment on "Liberating Lévy walk research from the shackles of optimal foraging" by A.M. Reynolds. *Phys Life Rev* 14:99–101.
53. Humphries NE (2015) Why Lévy Foraging does not need to be 'unshackled' from Optimal Foraging Theory: Comment on "Liberating Lévy walk research from the shackles of optimal foraging" by A.M. Reynolds. *Phys Life Rev* 14:102–104.
54. MacIntosh AJ (2015) At the edge of chaos-error tolerance and the maintenance of Lévy statistics in animal movement: Comment on "Liberating Lévy walk research from the shackles of optimal foraging" by A.M. Reynolds. *Phys Life Rev* 14:105–107.
55. Miramontes O (2015) Divorcing physics from biology? Optimal foraging and Lévy flights: Comment on "Liberating Lévy walk research from the shackles of optimal foraging" by A.M. Reynolds. *Phys Life Rev* 14:108–110.
56. Sims DW (2015) Intrinsic Lévy behaviour in organisms—searching for a mechanism: Comment on "Liberating Lévy walk research from the shackles of optimal foraging" by A.M. Reynolds. *Phys Life Rev* 14:111–114.
57. Reynolds A (2015) Venturing beyond the Lévy flight foraging hypothesis: Reply to comments on "Liberating Lévy walk research from the shackles of optimal foraging". *Phys Life Rev* 14:115–119.
58. Bartumeus F, Raposo EP, Viswanathan GM, da Luz MGE (2014) Stochastic optimal foraging: Tuning intensive and extensive dynamics in random searches. *PLoS One* 9(9):e106373.
59. O'Keefe J, Nadel L (1978) *The Hippocampus as a Cognitive Map* (Oxford Univ Press, Oxford).
60. Hafting T, Fyhn M, Molden S, Moser MB, Moser EI (2005) Microstructure of a spatial map in the entorhinal cortex. *Nature* 436(7052):801–806.
61. Gutiérrez ED, Cabrera JL (2015) A neural coding scheme reproducing foraging trajectories. *Sci Rep* 5(18009):18009.
62. Yoshida W, Ishii S (2006) Resolution of uncertainty in prefrontal cortex. *Neuron* 50(5):781–789.
63. Kable JW, Glimcher PW (2007) The neural correlates of subjective value during intertemporal choice. *Nat Neurosci* 10(12):1625–1633.
64. Welch BL (1947) The generalization of student's problem when several different population variances are involved. *Biometrika* 34(1-2):28–35.
65. Frank MJ, Doll BB, Oas-Terpstra J, Moreno F (2009) Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nat Neurosci* 12(8):1062–1068.
66. White EP, Enquist BJ, Green JL (2008) On estimating the exponent of power-law frequency distributions. *Ecology* 89(4):905–912.

Supporting Information

Namboodiri et al. 10.1073/pnas.1601664113

SI Discussion

1.1) Human Exploratory Task. The human exploratory task presented here has superficial differences from the general model of exploration of relative space that was proposed earlier for foragers in the wild. The main difference is that whereas foragers in the wild were predicted to explore relative locations in space (i.e., locations relative to their current position), humans in this task are exploring absolute locations in space (i.e., locations relative to a fixed location). However, there is a fundamental equivalence between both cases in that exploration is performed with respect to the distance traveled on each search bout. In this sense, exploration of absolute space in one dimension is just a special case of exploration in relative space. Consequently, the discounting of rewards for future exploitation is defined with respect to the distance traveled during exploratory bouts in both cases. Thus, our laboratory task provides a test for the conceptual framework we developed here.

1.2) Potential Predicted Deviations from the Simplified Model Presented Here. We presented a model for animal movements in the wild by assuming that animals are building a model of the subjective value of reward distributions across relative space. Our calculations would thus be expected to be true only when animals are in fact building such a model. Because animals would likely not spend their entire foraging time building such a model, in these cases, one must expect deviations from the hyperbolic distribution of path lengths predicted here. One such instance would be when animals are exploiting knowledge gained from the aforementioned exploration. If foragers realized that there is indeed a scale over which prey are distributed across space, they would fly these distances under exploitation. Thus, for exploitation, one would predict unimodal distributions of path lengths if the foragers are flying an optimal distance. Another prediction of our framework would be that paths of foragers might show directional preferences. If there is anisotropic autocorrelation in the spatial distribution of prey, the foragers would likely also learn the optimal angle to travel. Thus, one would predict nonuniform spread of directions of travel.

Further, even when foragers are building a model of subjective value of reward distributions across relative space, our calculations have ignored complications such as models of risk associated, for instance, with competition from other foragers. These complications would introduce quantitative deviations from the simplified framework presented here (*SI Results, 2.1.6) Modeling risk due to competition*). On a related note, it must also be pointed out that optimizing discriminability between all choices is not required to optimize the ability to pick the maximum reward in a static environment. However, because reward environments are rarely static, it is likely that animals evolved a mechanism to build models of the world appropriate for mapping the entire distribution of subjective values. Another caveat is in relation to the dimensionality of the environment to explore. Here, the assumption was that the dependence of interest to the animal is on the distance from the previous reward (i.e., autocorrelation). This results in a 1D exploration problem irrespective of the dimensionality of the environment. Further, we have assumed radial symmetry in autocorrelation. Despite these caveats, our calculations illustrate how path lengths of foragers in the wild can be heavy tailed, as experimentally observed.

SI Results

2.1) Optimal Exploration of Reward Distributions Across Relative Space. As explained in the main text, we assume that animals are mapping out a distribution of subjective reward values across relative space

(i.e., distances traveled). To this end, we assume that they maintain a constant speed and are mapping out the distribution of rewards against the time flown on each step length (Fig. 1). The aim for such an exploration is to sample each option a given number of times to build a map of the reward distribution across relative space for future exploitation. Because it is known that the subjective value of a delayed reward decays systematically with respect to the delay (e.g., “Which do you prefer: \$100 now or \$100 in a year?”), exploration of relative space should also consider the subjective value of a reward for a given flight time. In other words, the exploration of a given flight time must be done under consideration of its utility for future exploitation. We also assume that the search space for exploration is bounded by the forager to be between a minimum (t_{\min}) and a maximum flight time (t_{\max}). Note that we are using symbols for time here but they can easily be converted to distance as time = distance/speed. This conversion obviously applies to the case of constant speed, but also to the case where speed is variable, but independent of distance flown.

If there were no discounts (subjective value of a delayed reward divided by the subjective value of that reward when obtained immediately) associated with time and all flight times are expected to have the same reward distribution, the optimal manner to map out the reward-flight time relationship would be to sample all possible flight times equally and note the corresponding reward amounts. Such a sampling strategy would lead to a uniform distribution of sample times between t_{\min} and t_{\max} .

However, when different time intervals have different associated discounts, the problem of optimal sampling becomes nontrivial. We first solve this problem for the simpler case of two possible flight time options (similar to the “two-armed bandit” problem), under the following assumptions:

- i) As in our behavioral task, the environment is stationary such that the reward distribution associated with any option does not change in time.
- ii) The aim of exploration is to ascertain the optimal option for future exploitation.
- iii) The total number of trials to explore is fixed. The aim of the agent is to calculate how to sample the two different options while keeping the total number of trials fixed. This assumption will be relaxed later to also consider (a) the case where the total time for exploration is fixed and (b) a case where the stopping of exploration has to be determined by the forager.
- iv) The agent has access to the SEM of an option after n samplings.

We first consider the case where the total number of trials for exploration is fixed.

2.1.1) Total number of trials for exploration fixed. With the above assumptions, the problem faced by the agent is exactly the same as the problem faced in designing an optimal experiment to maximize one’s ability to distinguish between the mean of two distributions. The solution for the optimal experiment design is to sample each option in such a way that the t statistic between the two distributions can be maximized. Hence, for the agent trying to solve the optimal exploration problem, the task is to maximize the t statistic between the subjective value distributions of the two options. Because the variances of the subjective values for the options will not be the same due to the multiplicative discounts, one has to use Welch’s t statistic for unequal variances (64), which is defined as

$$t = \frac{\langle x_1 \rangle - \langle x_2 \rangle}{\sigma_d},$$

where x_i is the subjective value for option “ i ,” $\langle x \rangle$ denotes the expected value of x , and σ_d refers to the SE of the difference of means. Because the number of samplings of either option affects only σ_d , maximizing the t statistic is the same as minimizing σ_d .

If the distributions of rewards for each option are independent and identical, with a SD σ , σ_d —the SE of the difference of subjective means of the two options—would be

$$\sigma_d = \sigma \sqrt{\frac{d_1^2}{n_1} + \frac{d_2^2}{n_2}},$$

where d_1 and d_2 are the discounts associated with the two options, and n_1 and n_2 are the number of times the two options were sampled. Given that the total number of trials is constant (N), $n_2 = N - n_1$.

Minimizing σ_d is equivalent to minimizing the square of σ_d . At the minimum of the square of σ_d , the derivative of σ_d^2 with respect to n_1 will be zero. Therefore,

$$\begin{aligned} \frac{d(\sigma_d^2)}{dn_1} &= \sigma^2 \left(-\frac{d_1^2}{n_1^2} + \frac{d_2^2}{(N - n_1)^2} \right) = 0 \\ \Rightarrow \frac{n_1}{d_1} &= \frac{n_2}{d_2}. \end{aligned}$$

The second derivative can be verified to be positive at this solution, showing that σ_d is indeed at its minimum for the above solution.

Hence, for a binary choice between options with independent and identically distributed (i.i.d.) reward distributions but with different discounts, optimal exploration requires sampling in proportion to the discounts.

Next, we extend this analysis to the case of k options (“ k -armed bandits”), each having i.i.d. reward distributions with different discounts, d_k . Although it is hard to identify a single metric whose optimality can identify the option with the maximum subjective value, one can define the optimality metric as the sum of variances of the difference distributions for each distinct pair of options. Minimizing this metric will lead to maximum discriminability between all options. Hence, for the k -option case, we assume that the aim of the agent is to maximize the ability to distinguish between the subjective rewards of the k options. With this assumption, the agent has to minimize

$$\sigma_{\text{net}}^2 = \sum_{i < j} \sigma_d^{ij2},$$

where σ_d^{ij} is the SE of the difference of the expected subjective rewards of options i and j .

Expanding the expression on the right-hand side (RHS) gives

$$\begin{aligned} \sigma_{\text{net}}^2 &= \sigma^2 \left(\frac{d_1^2}{n_1} + \frac{d_2^2}{n_2} + \frac{d_1^2}{n_1} + \frac{d_3^2}{n_3} + \frac{d_1^2}{n_1} + \frac{d_4^2}{n_4} + \dots + \frac{d_1^2}{n_1} + \frac{d_k^2}{n_k} + \frac{d_2^2}{n_2} + \frac{d_3^2}{n_3} \right. \\ &\quad \left. + \frac{d_2^2}{n_2} + \frac{d_4^2}{n_4} + \dots + \frac{d_2^2}{n_2} + \frac{d_k^2}{n_k} + \dots \right). \end{aligned}$$

Collecting like terms,

$$\sigma_{\text{net}}^2 = \sigma^2 (k - 1) \left(\sum_i \frac{d_i^2}{n_i} \right).$$

Because the total number of trials is constant, n_k can be written as $n_k = N - \sum_{i \neq k} n_i$.

At the minimum of σ_{net}^2 , its partial derivative with respect to any $n_{i,i < k}$ will be zero. Therefore,

$$\begin{aligned} \frac{\partial \sigma_{\text{net}}^2}{\partial n_i} &= \sigma^2 \left(-\frac{d_i^2}{n_i^2} + \frac{d_k^2}{(N - \sum_{j < k} n_j)^2} \right) = 0 \\ \Rightarrow \frac{n_i}{d_i} &= \frac{n_k}{d_k}, \quad i < k. \end{aligned}$$

Hence, even in the k -option case between options with i.i.d. reward distributions but with different discounts, optimal exploration requires sampling in proportion to the discounts.

It is important to point out that the optimal exploration requires sampling in proportion to the discounts only when comparing between options with i.i.d. reward distributions. The general solution as we worked out for the case with different variances associated with the real values of each option is to sample in proportion to the estimated variance of the mean of the subjective values associated with that option ($n \alpha d^2 \sigma^2 / n$). It is interesting to point out that a similar strategy, in which exploration of an option was proportional to uncertainty, was assumed for some previous studies of exploratory behavior in humans (34, 65).

Extending this to the continuous case, if we denote the probability of sampling the flight time t by $p(t)$ and substitute a hyperbolic or a hyperbolic-like (26) discounting function for $D(t)$, we get that for optimal exploration, assuming that flight time is proportional to distance traveled,

$$p(t) \propto \frac{1}{(c + t)^\mu}; \quad t_{\min} < t < t_{\max}, \quad \text{[S1]}$$

where c and μ represent the two constants in the hyperbolic-like function. In other words, an optimal agent samples flight times in proportion to its hyperbolic-like discounting function. Eq. S1 has been derived using the assumption that the bins in flight time (or relative spatial location) are linearly spaced. This in turn results from the assumption that the error in perception of time for each flight time is constant. This is an inaccurate assumption as it is known that errors in the representation of longer temporal intervals are larger. For such a case of noisy temporal representation, see *SI Results, 2.1.5) Exploration under noisy temporal estimation*.

2.1.2) Total time for exploration fixed. Now, we solve for optimal exploration when the total time for exploration is fixed. As before, the aim is to minimize σ_d^2 but under the constraint that $n_1 t_1 + n_2 t_2 = T$, where T is the total time available for exploration. Taking the derivative with respect to n_1 , we get

$$\frac{d(\sigma_d^2)}{dn_1} = \sigma^2 \left(-\frac{d_1^2}{n_1} + \frac{d_2^2}{((T - n_1 t_1)/t_2)} \right).$$

Thus,

$$\frac{d(\sigma_d^2)}{dn_1} = \sigma^2 \left(-\frac{d_1^2}{n_1^2} + \frac{t_1 t_2 d_2^2}{(T - n_1 t_1)^2} \right).$$

Thus, the minimum of σ_d^2 is when

$$\sigma^2 \left(-\frac{d_1^2}{n_1^2} + \frac{t_1 t_2 d_2^2}{(n_2 t_2)^2} \right) = 0;$$

i.e.,

$$\frac{n_1}{n_2} = \frac{d_1}{d_2} \sqrt{\frac{t_2}{t_1}}.$$

For the case with k options, we have to minimize

$$\sigma_{\text{net}}^2 = \sigma^2(k-1) \left(\sum_i \frac{d_i^2}{n_i} \right)$$

under the constraint $n_k = T - (n_1 t_1 + \dots + n_{k-1} t_{k-1})$. Taking the partial derivative with respect to any $n_{i < k}$, we get

$$\frac{\partial \sigma_{\text{net}}^2}{\partial n_i} = \frac{\partial}{\partial n_i} \sigma^2 \left(\frac{d_i^2}{n_i} + \frac{t_k d_k^2}{(T - \sum_{j < k} n_j t_j)} \right).$$

Setting the derivative to zero, we get

$$\sigma^2 \left(-\frac{d_i^2}{n_i^2} + \frac{t_k d_k^2}{(n_k t_k)^2} \right) = 0.$$

Thus,

$$\frac{n_i}{n_k} = \frac{d_i}{d_k} \sqrt{\frac{t_k}{t_i}}, \quad i < k.$$

Extending this to the continuous case, we get that when the total time for exploration is held constant, the probability of sampling a particular flight time must be

$$p(t) \propto \frac{1}{\sqrt{t}} \frac{1}{(c+t)^{\mu}}, \quad t_{\min} < t < t_{\max}. \quad [\text{S2}]$$

2.1.3) Caveat associated with the above model for optimality. It is important to mention the caveats associated with the above model for optimality. It is defined as an extension of the optimal experimental design concept wherein the only aim of the forager is to maximize the ability to discriminate between the means of the subjective values for different flight times. That is, the forager is seeking to purely explore the environment without exploiting its knowledge. The optimality is thus defined by this constraint. The second caveat associated with the above model is that it is only a steady-state solution for optimality. In other words, it says only that at the steady state, the probability of sampling a given flight time is as expressed by Eqs. S1 and S2. It does not provide a dynamical solution and, hence, does not predict how the sampling will develop with experience. For a discussion of this point, see the Movie S1 legend. These equations also assume that every flight time is equally likely to contain the largest fish, i.e., that the prior is uniform.

2.1.4) Self-initiated stopping rule for exploration. With the above caveats in mind, we can now attempt to define a stopping rule for exploration. Such a rule would be useful when the exploration phase is not well defined, as we assumed previously with fixed total samplings or fixed total time. A reasonable stopping rule can be defined as the moment when the net discriminability ($1/\sigma_{\text{net}}^2$) goes above a threshold. This threshold might represent the maximum resolution available to the agent. In other words, sampling beyond this point affords no benefit to the agent in terms of new information about subjective value. A caveat of the above statement is that in some cases, there might be no need to exploit even after this threshold has been met. For instance, the total time available for exploration to a forager could be much more than that needed to meet the maximum discriminability threshold. In this case, further exploration can be aimed at gaining more information about the rewards themselves. Under this framework, if there is infinite time to explore (i.e., no opportunity to exploit any knowledge gained), the forager will start sampling uniformly within the bounded search region.

Mathematically, we can quantify the abovementioned stopping rule of maximum discriminability as the sampling when $\sigma_{\text{net}}^2 = \sigma_{\text{stop}}^2$. For this calculation, we assume that the forager is sampling each option in proportion to its uncertainty (or discount, assuming a uniform prior). To calculate the sampling at the stoppage point, let us represent $n_i^{\text{stop}} = \beta d_i$. β can now be calculated as the value that satisfies

$$\sigma_{\text{net}}^2 = \sigma^2(k-1) \left(\sum_i \frac{d_i^2}{n_i^{\text{stop}}} \right) = \sigma_{\text{stop}}^2.$$

Substituting $n_i^{\text{stop}} = \beta d_i$, we get

$$\sigma^2(k-1) \left(\sum_i \frac{d_i^2}{\beta d_i} \right) = \sigma_{\text{stop}}^2.$$

Therefore,

$$\sigma^2(k-1) \left(\frac{\sum_i d_i}{\beta} \right) = \sigma_{\text{stop}}^2.$$

Or,

$$\beta = \frac{\sigma^2}{\sigma_{\text{stop}}^2} (k-1) \left(\sum_i d_i \right).$$

Thus,

$$n_i^{\text{stop}} = \frac{\sigma^2}{\sigma_{\text{stop}}^2} (k-1) d_i \left(\sum_i d_i \right). \quad [\text{S3}]$$

As mentioned previously, if there is no possibility to exploit even after the above threshold has been met, the sampling will then be directed toward gathering more information about the rewards themselves. In this case, the steady-state sampling will be uniform. **2.1.5) Exploration under noisy temporal estimation.** In the prior sections, we have assumed that errors in the representation of time are constant for every flight time. This is, however, not true. Hence, the sampling of different flight times will be done linearly with respect to the subjective representation of those intervals (Fig. 4), rather than their real time values (as was previously assumed). In a previous theory on intertemporal decision making and time perception (43, 44), we showed that a decision-making algorithm that considers reward rate maximization over a limited temporal interval (including a past interval over which reward rate is estimated as well as the expected interval to future reward) explains well-established observations in intertemporal decision making and time perception. In our theory, the subjective value of a delayed reward was calculated as

$$SV(r, t) = \frac{r - a_{\text{est}} t}{1 + t/T_{\text{ime}}}, \quad [\text{S4}]$$

where $SV(r, t)$ is the subjective value of a delayed reward of magnitude r and delay t , and a_{est} is the experienced reward rate estimated over the duration T_{ime} (referred to as “past integration interval”). The above equation holds when the average reward rate estimated in the past is assumed to not be available during the delay to the current reward, resulting in an opportunity cost of the delay. If such an opportunity cost is not present or accounted for, the numerator in Eq. S4 will not contain the $a_{\text{est}} t$ term. We also showed that the subjective representation of the delay t can be represented by

$$ST(t) = \frac{t}{1+t/T_{ime}}. \quad [S5]$$

The above equation represents approximate scalar timing [error in timing grows in proportion to interval being timed (42)] and thus takes into account noisy temporal estimation. It says that the representation of time is nonlinear and that this nonlinearity is controlled by the past integration interval, T_{ime} .

The subjective value of a delayed reward can be expressed in terms of the subjective representation of the delay as

$$SV(r, t) = r \left(1 - \frac{ST(t)}{T_{ime}} \right) - a_{est} ST(t).$$

Thus, the discounting of a delayed reward is linear with respect to the subjective representation of the delay. It also has two components: One is an explicit cost of time (the first term on the RHS) and the second one is the opportunity cost of time, expressed in terms of the subjective representation of the delay (second term on the RHS).

For exploration, the aim is to sample intervals, keeping in mind the explicit cost of time. Thus, the subjective value considered during exploration will not include the opportunity cost term. Hence, for exploration,

$$SV(r, t) = r \left(1 - \frac{ST(t)}{T_{ime}} \right).$$

Thus, the discounting function in terms of subjective representation of time is expressed as

$$D(ST(t)) = \frac{SV(r, t)}{r} = \left(1 - \frac{ST(t)}{T_{ime}} \right).$$

As was mentioned previously, the sampling under exploration is done in subjective time and not real time. Thus, the sampling probability of a flight time t can be expressed (similar to Eq. S1) as

$$p(ST(t)) \propto 1 - \frac{ST(t)}{T_{ime}}. \quad [S6]$$

The corresponding sampling in real time can then be calculated as

$$p(t) = p(ST(t)) \frac{dST(t)}{dt}.$$

From Eq. S5, $\frac{dST(t)}{dt} = \frac{1}{(1+t/T_{ime})^2}$ and from Eq. S6, $p(ST(t)) \propto \frac{1}{(1+t/T_{ime})}$. Thus,

$$p(t) \propto \frac{1}{(1+t/T_{ime})^3}. \quad [S7]$$

Or equivalently,

$$p(t) \propto \frac{1}{(T_{ime} + t)^3}. \quad [S8]$$

This equation is similar in form to Eq. S1 but has a fixed power of 3. When $T_{ime} \rightarrow 0$, Eq. S8 predicts exploratory sampling with a power law with exponent = 3. When $T_{ime} \rightarrow \infty$, the RHS will be dominated by a constant and thus, the sampling will be uniform. Uniform sampling can also be approximated by a power law of exponent = 0. Thus, Eq. S8 predicts that optimal exploratory sampling of foragers in the wild will depend on their

past integration interval—a quantity that measures future tolerance to delay in decision making and the nonlinearity of time perception.

For a given value of T_{ime} , the closest power law fit to Eq. S8 can be calculated by equating the medians of the two distributions. This calculation is worked out below.

From Eq. S19, the median for Eq. S8 can be expressed as

$$\text{median} = \left[(x_{\min} + T_{ime})^{-2} + 0.5 \left\{ (x_{\max} + T_{ime})^{-2} - (x_{\min} + T_{ime})^{-2} \right\} \right]^{-1/2} - T_{ime}.$$

From Eq. S18, the median for the closest truncated power law can be expressed as

$$\text{median}(tp) = \left[x_{\min}^{1-\mu_{opt}} + 0.5 \left(x_{\max}^{1-\mu_{opt}} - x_{\min}^{1-\mu_{opt}} \right) \right]^{1/(1-\mu_{opt})}.$$

Equating the two medians, the exponent (μ_{opt}) for the closest power law to Eq. S8 is the solution to

$$\begin{aligned} & \left[x_{\min}^{1-\mu_{opt}} + 0.5 \left(x_{\max}^{1-\mu_{opt}} - x_{\min}^{1-\mu_{opt}} \right) \right]^{1/(1-\mu_{opt})} \\ & = \left[(x_{\min} + T_{ime})^{-2} + 0.5 \left\{ (x_{\max} + T_{ime})^{-2} - (x_{\min} + T_{ime})^{-2} \right\} \right]^{-1/2} - T_{ime}. \end{aligned} \quad [S9]$$

This can be calculated only numerically and depends on the value of T_{ime} . When $T_{ime} = 0$, it is easy enough to see that $\mu_{opt} = 3$. When $T_{ime} \rightarrow \infty$, the RHS of Eq. S9 tends to the limit $0.5(x_{\max} + x_{\min})$. Hence, the solution for $\mu_{opt} = 0$ in this case. Thus, in all cases, the exponent of the best fit power law to Eq. S8 will be between 0 and 3. Numerical solutions for different values of T_{ime} , x_{\min} , and x_{\max} are shown in Fig. S6.

An important caveat needs to be mentioned regarding Eq. S8. Its derivation assumes that the discounting function is calculated without any associated model of risk such as those resulting from competition due to other predators. These factors are quite likely important in determining the success of foragers in the wild and hence would likely be included in their decision making. However, to preserve simplicity, we have chosen to ignore such factors. Simple models of such risk can be found in the supplement of our prior theoretical work (43). When such factors are included in the discounting function, the resultant path length distribution would be much more complicated. Further, as mentioned in *SI Results, 2.1.3) Caveat associated with the above model for optimality*, the above model assumes a uniform prior. Hence, real life path lengths will certainly be more complicated than the simple model presented here. Nevertheless, our model shows that the resultant path lengths will be heavy tailed.

2.1.6) Modeling risk due to competition. In *SI Results, 2.1.5) Exploration under noisy temporal estimation*, we considered the simple case where animals in the wild have to account only for the passage of time in calculating the temporal discounting. Specifically, we assumed that they do not face explicit risks of losing rewards or, at least, that they do not model such risks. However, this assumption is almost definitely incorrect. In the presence of such competition, during the course of a foraging path, the value of a reward might reduce because other animals might consume it. Considering such a risk as a stochastic process with a mean decay proportional to the magnitude of the reward (i.e., larger rewards are more sought after), we showed previously that an appropriate model of risk can be mathematically expressed as

$$r(\text{after path of duration } t) = \frac{r(\text{start of path})}{(1 + kar(\text{start of path})^\alpha t)^{1/\alpha}}.$$

Here, $r(\text{start of path})$ is the reward magnitude at the start of a given path of duration t and k and α represent the degree of competition—the larger their values are, the more the competition. In the time t , the mean reward is expected to have decayed to the value $r(\text{after path of duration } t)$ as expressed above. As is clear, this introduces another power law form to the model expressed in *SI Results, 2.1.5) Exploration under noisy temporal estimation*. Thus, a more complete model of path length distribution can be obtained by combining temporal uncertainty (shown in Eq. S8) and competition risk (shown above) as

$$p(t) \propto \frac{1}{(T_{ime} + t)^3 (1 + kar^\alpha t)^{1/\alpha}}. \quad \text{[S10]}$$

Here, r can be thought of as the mean reward expected in an environment.

2.2) Statistical Characterization of the Hyperbolic Distribution in Eq. S1. Rewriting Eq. S1 with a proportionality constant k , we get

$$p(t) = k(c + t)^{-\mu}; \quad t_{\min} < t < t_{\max}.$$

For this expression to represent a probability distribution, k must have a value such that it is normalized over the domain. We first consider the case of continuous data for which the integral of the probability distribution should be 1; i.e.,

$$\int_{t_{\min}}^{t_{\max}} k(c + t)^{-\mu} dt = 1.$$

Solving for k , we get

$$k = \frac{1 - \mu}{(t_{\max} + c)^{1-\mu} - (t_{\min} + c)^{1-\mu}}.$$

Thus,

$$p(t) = \frac{1 - \mu}{(t_{\max} + c)^{1-\mu} - (t_{\min} + c)^{1-\mu}} (t + c)^{-\mu}. \quad \text{[S11]}$$

The cumulative distribution function can be calculated as

$$F(t) = \int_{t_{\min}}^t p(t) dt = \int_{t_{\min}}^t \frac{1 - \mu}{(t_{\max} + c)^{1-\mu} - (t_{\min} + c)^{1-\mu}} (t + c)^{-\mu} dt.$$

Solving and simplifying, we get

$$F(t) = \frac{(t + c)^{1-\mu} - (t_{\min} + c)^{1-\mu}}{(t_{\max} + c)^{1-\mu} - (t_{\min} + c)^{1-\mu}}. \quad \text{[S12]}$$

For testing whether observed data are consistent with the probability distribution shown in Eq. S11, it is important to be able to generate random numbers following that distribution. This can be done using inverse transform sampling, provided one has access to a uniform random variate u . Inverse transform sampling states that solving for t in the equation below will require that t is distributed according to the distribution shown in Eq. S11:

$$F(t) = u;$$

i.e.,

$$\frac{(t + c)^{1-\mu} - (t_{\min} + c)^{1-\mu}}{(t_{\max} + c)^{1-\mu} - (t_{\min} + c)^{1-\mu}} = u.$$

Solving for t , we get

$$t = \left[(t_{\min} + c)^{1-\mu} + \left\{ (t_{\max} + c)^{1-\mu} - (t_{\min} + c)^{1-\mu} \right\} u \right]^{1/(1-\mu)} - c. \quad \text{[S13]}$$

Eq. S13 describes the random number generator for the distribution shown in Eq. S11 with u being a uniform random variate between 0 and 1.

The next question to be addressed is that of parameter estimation for Eq. S11 when attempting to fit experimental data. This can be done using maximum-likelihood estimation as shown below.

Say that we have n independent observations that are assumed to be from the hyperbolic distribution in Eq. S11. Let the i th observation be t_i . Then the likelihood of the data given parameters c and μ is

$$\mathcal{L} = \prod_i p(t_i).$$

The log-likelihood is

$$\log(\mathcal{L}) = \sum_i \log(p(t_i)).$$

Substituting $p(t_i)$ from Eq. S11,

$$\log(\mathcal{L}) = \sum_i \log \left(\frac{1 - \mu}{(t_{\max} + c)^{1-\mu} - (t_{\min} + c)^{1-\mu}} (t_i + c)^{-\mu} \right).$$

Simplifying,

$$\log(\mathcal{L}) = \sum_i \left[\log(1 - \mu) - \mu \log(t_i + c) - \log \left((t_{\max} + c)^{1-\mu} - (t_{\min} + c)^{1-\mu} \right) \right].$$

Performing the sums, we get

$$\log(\mathcal{L}) = n \log(1 - \mu) - \mu \sum_i \log(t_i + c) - n \log \left((t_{\max} + c)^{1-\mu} - (t_{\min} + c)^{1-\mu} \right).$$

At the maximum of the log-likelihood function, its partial derivatives with respect to c and μ will be zero.

Therefore,

$$\frac{\partial \log(\mathcal{L})}{\partial \mu} = \frac{\partial}{\partial \mu} \left[n \log(1 - \mu) - \mu \sum_i \log(t_i + c) - n \log \left((t_{\max} + c)^{1-\mu} - (t_{\min} + c)^{1-\mu} \right) \right] = 0.$$

Thus, we get

$$\frac{1}{1 - \mu} + \langle \log(x_i + c) \rangle = \frac{(x_{\max} + c)^{1-\mu} \log(x_{\max} + c) - (x_{\min} + c)^{1-\mu} \log(x_{\min} + c)}{(x_{\max} + c)^{1-\mu} - (x_{\min} + c)^{1-\mu}}, \quad \text{[S14]}$$

where $\langle \rangle$ signifies the mean.

Similarly,

$$\begin{aligned} \frac{\partial \log(\mathcal{L})}{\partial c} &= \frac{\partial}{\partial c} \left[n \log(1 - \mu) - \mu \sum_i \log(t_i + c) \right. \\ &\quad \left. - n \log \left((t_{\max} + c)^{1-\mu} - (t_{\min} + c)^{1-\mu} \right) \right] = 0 \\ -\mu \langle (x_i + c)^{-1} \rangle &= (1 - \mu) \frac{(x_{\max} + c)^{-\mu} - (x_{\min} + c)^{-\mu}}{(x_{\max} + c)^{1-\mu} - (x_{\min} + c)^{1-\mu}}. \end{aligned} \quad [\text{S15}]$$

Eqs. **S14** and **S15** have to be numerically solved simultaneously to calculate the maximum-likelihood estimates for parameters c and μ .

In practice, this numerical estimation has to be multisteped because typical numerical solvers provide only local solutions. To ensure that the initial values for numerical solution are close to the global solution, we used the following procedure. First, we calculated the pure power law fit to data using the version of Eq. **S14** with $c = 0$. This is also the ML estimator for a pure power law as calculated previously (66) and is shown below:

$$\frac{1}{1 - \mu} + \langle \log(x_i) \rangle = \frac{(x_{\max})^{1-\mu} \log(x_{\max}) - (x_{\min})^{1-\mu} \log(x_{\min})}{(x_{\max})^{1-\mu} - (x_{\min})^{1-\mu}}. \quad [\text{S16}]$$

Eq. **S16** also needs to be estimated numerically. The initial value for this solution was taken as the ML exponent for a nontruncated power law that has an analytical expression shown below (66). The expression can be obtained as the limit $x_{\max} \rightarrow \infty$ in Eq. **S16**:

$$\mu = 1 - \frac{1}{\log(x_{\min}) - \langle \log(x_i) \rangle}. \quad [\text{S17}]$$

Once the solution to Eq. **S16** was obtained, the numerical solutions for Eqs. **S14** and **S15** were calculated using the procedure explained below.

- i) Call μ_{mle} for Eq. **S16** as μ_{tp} because this is the MLE exponent for a pure truncated power law model. The initial values of μ for the numerical solution for Eqs. **S14** and **S15** were taken as $[\mu_{\text{tp}}, \mu_{\text{tp}} + 0.05, \mu_{\text{tp}} + 0.10, \dots, \mu_{\text{tp}} + 2]$.
- ii) The corresponding initial value of c for each of the above-mentioned μ was calculated as the value that would produce the same median for the truncated hyperbolic distribution as the median for the best fit truncated power law with exponent μ_{tp} . This is calculated as shown below:

$$\text{median}(\text{tp}) = \left[x_{\min}^{1-\mu_{\text{mle}}} + 0.5 \left(x_{\max}^{1-\mu_{\text{mle}}} - x_{\min}^{1-\mu_{\text{mle}}} \right) \right]^{1/(1-\mu_{\text{mle}})} \quad [\text{S18}]$$

$$\begin{aligned} \text{median}(\text{hyp}) &= \left[(x_{\min} + c)^{1-\mu} + 0.5 \left\{ (x_{\max} + c)^{1-\mu} \right. \right. \\ &\quad \left. \left. - (x_{\min} + c)^{1-\mu} \right\} \right]^{1/(1-\mu)} - c. \end{aligned} \quad [\text{S19}]$$

Equating the above two values and solving for c for each value of μ from the list $[\mu_{\text{tp}}, \mu_{\text{tp}} + 0.05, \mu_{\text{tp}} + 0.10, \dots, \mu_{\text{tp}} + 2]$ provides the appropriate initial point for the numerical solution for Eqs. **S14** and **S15**. However, the above equations need to be solved numerically as well. The initial value for this solution was set sequentially. For $\mu = \mu_{\text{tp}}$, the initial value for c was taken as zero. The solution to this equation provided the initial value for $\mu = \mu_{\text{tp}} + 0.05$. The solution for this equation provided the initial value for $\mu = \mu_{\text{tp}} + 0.1$ and so on.

- iii) For each of the above combinations of c and μ as initial values, the corresponding log-likelihood of the data was calculated.
- iv) The maximum-likelihood c and μ for solving Eqs. **S14** and **S15** were taken as the pair that maximized the global (against initial values for numerical optimization) log-likelihood calculated above.

To appropriately compare the exponential model to truncated hyperbolic and power law models, it is important to use a truncated exponential model. If k is the normalization constant, the exponential distribution can be defined as

$$p(t) = ke^{-\lambda t}; \quad t_{\min} \leq t \leq t_{\max}$$

with

$$\int_{t_{\min}}^{t_{\max}} p(t) dt = 1.$$

Solving for k , we get

$$k = \frac{\lambda e^{\lambda t_{\min}}}{1 - e^{-\lambda(t_{\max} - t_{\min})}}.$$

Thus, the truncated exponential model used for calculating likelihoods is

$$p(t) = \frac{\lambda e^{-\lambda(t - t_{\min})}}{1 - e^{-\lambda(t_{\max} - t_{\min})}}; \quad t_{\min} \leq t \leq t_{\max}. \quad [\text{S20}]$$

It is Eq. **S20** that must be used to calculate \mathcal{L}_{exp} in Eq. **S30**.

The CDF over the truncated domain for the exponential distribution can be calculated as

$$F(t) = \int_{t_{\min}}^t p(t) dt = \int_{t_{\min}}^t \frac{\lambda e^{-\lambda t}}{1 - e^{-\lambda(t_{\max} - t_{\min})}}; \quad t_{\min} \leq t \leq t_{\max}.$$

Solving and simplifying, we get

$$F(t) = \frac{1 - e^{-\lambda(t - t_{\min})}}{1 - e^{-\lambda(t_{\max} - t_{\min})}}. \quad [\text{S21}]$$

A random number generator for this distribution can again be obtained using inverse transform sampling by solving for t in the equation below,

$$F(t) = u,$$

where u is a uniform random variate.

Solving for t , we get

$$t = t_{\min} - \frac{1}{\lambda} \log \left(1 - u \left\{ 1 - e^{-\lambda(t_{\max} - t_{\min})} \right\} \right). \quad [\text{S22}]$$

To get a maximum-likelihood estimate of λ , the log-likelihood of the data can be expressed as

$$\log(\mathcal{L}) = \sum_i \log \left(\frac{\lambda e^{-\lambda(t_i - t_{\min})}}{1 - e^{-\lambda(t_{\max} - t_{\min})}} \right). \quad [\text{S23}]$$

We found the MLE λ by numerically maximizing the above equation.

2.3) Discrete Distributions. Because previously collected data (Table 1) also contained discrete data, we derive the above procedure for discrete data here. For truncated discrete data, the only difference from the above derivation is that the sum of the probability distribution

should equal 1, rather than the integral. For the truncated hyperbolic model, the appropriate probability distribution function can be written as

$$p(t) = \frac{(c+t)^{-\mu}}{\sum_{t_{\min}}^{t_{\max}} (c+t)^{-\mu}}; \quad t_{\min} \leq t \leq t_{\max}.$$

The ML estimates of c and μ were calculated by numerically maximizing the log-likelihood of the data:

$$\log(\cdot) = \langle -\mu \log(c+t_i) \rangle - \log \left(\sum_{t_{\min}}^{t_{\max}} (c+t)^{-\mu} \right).$$

We used the same procedure as described in the above section for global maximization. The same equation above was used to estimate the best fit power law distribution by setting $c = 0$.

For a discrete truncated exponential distribution, the probability distribution function can be written as

$$p(t) = (1 - e^{-\lambda}) \frac{e^{-\lambda(t-t_{\min})}}{1 - e^{-\lambda(t_{\max}-t_{\min}+1)}}. \quad [\text{S24}]$$

The ML estimate for λ was calculated by maximizing the log-likelihood. It can be shown that this is equivalent to numerically solving the following equation:

$$\frac{e^{-\lambda}}{(1 - e^{-\lambda})} - (t_{\max} - t_{\min} + 1) \frac{e^{-\lambda(t_{\max}-t_{\min}+1)}}{1 - e^{-\lambda(t_{\max}-t_{\min}+1)}} = \langle t_i - t_{\min} \rangle. \quad [\text{S25}]$$

2.4) Truncation Due to Prey Encounter. A curious observation regarding forager path lengths is that they tend to become diffusive when food abundance is high (8, 10). Whereas it was previously argued that Brownian walks are sufficiently productive under high food density (8, 10), a recent experimental finding demonstrated that such path lengths result from truncation of search due to prey encounter (7). A simple theoretical argument provided an explanation for this result. However, their calculation was in one dimension. Here, we extend their argument to search in 2D (Fig. S2).

We first calculate the probability distribution of truncated path lengths. An easy solution is to assume that foragers are moving in a straight line until they hit a food item and then calculate the resultant path length distribution. We assume that food is homogeneously distributed with density (number per unit area) ρ . Denote the diameter of the prey (or a cluster of prey) as d_{prey} , the CDF of path lengths as $F(r)$, path lengths as r , and probability density function of path lengths as $p(r)$. Then the infinitesimal change in the CDF, $dF(r)$ over a distance dr from r corresponds to the probability that the path length lies between r and $r + dr$. This is equal to the probability that the forager at least moved r ($= 1 - F(r)$) multiplied by the probability that the forager hit a target between r and $r + dr$. The second probability can be calculated as the total angle covered by prey in the ring between r and $r + dr$ divided by 2π . The total number of foragers (on an average) in this ring equals $2\pi r dr \rho$. Thus, we can write

$$dF(r) = (1 - F(r)) \cdot 2\pi r dr \rho \cdot \frac{d_{\text{prey}}/r}{2\pi}$$

If the forager had a sensory diameter of d_{perc} , the above equation would remain exactly the same but with the change that d_{prey} would change to $d_{\text{prey}} + d_{\text{perc}}$.

Solving the above equation, we get

$$F(r) = 1 - e^{-\rho d_{\text{prey}} r}. \quad [\text{S26}]$$

Thus, the probability density function of path lengths is given by

$$p(r) = \frac{dF(r)}{dr} = \rho d_{\text{prey}} e^{-\rho d_{\text{prey}} r}.$$

Hence, the prey-truncated path lengths are exponentially distributed.

If the intended path length distribution (CDF) for exploration was $F_{\text{intended}}(r)$, the probability that the path length is at least greater than r is the probability that the intended path length is at least greater than r multiplied by the probability that there was no prey truncation within r . The latter probability, from Eq. S26, is $1 - F(r)$. Thus, the observed path lengths would be

$$1 - F_{\text{observed}}(r) = (1 - F_{\text{intended}}(r)) e^{-\rho d_{\text{prey}} r}. \quad [\text{S27}]$$

Thus, as the density of prey increases, path lengths approach the exponential distribution.

SI Methods

3.1) Model Fits. We fitted the data using three models: exponential, truncated power law, and truncated hyperbolic. As discussed in prior work (1, 2), upper truncation is important for power law and hyperbolic models, but not for exponential models, as such heavy-tailed models cannot last for an infinite domain in the real world. Additionally, a lower truncation is necessary for the power law model as it is not defined for a flight time of zero. Consequently, as done previously (1, 2), we used a truncation for all of the models tested. Unlike before, however, we did not tune the truncation parameters to each model. Rather, we tuned them across all models by picking the truncation that resulted in the lowest KS D statistic across all models and all truncations. The different values of the truncation tested were as follows: For discrete data, the low truncation possibilities were set to 1, 2, ..., 7 and the upper truncation to 80th, 84th, ..., 100th percentiles of the unique observations in the data, whereas for continuous data, the possible options for the minimum truncation were set to 0th, 4rth, ..., 20th percentiles and the upper truncation was chosen from 80th, 84th, ..., 100th percentiles for which at least 80% of the data were retained. Thus, if the lower truncation was set to the 20th percentile, the upper truncation had to be the 100th percentile, whereas the possible upper truncations with a lower truncation of 16th percentile were 96th and 100th percentiles.

The ML estimate for the exponent in the truncated power law model was numerically calculated by solving Eq. S16. Similarly, the ML estimate for c and μ for the truncated hyperbolic model was numerically calculated by solving Eqs. S14 and S15. Finally, the ML estimate for the truncated exponential model was calculated by solving Eqs. S23 and S25.

3.2) Model Comparisons. We used the AIC to compare the three models. Because the truncation parameters were set to be the same for all models, these were not counted as free-fit parameters in the calculation of the AIC. Thus, the AICs for the different models were calculated as shown below (each AIC was calculated using the correction for small sample sizes; i.e., they were AICc). The numbers of free-fit parameters for the different models were one for “tp,” one for “exp,” and, two for “hyp”:

$$\text{AIC}_{\text{tp}} = 2 - 2 \log(\mathcal{L}_{\text{tp}}) + \frac{4}{n-2} \quad [\text{S28}]$$

$$\text{AIC}_{\text{hyp}} = 4 - 2 \log(\mathcal{L}_{\text{hyp}}) + \frac{12}{n-3} \quad [\text{S29}]$$

$$\text{AIC}_{\text{exp}} = 2 - 2 \log(\mathcal{L}_{\text{exp}}) + \frac{4}{n-2}. \quad [\text{S30}]$$

Aim: optimally search for randomly distributed rewards in space

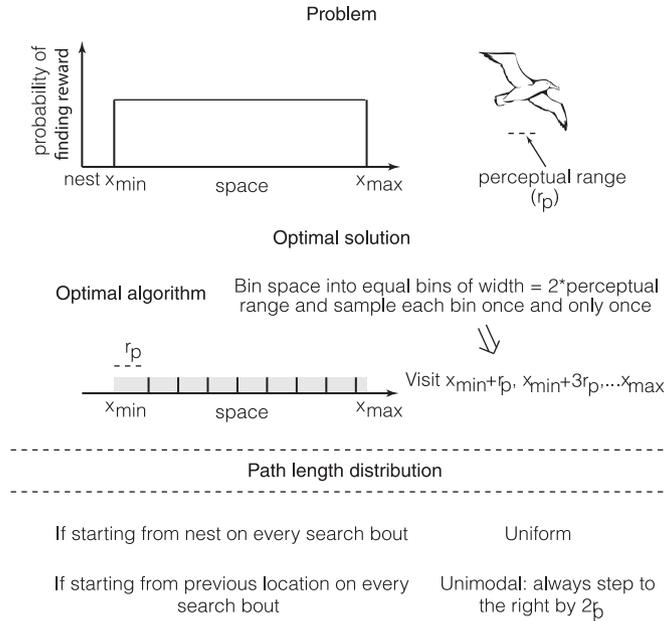


Fig. S1. Optimal search of a bounded area containing reward by a forager with spatial memory. A forager with spatial memory searches a bounded area containing uniformly distributed rewards. The optimal solution, assuming that the forager has a limited perceptual range, is to tessellate the search region into bins defined by the perceptual range so that every location in the search space can be sensed by single visits to the locations $x_{\min} + r_p, x_{\min} + 3r_p, x_{\min} + 5r_p, \dots, x_{\max}$. If the forager has no spatial memory and hence cannot remember which location it visited on the previous search bout, it will be suboptimal.

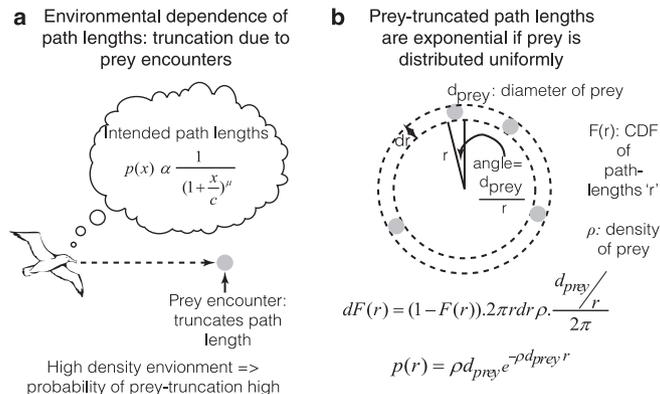


Fig. S2. Prey truncation leads to exponential path lengths when the density of food is high. Observed path lengths of foragers no longer reflect the intended path lengths when the distribution of prey is high due to prey-encounter truncation. The resultant path lengths can be shown to be exponential for a 2D environment with randomly distributed prey (*SI Results, 2.4 Truncation Due to Prey Encounter*).

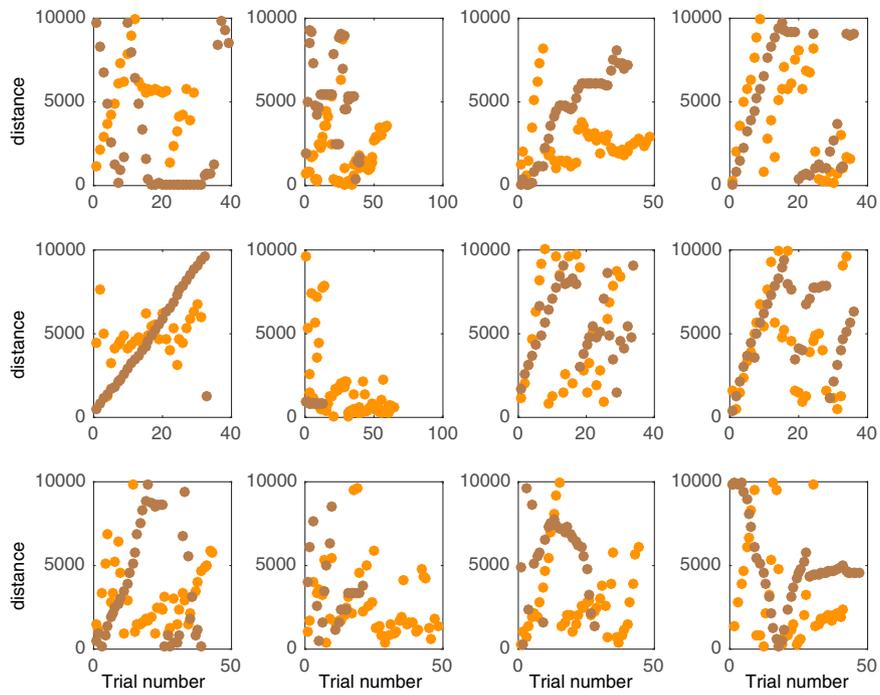


Fig. S3. Search pattern for all 12 human subjects for phase 1 and phase 2. Data from phases 1 and 2 are shown in brown and orange, respectively.

Data generated from hyperbolic distributions are better-fit
by **power law** than **exponential** distributions

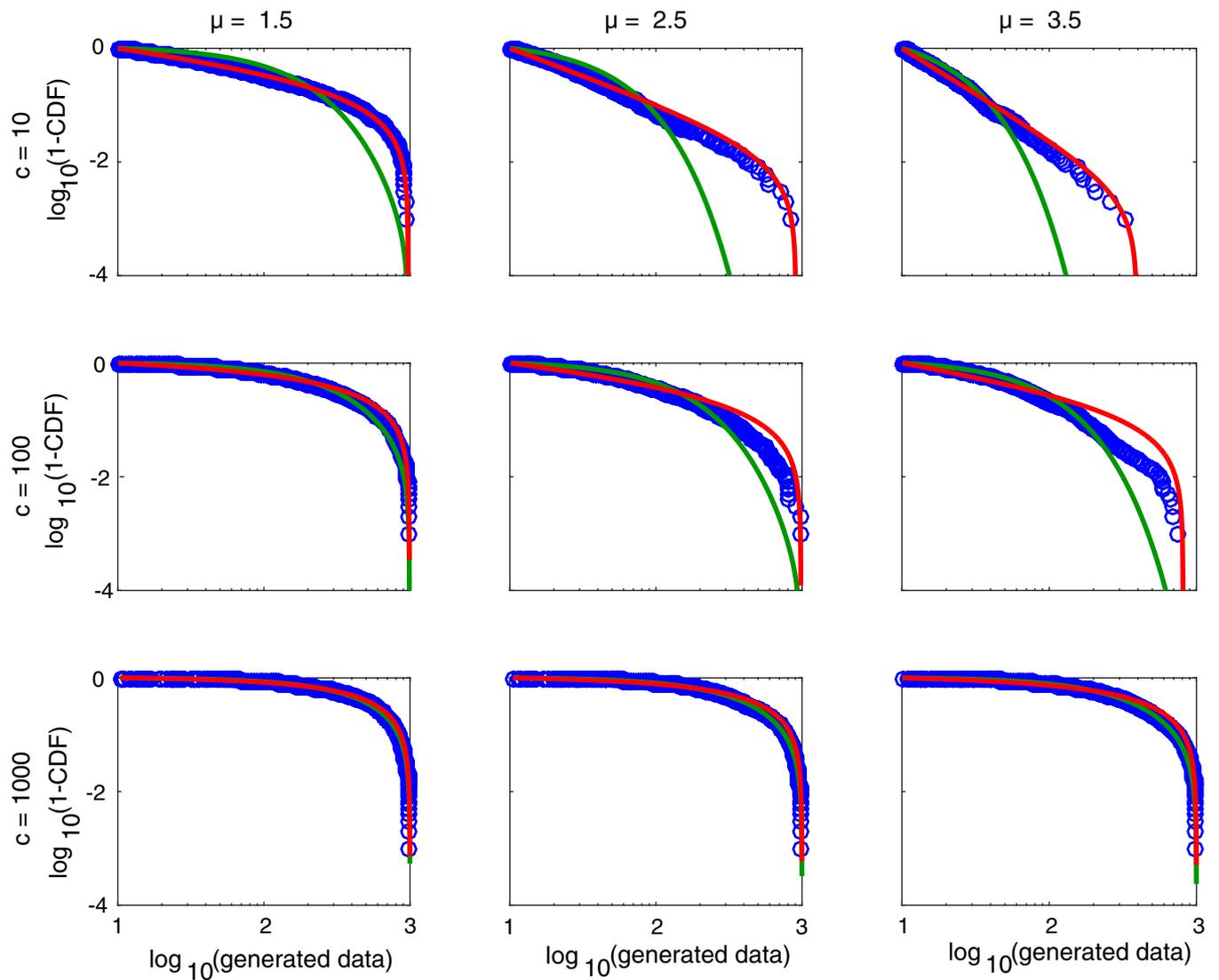


Fig. S4. Simulation shown in Fig. 3A repeated for more parameters. Truncation of random data generated from a truncated hyperbolic distribution was set to $x_{\min} = 10$ and $x_{\max} = 1,000$. The AIC strongly favored the truncated power law fit over exponential ($wAIC_{tp} = 1.000$ and $wAIC_{exp} = 0.000$) in all cases.

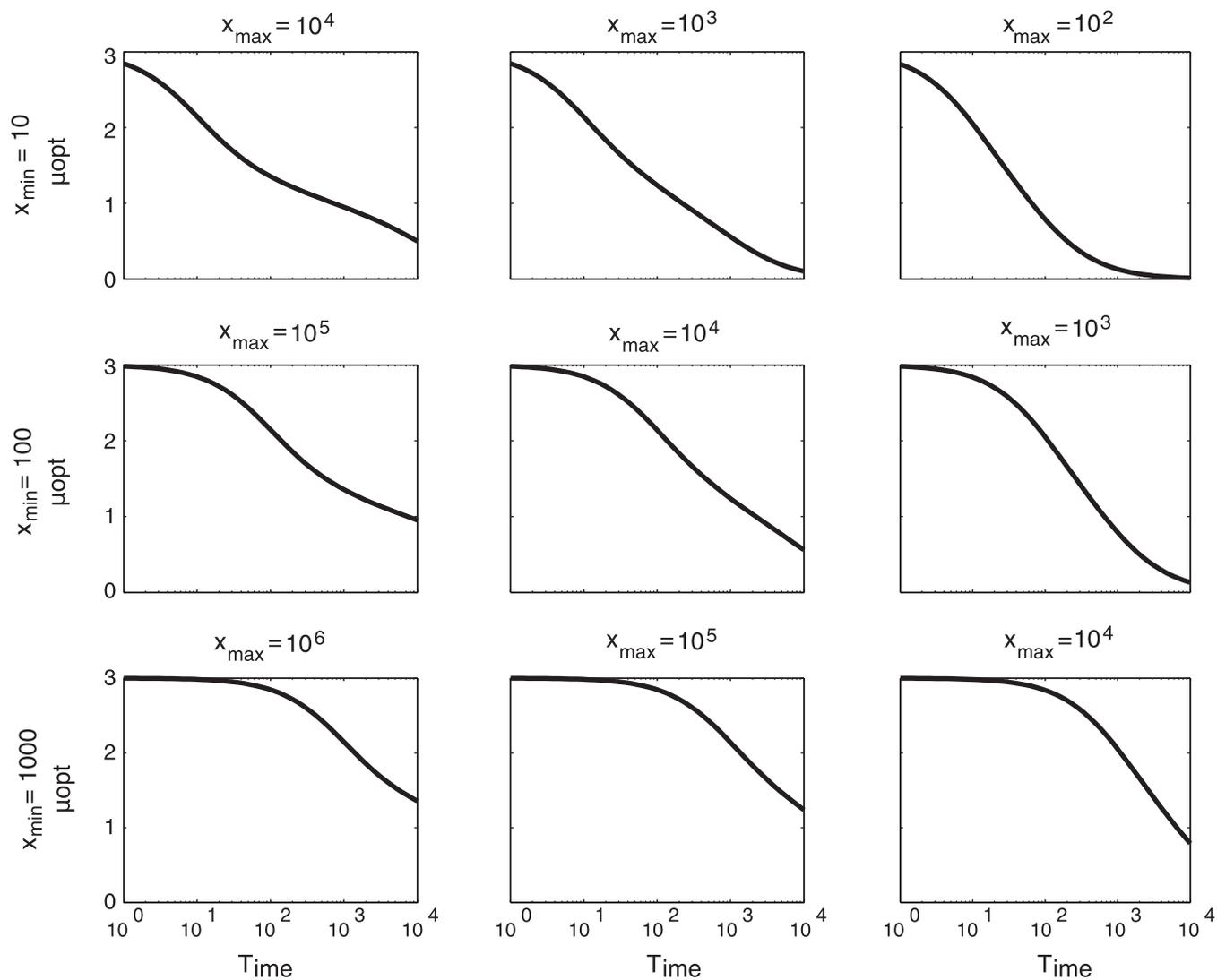


Fig. S6. Optimal exponents for power law fits to Eq. S8. If path length data conformed to Eq. 3 from the main text (derived for uncertainty in time perception), the closest approximation with a power law fit would have exponents as shown here. This is a numerical solution of Eq. S9 (SI Results, Eq. S9 (2.1.5) *Exploration under noisy temporal estimation*).



Movie S1. Shown is a sample from an experimental session of the human exploratory task. Each chapter shows a segment of the task. The segments are shortened in the movie for illustrative purposes. For complete details of the task, please see *Methods* and *SI Discussion, 1.1) Human Exploratory Task*.

[Movie S1](#)