A Linear Threshold Model for Optimal Stopping Behavior

Christiane Baumann[a,1], **Henrik Singmann**[b], **Samuel J. Gershman**[c], **and Bettina von Helversen**[d,a]

[a]Department of Psychology, University of Zürich, Binzmühlestrasse 14/19, 8050 Zurich, Switzerland; [b]Department of Psychology, University of Warwick, University Road, Coventry, CV4 7AL, United Kingdom; [c]Department of Psychology, Harvard University, 52 Oxford Street, Cambridge, MA 02138; [d]Department of Psychology, University of Bremen, Hochschulring 18, 28213 Bremen, Germany

This manuscript was compiled on April 30, 2020

In many real life decisions, options are distributed in space and time, making it necessary to search sequentially through them, often without a chance to return to a rejected option. The optimal strategy in these tasks is to choose the first option that is above a threshold that depends on the current position in the sequence. The implicit decision making strategies by humans vary but largely diverge from this optimal strategy. The reasons for this divergence remain unknown. We present a new model of human stopping decisions in sequential decision making tasks based on a linear threshold heuristic. The first two studies demonstrate that the linear threshold model accounts better for sequential decision making than existing models. Moreover, we show that the model accurately predicts participants' search behavior in different environments. In the third study, we confirm that the model generalizes to a real-world problem, thus providing an important step towards understanding human sequential decision making.

optimal stopping | cognitive modeling | sequential decision making | adaptive behavior

Decisions that arise in everyday life often have to be made when options are presented sequentially. For example, searching for a parking spot, deciding when to take a vacation day, or finding a partner, all require that the decision maker accepts or rejects an option without knowing if future options will be more attractive. Decisions in such problems involve a trade-off between accepting a possibly suboptimal option prematurely and rejecting the current offer out of false hopes for better options in the future.

Despite the importance of such decisions, relatively little work has been made toward characterizing the process by which humans decide to stop searching in natural settings of this task.

Earlier research has focused on a simplified version of optimal stopping problems, the so-called secretary problem, where only the rank of the option relative to those already seen is shown (1–3) and only the overall best alternative is rewarded. In the secretary problem, the optimal strategy is to ascertain the maximum of the first 37% options and choose the next option that exceeds this threshold (4). Empirical studies suggest that in general people follow a similar strategy but usually set the cut-off (i.e., from which point on they will accept an option that exceeds the previous options) earlier than the 37% prescribed by the optimal solution (1, 5).

Some studies have investigated tasks closer to real sequential choice problems by presenting the actual value of the option to the decision makers (6–10). In this version, the optimal is based on calculating the probability of winning on the later positions. From this probability, a threshold is calculated for each option in the sequence as described by Gilbert and Mosteller (4, Section 3). Lee (6) estimated a family of threshold-based models and showed that most participants decreased their choice thresholds as sequences progress. Although people are overall quite heterogeneous in their search behavior, they tend to cluster around the optimal solution (7, 8). Importantly, these studies still kept the restriction that only the best alternative is rewarded—a payoff function that does not correspond well with everyday experiences. Humans do find a mate, an apartment to live, or a ticket to fly to their vacation destination, and thus receive some payoff, even if that may not be the highest possible payoff.

In the present research, we propose a model of human decision making in optimal stopping problems using payoffs that are based on the actual values. In this variant of the search problem, the optimal decision thresholds are calculated based on the expected reward of the remaining options ((4, Section 5b) and SI Appendix, Text A). This leads to a decision threshold that changes notably nonlinear over the sequence.

In contrast, we propose that people rely on a mental shortcut and adapt their thresholds linearly over the sequence. We show that a model with this linearity assumption accurately captures when people stop search and accept an option, even in a real-world setting. Furthermore, this model allows us to predict under which conditions people search more or less than the optimal model, making it a useful tool to understand human sequential decision making.

We first sketch a family of cognitive models for describing behavior in optimal stopping problems. We then present results from three behavioral experiments that provide evidence for the validity of the linear model in a laboratory setting as well as in a real-world scenario.

Significance Statement

Behavioral research has made rapid progress toward revealing the processes by which we make choices between options that are presented simultaneously. Decisions in everyday life are typically more complex. We often encounter choices where options are separated in space and time and therefore the question is: "When is the right time to stop searching?" We suggest that humans use a probabilistic threshold. A model in which this threshold changes linearly over time, where the optimal policy prescribes a non-linear change, provides an excellent account to the data, even in real-life settings.

C.B., B.v.H., H.S. and S.G. designed research; C.B. and B.v.H. performed research; C.B., B.v.H. and H.S. contributed analytic tools; C.B. analyzed data; and C.B., B.v.H., H.S. and S.G. wrote the paper.

The authors declare no conflict of interest.

[1]To whom correspondence should be addressed. E-mail: cbaumann@psychologie.uzh.ch

Computational models. We explain the computational models based on a typical optimal stopping problem that we also used in our first two experiments. The decision maker (here a customer) is planing a vacation and decides to buy the plane ticket online. Ticket prices vary randomly from day to day and the customer wants to find the cheapest ticket. The customer checks the ticket price every day and decides if she wants to accept or reject the ticket, without having the option to go back in time to a previously rejected offer. Search time is limited by her vacation schedule (i.e., 10 decisions per trial) and, once accepted, the search ends.

More formally, we consider a decision maker who encounters a sequence of tickets with values denoted by $x_1, \ldots, x_{10}$ and she wants to find the minimum value in the sequence. If the decision maker accepts ticket x_i, the sequence terminates and she has to pay x_i; otherwise, she continues to the next ticket. When the last ticket is reached, it must be accepted.

All models assume that the decision maker relies on a probabilistic threshold to make the decision to accept or reject a ticket—i.e., ticket x_i on position i is compared to a position dependent threshold t_i. This comparison yields an acceptance probability θ_i based on a sigmoid choice function with sensitivity parameter β and

$$\theta_i = \frac{1}{1 + \exp\{\beta(x_i - t_i)\}}. \quad [1]$$

Small values of β produce more stochasticity in decisions, whereas the policy approaches determinism when $\beta \to \infty$.

We examine the setting of thresholds by comparing the performance of four different models.

- The *Independent Threshold Model (ITM)* serves as our baseline. It assumes no dependency between the thresholds. It entails N independent threshold parameters $t_1, ..., t_N$, one for each position in the sequence, where the decision maker can decide to accept or reject an offer (at position $N + 1$ the ticket must be accepted). The thresholds can take any value across positions. The model maintains maximal flexibility and provides an upper limit how well any threshold model can describe a person's decision given the assumption of a probabilistic threshold.
- The *Linear Threshold Model (LTM)* postulates that the thresholds are constrained by a linear relation to each other and therefore are completely defined by the first threshold t_0 and the linear increase δ as the sequence progresses:

$$t_{i+1} = t_i + \delta \cdot i, \quad [2]$$

  This model entails three free parameters, the first threshold t_0, the increase of the threshold δ and the choice sensitivity β.
- The *The Biased Optimal Model (BOM)* is based on the Bias-from-Optimal threshold model proposed by Guan et al. (8), assuming that humans are using thresholds that deviate systematically from the optimal thresholds.. The optimal thresholds t_i^* for each position i are derived by determining the expected reward of the remaining options (derivation in (4, Section 5b) and in SI Appendix, Text A). The model entails a systematic bias parameter γ that reflects the divergence of the human threshold from the optimal one. Additionally, the thresholds depend on a parameter α that determines how much their bias increases or decreases as the sequence progresses.

$$t_i = t_i^* + \gamma + \alpha \cdot i, \quad [3]$$

  When γ and α are set to 0, the thresholds represent the optimal thresholds that lead to best performance. This model is therefore defined by three free parameters, γ, α and the choice sensitivity β.
- The *Cut-off Model (CoM)* is inspired by the optimal decision rule for the rank information version of the secretary problem where the distribution of the prices is unknown. It assumes that the DM has a fixed cut-off value k that determines how long she explores in the beginning of the sequence. The highest value seen in that initial sample of k tickets is then set as her threshold, and the first value that exceeds this threshold in the remainder of the sequence is chosen. This model has two free parameters, the cut-off value k and the sensitivity parameter β.

Models were implemented in a *hierarchical-Bayesian statistical framework* using JAGS (11) (SI Appendix, Text B).

Experiment 1. We asked 129 participants to solve a computer-based optimal stopping problem following the ticket-shopping task described above. Tickets were normally distributed with a mean value of \$180 and a standard deviation of \$20. In the first phase, subjects learned the distribution using a graphical method proposed by (12) (*Methods*). SI Appendix, Fig. S1A shows that this procedure was successful in ensuring participants learned the distribution.

In the second phase, participants performed 200 trials of the ticket-shopping task. In each trial, participants searched through a sequence of ten ticket prices. For each ticket, they could decide to accept or reject it at their own pace. Participants were aware that they could see up to 10 tickets in each trial, and they were always informed about the actual position and the number of remaining tickets (SI Appendix, Fig. S2E for a screen shot). It was not possible to go back to an earlier option after it was initially declined. If they reached the last ticket (10*th*) they were forced to choose this ticket. When participants accepted the ticket, they received feedback about how much they could have saved if they had chosen the best ticket in the sequence. Performance was incentivized based on the value of the chosen ticket (*Methods*).

Behavioral results. Subjects earned on average 17.1 points (SD: 4.2) in each trial (maximum points = 20), which represents a 6% loss on optimal earnings. Participants' marginal accept probabilities steadily increased as the sequence progressed (Fig. 1A, black line), but differed systematically from the optimal agent's accept probability (Fig. 1A, yellow line). On the second-to-last (9*th*) position, participants accepted the ticket only with a 28%, 95%-CI [26%, 29%], probability, whereas following the optimal policy would result in a significantly higher acceptance rate of 50%.

Overall, subjects stopped earlier than optimal. The average position at which a ticket was accepted was 4.7 (SD: 2.9), whereas an optimal agent would have stopped at an average stopping position of 5.2 (SD: 2.8). However, a closer look at Fig. 1A reveals that whether subjects accept too early or too late depends on the position: on earlier positions they accept options although they should continue to search, whereas, if

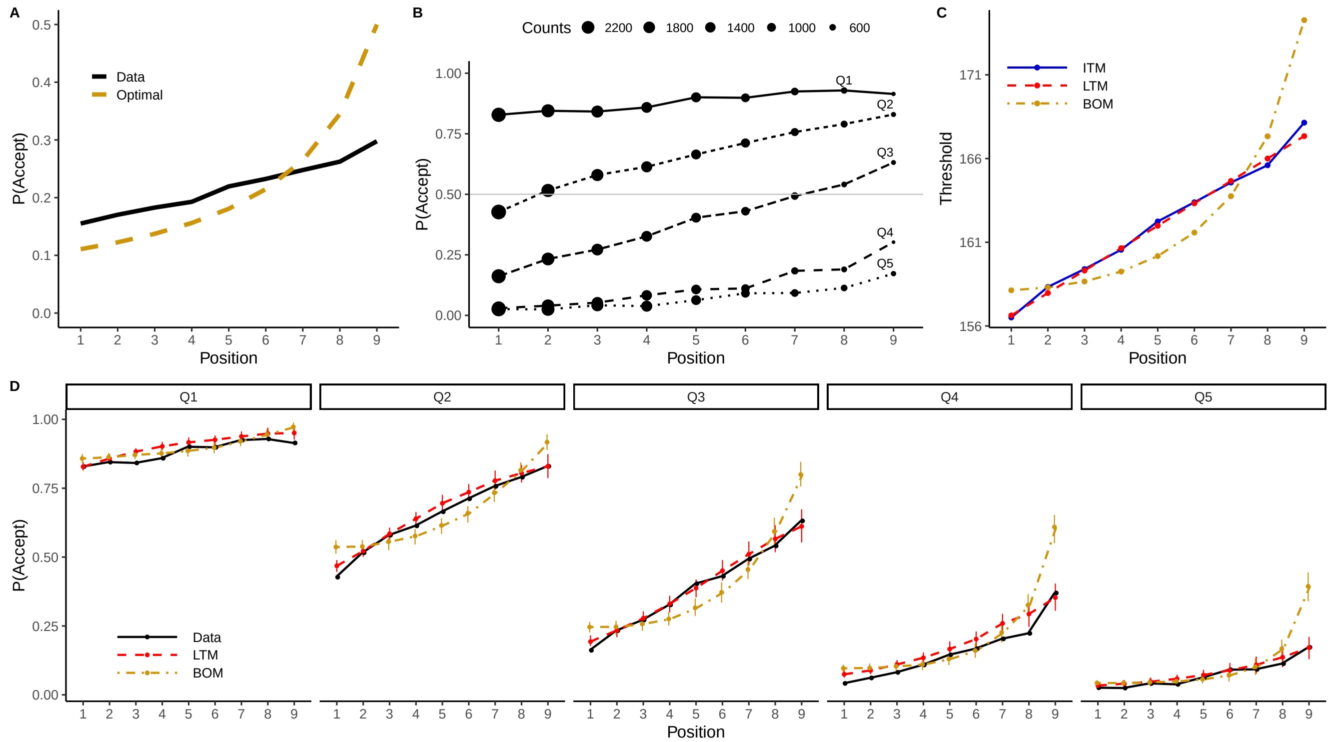

Fig. 1. (A) Probability to accept a ticket on each position across all prices. The dark line represents participant's frequency to accept, the dashed yellow line an optimal agent's probability to accept. (B) Participants' probability to accept. Each line represents ticket prices ranging from the first quantile to the fifth quantile. Q1: Tickets in first quantile, Q2: Tickets ranging from the first to the second quantile etc. The size of circles correspond to the number of data points on each position. (C) Estimated thresholds for the ITM with 9 free threshold parameters (solid blue line), the LTM with 2 free threshold parameters (dashed red line) and the BOM with 2 free threshold parameters (dash-dotted yellow line) (D) Posterior predictive mean and 95% HDI of the LTM (dashed red line) and the BOM (dash-dotted yellow line) for Q1 to Q5, as indicated in (B). Data: solid black lines

they get to position 7, they continue searching even for options that should be accepted according to the optimal policy.

Fig. 1B shows the accept probabilities conditional on ticket prices, split into the first five quantile ranges $Q1$ - $Q5$ (out of a total of ten quantile ranges). Qi is defined as the range of ticket prices from the $0.ith$ to the $(0.i - 0.1)th$ quantile of the ticket price distribution. In this experiment, the ticket distribution corresponds to a Gaussian distribution with mean 180 and standard deviation of 20. Accept probabilities for $Q4$ and $Q5$ did not reach 50% at position 9, in contrast to the optimal strategy that predicts much higher acceptance probabilities at this position.

Our models did not assume any learning over trials. This assumption was supported by an analysis of performance across trials. A linear mixed model on points per trial with trial number as fixed effect and by-participant random intercepts and random slopes for trial number showed no significant effect of trial number, $F(1, 64.00) = 0.02$, $p = 0.88$.

Modeling results and discussion. First, we checked whether the key assumptions of the modeling framework were supported. We calculated, per participant and model, posterior predictive p-values (p_{pp}) that compared misfit (i.e., deviance) of the observed data with misfit of synthetic generated data from the model. For the baseline model, ITM, this analysis indicated that the absolute fit was very good, and a probabilistic threshold adequately describes participants' responses; $p_{\mathrm{pp}} < .05$ for only 8% of participants (SI Appendix, Fig. S3A). For the vast majority of participants the observed misfit was consistent with the assumptions of the ITM plus sampling variability.

The performance of the LTM was almost identical to the ITM, suggesting that the considerably more parsimonious LTM (three free parameters for LTM compared to ten for ITM) adequately describes behaviour in optimal stopping tasks. The distribution of p_{pp}-values of the LTM was almost identical to the ITM (SI Appendix, Fig. S3A-B). Fig. 1D provides qualitative evidence of the agreement between LTM and data; the LTM adequately predicts accept probabilities for each quantile at every position (see SI Appendix, Fig. S4 for agreement between ITM and data). Fig. 1C compares the recovered thresholds of ITM and LTM and shows that the ITM thresholds essentially form a straight line lying exactly on top of the LTM thresholds.

The absolute fit of the BOM is clearly worse than for ITM/LTM; $p_{\mathrm{pp}} < .05$ for 35% of participants (SI Appendix, Fig. S3C). The source for this increased misfit can be seen in Fig. 1D. Only for Q1 and early positions of Q4 and Q5 did the BOM provide an adequate account. Furthermore, the recovered thresholds (Fig. 1C) of the BOM clearly differ from the ITM in almost all positions. Results of the CoM are not shown explicitly as its performance was extremely poor. All $p_{\mathrm{pp}} = 0$; there was not a single posterior sample for which the observed misfit of the CoM was smaller than for synthetic data generated from the CoM. Furthermore, choices were essentially random for CoM with $\beta_{\mathrm{CoM}} = 0.02$ [0.01, 0.06] (for the other models, $\beta \approx 0.21$).

Participants differed in their first threshold and slope pa-

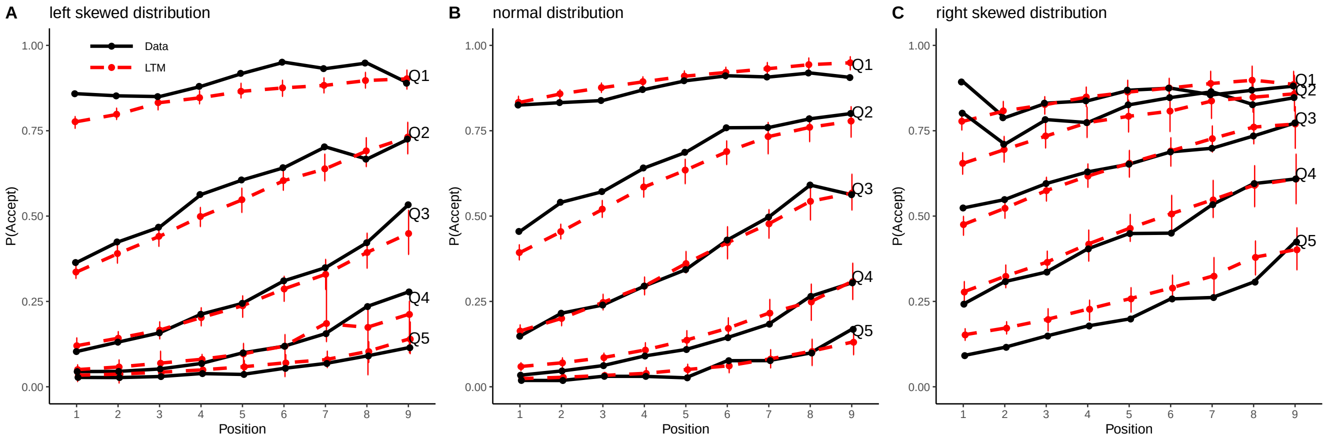

Fig. 2. Results of experiment 2: Empirical data appear in black lines and the posterior predictive means of the LTM in red lines. Bars represent the 95% HDI. The different lines represent the tickets ranging in from the Q1 to Q5. Q1: Tickets in first quantile, Q2: Tickets between the first and second quantile etc. (A) Condition 1: Tickets are left skewed distributed (PERT(40,195,200)) corresponding to a scare environment. (B) Condition 2: Tickets are normally distributed (PERT(90,140,190)). (C) Condition 3: Tickets are right skewed distributed (PERT(120,125,400)) corresponding to a plentiful environment.

rameters estimated by the LTM. However, all slope parameters are larger than 0 indicating that all participants increased the thresholds over the sequence (see also SI Appendix, Text C).

These results suggest that humans use a linear threshold when searching for the best option. In the present tests we found that the human performance is only 6% off from the performance of an optimal agent, indicating that the linear strategy performs quite well. Therefore, using linear thresholds could be an ecologically sensible adaptation to sequential choice tasks. However, it could also mean that the LTMs good performance might not generalize to new task environments, in which the linear model performs less well – an ability that would be crucial for the LTM to be a useful model of human behavior.

Search behavior in Experiment 1 indicated that people deviate from the optimal model depending on the price structure of the sequence: In trials with good options in the beginning people tended to accept them too early. However, in trials with few or no good options they continued search longer than the optimal model prescribed (SI Appendix, Fig. S5). Accordingly, in tasks with plenty of good options people might search less than optimal. However, in tasks in which good options are rare they might be tempted to search too long.

To find out and further predict how people will adapt to the tasks, we conducted a simulation study comparing the optimal solution with a best performing linear model (using a grid search to find the best performing parameter values for the linear model) and an empirical study manipulating the distributions of ticket prices across three conditions: (1) a left skewed distribution simulating a scarce environment, (2) a normal distribution, (3) a right skewed distribution simulating an environment with plentiful desirable alternatives. As illustrated in SI Appendix, Fig. S6B, the simulation study showed that the optimal model predicts more search in a plentiful environment, whereas a linear model predicts more search in the scarce environment. Furthermore, the linear model predicts a stronger decline in performance in the scarce environment than the optimal model (SI Appendix, Fig. S6A).

Experiment 2. To show that the LTM can capture people's choice behavior across different tasks and allows us to predict when people will search too much or too little we conducted a second experiment changing the distribution of options. We manipulated the different task environments by sampling tickets from (1) a left skewed (PERT*(40,195,200)), (2) a normal (PERT(90,140,190)) or (3) a right skewed distribution (PERT(120,125,400)), representing a scarce, a normal and a plentiful environment, respectively (SI Appendix, Fig. S1B-D, red lines). Each participant was assigned to only one condition. The final sample included 172 participants. The procedure was identical to Experiment 1, consisting of a learning phase, where participants got acquainted with the distribution (SI Appendix, Fig. S1B-D, participant's estimate in black lines), and a testing phase. In the testing phase, participants had to choose the lowest-priced ticket out of a sequence of 10 tickets with 200 trials (*Methods*).

Behavioral results. Participants' performance increased from the left-skewed (scarce) environment to the right-skewed (plentiful) environment ($F(2, 268) = 114$, $p < .0001$). As predicted by the best performing linear model, the loss compared to optimal performance was largest in the left-skewed condition, where only few good tickets occur (SI Appendix, Fig. S6A).

The average search length decreased from the left skewed scarce environment to the right skewed plentiful environment, $F(2, 268) = 11.5, p < .0001$. This pattern also follows the predictions of the best performing linear model in the simulation study but is in contrast to the optimal model's predictions (SI Appendix, Fig. S6B). Specifically, in the left skewed environment, where good tickets occur very rarely participants searched too long compared to an optimal agent, whereas in the environment where good tickets are abundant, participants ended their search too early compared to the optimal strategy.

Modeling Results and Discussion. Modeling results replicate the results from Experiment 1 and indicate that the LTM but not the BOM performed extremely well ($p_{\text{pp}} < .05$ for 7% to 10% of participants across the three conditions for LTM,

*The PERT distribution is a special case of the beta distribution defined by the minimum (a), most likely (b) and maximum (c) values that a variable can take and an additional assumption that its expected value is $\mu = \dfrac{a + 4b + c}{6}$.

but $p_{\text{pp}} < .05$ for 20% to 55% of participants for BOM, SI Appendix, Fig. S7). The observed accept probabilities (Fig. 2A-C, black lines, where each line represents a ticket price within the specified quantile range) are adequately described by LTM predictions (red lines) on almost all positions and in all three environments. Moreover, the threshold parameters for the ITM are again on top of the threshold parameters estimated by the LTM in all the three environmental conditions (SI Appendix, Fig. S8A-C).

These results indicate that humans use a linear threshold in optimal stopping problems, independent of the distributional characters of the task. However, this does not mean that people do not adapt to the task at all. Participants are responsive to task features and adapt their first threshold and the slope to the distributional characteristics of the task within the constraints of the linear model (SI Appendix, Fig. S8A-C).

Experiment 1 and 2 show that the linear model reflects a robust psychological process when deciding between sequentially presented options. However, in both experiments deciders were explicitly trained on the distribution of options, something not common in real life decision making. The next experiment tests if the linear strategy can also explain choices in a realistic optimal stopping task where initial learning is omitted.

Experiment 3. The decision maker's goal is to buy online products at the lowest rate where prices for this product are presented sequentially. We selected commodity products from different categories (e.g food, leisure, kitchen tools) and collected for each product a set of prices from Amazon.com. Only products with approximately normal price distributions were selected for a final set of 60 products (SI Appendix, Table S1). In the experiment, prices were sampled from a normal distribution, with a mean and standard deviation estimated from the real prices. All participants worked on 120 trials, divided into two blocks of 60 trials. In these two blocks, the 60 products were displayed in a random order (each product was encountered twice). Participants were aware that they could see up to 10 prices in each trial, and we indicated the average price of each product on the screen to reflect that people often have an idea of familiar products' prizes and to minimize individual differences in these.

Behavioral Results. Data from 95 participants were analyzed and replicated the results from Experiments 1 and 2 (normal distribution condition). Again, participants accepted too early, on average at position 4.6 (SD: 2.9). Comparing the performance in detail to the optimal strategy showed that (SI Appendix, Fig. S9) participants accepted too frequently at the beginning of the sequence (i.e., too low threshold) and searched too long towards the end of the sequence (i.e., too high threshold). We again found no evidence for learning across trials (linear mixed model on points per trial with trial number as fixed effect and by-participant random intercepts and random slopes for trial number showed no significant effect of trial number $F(1, 94) = 0.13, p = 0.72$).

Modeling Results. To deal with the prices' variability we normalized all values using mean and SD prior to fitting our models. We could replicate the results from Experiment 1 and 2, despite the fact that participants did not explicitly learn the product's prices beforehand: The LTM (10% of $p_{\text{pp}} < .05$, SI Appendix, Fig. S10A), but not the BOM (31% of $p_{\text{pp}} < .05$, SI Appendix, Fig. S10C), was able to capture the observed accept probabilities accurately on each position and for each quantile (Fig. 3B&C). Furthermore, threshold parameters estimated by the LTM were very similar to threshold parameters estimated by the ITM (SI Appendix, Fig. S11).

Discussion. In this paper, we designed a variant of an optimal stopping task that allowed us to quantitatively characterize the deviations of human behaviour from optimality. We found that humans apply a simplifying strategy, where thresholds are linearly increased over time. We implemented this assumption in a computational framework and demonstrated that this model not only provided an excellent fit to the data, it also outperformed other models found in the optimal stopping literature. Furthermore, the linear threshold assumption makes a non-trivial prediction about search length, which we confirmed experimentally: Humans stop earlier in environments with many desirable alternatives compared to scarce environments.

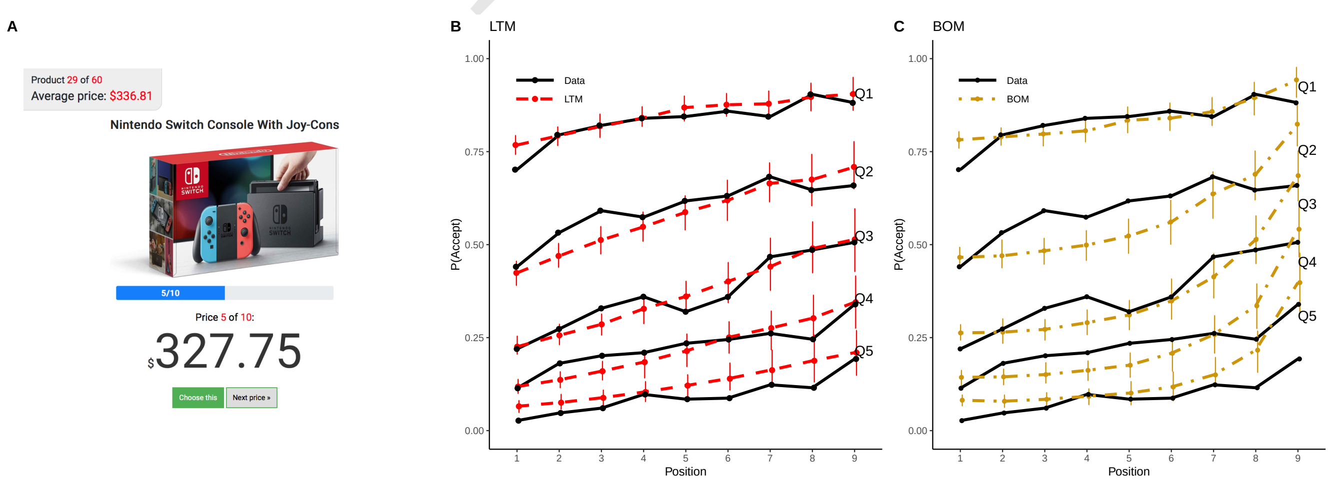

Fig. 3. (A) Screenshot of the product purchasing task. (B and C) Results of experiment 3: (B) Empirical data appear in solid black lines and the posterior predictive means of the LTM in dashed red lines. (C) Empirical data appear in solid black lines and the posterior predictive means of the BOM in dashed yellow lines. Bars represent the 95% HDI. The different lines represent the product prices ranging from the first quantile to the fifth quantile. Q1: Product prices in first quantile, Q2: Product prices between the first and second quantile, Q3: Product prices ranging from second to third quantile, etc.

These results contrast with the prediction from the optimal model. Finally, in a online product purchase paradigm we could show that our model generalizes to real-world sequential choice problems. Understanding how humans make sequential decisions will help quantify the conditions under which people may succeed or fail in such tasks.

But why are humans relying on a linear strategy in adapting their thresholds when an optimal policy is nonlinear? For one, our findings correspond well with recent studies demonstrating that human choice behavior in related explore-exploit paradigms is well described by a linear threshold rule (13, 14). But a human linearity bias seems to be more general. Indeed, a tendency to assume linear relationships has been reported in a range of domains such as function learning (15, 16) and reasoning (17–19). Crucially, simple strategies do not necessarily perform badly. In particular in uncertain and complex environments, simple heuristics can be efficient and powerful tools if they are adapted to the structure of the environment (20, 21). In this context, linearity could be considered as an adaptation of the human mind to its environment.

Materials and Methods

Participants. We recruited 438 participants (272 females; age range: 18-62; $N1 = 144$, $N2_{\text{left}} = 92$, $N2_{\text{normal}} = 110$, $N2_{\text{right}} = 92$, $N3 = 100$ in Experiments 1, 2 and 3, respectively) on Amazon Mechanical Turk to participate in the experiments. Participants gave informed consent, and the Harvard Committee on the Use of Human Subjects approved the experiments. Participants were excluded from analysis if they accepted the first option in a trial in more than 95% of the trials. After applying these criteria, we included data from 499 participants in the subsequent analysis $(\mathrm{N1} = 129, \mathrm{N2}_{\text{left}} = 86, \mathrm{N2}_{\text{normal}} = 102, \mathrm{N2}_{\text{right}} = 84, \mathrm{N3} = 95)$.

Task. In Exp. 1 and 2, participants performed the same online ticket shopping task that consisted of a learning and a testing phase. In the learning phase, participants experienced the distribution from which the ticket prices were drawn. In Exp. 1, the distribution from which the values were sampled was normal with $\mathcal{N}(\mu = 180, \sigma = 20)$. The procedure was as follows (SI Appendix, Fig. S2A-D): Participants encountered sequentially 50 ticket prices drawn from the predefined distribution. After every ten tickets, participants had to guess the average value of the tickets seen so far. After each guess, participants were told the correct response. At the end of the learning phase participants were asked to complete a histogram (by dragging the bars) for an additional 100 tickets that were drawn from the same predefined distribution. Participants received feedback by observing the correct distribution superimposed over their estimate (12).

In Exp. 2, we used three conditions to realize three different distributional environments, a left skewed distribution, PERT(40,195,200), a normal distribution, PERT(90,140,190), and a right skewed distribution, PERT(120,125,400). The procedure of the learning phase was identical to Exp. 1, except that we removed the section about reporting the mean for the skewed distributions (SI Appendix, Fig. S2B). Visual inspection of the performance in the histogram task suggested that participants learned the target distributions well (SI Appendix, Fig. S1).

In the second phase of Exp. 1 and 2, participants performed the ticket-shopping task. It started with a practice trial followed by 200 test trials. In each trial participants searched through a sequence of 10 ticket prices randomly drawn from the predefined distribution. For each ticket, they could decide to accept or reject it at their own speed. People were aware that they could see up to 10 tickets in each trial and they were always informed about the actual position and the number of remaining tickets (SI Appendix, Fig. S2E). It was not possible to go back to an earlier option after it was initially declined. If they reached the last (10^{th}) ticket they were forced to accept this ticket. When participants accepted the ticket, they received explicit feedback about how much they could have saved by choosing the lowest-priced ticket in the sequence (SI Appendix, Fig. S2F).

Participants were paid according to their performance. In each of the 200 trials there was a maximum of 20 points to earn. The participants received the maximum number of 20 points if they chose the lowest-priced ticket and 0 points for the worst ticket in the sequence. The payoff for a ticket that lied between the lowest-priced and the highest-priced was calculated proportional to the distance to the lowest-priced ticket in the sequence. The exact calculation for the points in each trial i was as follows:

$$points_i = \frac{20 \cdot (ticket_{max} - ticket_{chosen})}{ticket_{max} - ticket_{min}}, \qquad [4]$$

where $ticket_{max}$ represents the most expensive ticket in the sequence and $ticket_{min}$ the cheapest ticket in the sequence. Participants received a base payment of $4 and earned between $0 and $4 additionally depending on their performance.

In Exp. 3, participants performed an online product shopping task that started with a practice trial followed by 120 test trials divided into two blocks containing the same sixty products. In each trial, they encountered a product and searched trough a sequence of ten prices. Prices were randomly drawn from a normal distribution with a mean and standard deviation estimated from realistic prices collected from Amazon.com. Participants received a base payment of $2 and a performance contingent bonus between $0 and $4.

Data Availability. Data and modeling scripts are available on the Open Science Framework: https://osf.io/wqth3/.

ACKNOWLEDGMENTS. We thank Michael Lee and Peter Todd for helpful reviews and are grateful to Vassilios Kaxiras for helping with data collection. CB was supported by Swiss National Science Foundation (SNSF) (168889). HS was supported by SNSF (100014_179121). SJG was supported by the Office of Naval Research (N00014-17-1-2984). BvH was supported by SNSF (157432).

References.

1. Seale DA, Rapoport A (1997) Sequential Decision Making with Relative Ranks: An Exp Invest of the "Secretary Problem". *Organ Behav Hum Decis Process* 69(3):221–236.
2. Seale DA, Rapoport A (2000) Optimal stopping behavior with relative ranks: The secretary problem with unknown population size. *J Behav Decis Mak* 13(4):391–411.
3. Bearden JN, Rapoport A, Murphy RO (2006) Experimental studies of sequential selection and assignment with relative ranks. *J Behav Decis Mak* 19(3):229–250.
4. Gilbert JP, Mosteller F (1966) Recognizing the maximum of a sequence. *Springer Series in Statistics* 61:355.
5. Kahan JP, Rapoport A, Jones LV (1967) Decision making in a sequential search task. *Percept Psychophys* 2(8):374–376.
6. Lee MD (2006) A hierarchical bayesian model of human decision-making on an optimal stopping problem. *Cogn Sci* 30(3):1–26.
7. Guan M, Lee MD (2018) The effect of goals and environments on human performance in optimal stopping problems. *Decision* 5(4):339.
8. Guan M, Lee MD, Vandekerckhove J (2015) A hierarchical cognitive threshold model of human decision making on different length optimal stopping problems. in *Proc of the 37th annual meeting of the cognitive science society*. (Austin, TX), pp. 824–829.
9. Kogut CA (1990) Consumer search behavior and sunk costs. *J Econ Behav Organ* 14(3):381.
10. von Helversen B, Mata R (2012) Losing a dime with a satisfied mind: Positive affect predicts less search in sequential decision making. *Psychology and aging* 27(4):825.
11. Plummer M, , et al. (2003) Jags: A program for analysis of Bayesian graphical models using Gibbs sampling in *Proc 3rd int works on distr stat comp*. (Vienna, Austria), Vol. 124.
12. Goldstein DG, Rothschild D (2014) Lay understanding of probability distributions. *Judgm Decis Mak* 9(1).
13. Song M, Bnaya Z, Ma WJ (2019) Sources of suboptimality in a minimalistic explore–exploit task. *Nat Hum Behav* p. 1.
14. Sang K, Todd PM, Goldstone RL, Hills TT (2020) Simple threshold rules solve explore/exploit trade-offs in a resource accumulation search task. *Cognitive Science* 44(2).
15. Kalish ML, Griffiths TL, Lewandowsky S (2007) Iterated learning: Intergenerational knowledge transmission reveals inductive biases. *Psychonomic Bulletin & Review* 14(2):288–294.
16. Lucas CG, Griffiths TL, Williams JJ, Kalish ML (2015) A rational model of function learning. *Psychonomic bulletin & review* 22(5):1193–1215.
17. Little DR, Shiffrin R (2009) Simplicity bias in the estimation of causal functions in *Proceedings of the Annual Meeting of the Cognitive Science Society*. Vol. 31.
18. Wagenaar WA, Sagaria SD (1975) Misperception of exponential growth. *Percept Psychophys* 18(6):416–422.
19. Stango V, Yinman J (2009) Exponential growth bias and household finance. *J Finance* 64(6):2807–2849.
20. Gigerenzer G, Brighton H (2009) Homo heuristicus: Why biased minds make better inferences. *Topics in cognitive science* 1(1):107–143.
21. Todd PM (2001) Fast and frugal heuristics for environmentally bounded minds. *Bounded rationality: The adaptive toolbox* pp. 51–70.