Threshold Models of Human Decision Making on Optimal Stopping Problems in Different Environments

Maime Guan (hongyang@uci.edu)
Michael D. Lee (mdlee@uci.edu)
Department of Cognitive Science, University of California, Irvine
Irvine, CA 92617 USA

Andy Silva (aesilva@ucla.edu)
Department of Psychology, University of California, Los Angeles
Los Angeles, CA 90095 USA

Abstract

Optimal stopping problems require people to choose from a sequence of values, under the constraint that they cannot return to an earlier option once it is rejected. We study how people solve optimal stopping problems when the distribution of values they must choose from is not uniform, but is constructed to contain many high values or many low values. We present empirical evidence that people adapt to both sorts of environments, and make decisions consistent with using threshold-based models. We then fit a threshold model to our data, inferring the threshold people use, and finding they usually decrease their thresholds faster than is optimal as the sequence progresses. We also present empirical and model-based evidence that people generally do not adjust their thresholds on the basis of the values they see.

Keywords: optimal stopping; secretary problem; sequential decision-making; threshold models

Introduction

In optimal stopping problems, people must choose the maximum out of a set of numbers, under the constraint that a number can only be chosen when it is presented. People are told how many numbers are in the sequence, and that they must choose the last number if they do not choose an earlier number. For example, the sequence $73, 45, 56, 82, 27$ might be presented, one number at a time. The correct answer is 82, so the decision maker must not choose 73, 45, and 56 when they are presented, but must decide to choose 82 rather than be forced to take the final value 27.

Studying how people solve optimal stopping problems is interesting, because they have two features found in many real-world decision settings. The first feature is that *there is no going back*. When a choice must be made among a series of alternatives, it is often difficult or impossible to return to earlier options. In dating, once one potential partner is replaced by another, it is hard to go back. (The Tinder social application for dating make this "no going back constraint" explicit and non-negotiable). On long cross-country drives, once the gas station in a town is passed, there is a strong disincentive to turn around. In job recruitment, a candidate initially not offered a position may no longer be on the job-market if they are later sought.

The second feature of optimal stopping problems is that *only the best will do*. Often when a choice is made it needs to be the best one, and any inferior choice is not acceptable. In eye-witness identification line-ups, choosing the identical twin of the perpetrator is no better than choosing any other innocent suspect. Searching a key-chain requires finding exactly the right key, and it is equally a waste of time to try any other key, whatever its similarities to the correct one.

How people solve optimal stopping problems has been studied experimentally in a variety of contexts. Many studies focus on the classic "rank order" version, in which only the rank of the current alternative relative to those already seen is presented (e.g., Seale & Rapoport, 1997, 2000). Other studies focus on the "full information" version, in which continuously-scaled values for the alternatives are presented (e.g., Lee, 2006). For both versions of the problem, there are known optimal solutions processes, so that human performance can be compared to optimal performance (Ferguson, 1989; Gilbert & Mosteller, 1966). In the rank order version, the optimal solution involves waiting until a critical point in the sequence, then choosing the first option with rank one (if such an option exists) after that point. In the full information version, the optimal solution involves choosing the first currently maximal number that is above a threshold for the current position in the sequence.

In this paper, we focus on an under-explored but natural manipulation in optimal stopping problems. We change the nature of the environment from which the values are drawn, so that environments can either be *plentiful*, with lots of high values, or *sparse* with lots of low values. This manipulation is not very interesting in the rank order version of the problem, since the values underlying the ranks are not available to people. In the full-information version of the problem, however, people have access to the values, and so can learn about the distributional properties of the environment.

Optimal decision making involves setting higher thresholds in plentiful environments, and lower thresholds in scarce environments. A job-market awash with strong candidates allows for more selective recruiting

than one with weak candidates. Surprisingly, there appear to be very few studies of how people solve optimal stopping problems that manipulate the environment. Most studies use a single environment, usually with uniformly distributed values (e.g., Campbell & Lee, 2006; Kogut, 1990; Lee, 2006; Sonnemans, 1998). The only exception we are aware of is the early study by Kahan, Rapoport, and Jones (1967) that used three sets of 200 numbers, with the sets having the same mean but different variances, although there are related searching-and-stopping tasks for studying economic decision-making that have been studied with environment manipulation (e.g., Brickman, 1972; Hey, 1982).

Given the lack of previous work in studying how people solve optimal stopping problems in different environments, our goals are simple. We study how people solve optimal stopping problems in a plentiful and a sparse environment. In the next section, we describe the experimental method, and present basic empirical results relating to the accuracy of the decisions people make. We then use Bayesian methods to fit threshold models, which leads to some more detailed findings, including how people's decision-making relates to optimality.

Experiment

Method

Participants   A total of 56 UC Irvine undergraduate students participated in the experiment. Each participant was randomly assigned to either the plentiful or scarce environment condition, so that there were 28 participants in each condition.

Procedure   Participants were told to choose the highest out of a sequence of five random numbers ranging from 0 to 100, presented to two decimal places, under the constraint that they must choose a number when it is presented. They were also told that the only correct answer was the (unique) highest number out of the five, and that any incorrect answer is equally and completely incorrect. Each participant completed a total of 64 five-point optimal stopping problems, using a simple computer interface that presented the current value, showed its position in the sequence (e.g., "2/5" for the second position) and allowed the participants to choose or not choose the value with "Yes" and "No" buttons.. The interface provided feedback after each trial, and showed a cumulative record of the number of correct responses the participant had made over all of their problems.

Two different distributions were used to generate the stimuli. In the plentiful condition, the presented values were based on values generated as $v_{ijk} \sim \text{Beta}(4, 2)$, where v_{ijk} is the value the ith participant saw in the kth position on the jth problem they completed. In the scarce condition, $v_{ijk} \sim \text{Beta}(2, 4)$. Thus, participants in the plentiful environment were presented with numbers that were relatively large, and participants in the

scarce environment were presented with numbers that were relatively small. The environment condition manipulation was done between-subjects.

Empirical Results

Choosing the Current Maximum   We first checked whether participants completed basic components of the task properly. In order to choose the maximum out of a set of five, participants should not choose a value that is lower than an earlier alternative, excluding being forced to choose the final number. For example, if a 92 was not chosen when it was presented in the first position, then a 91 must not be chosen later on in the third position because it can no longer be the maximum. All but 4 participants chose the currently maximum value on over 90% of their problems. The remaining 4 participants (2 in the plentiful condition, and 2 in the sparse condition) met this standard on fewer than 65% of their problems Accordingly, we treated these 4 participants as contaminants, and excluded them from all of our analyses.

Accuracy   Figure 1 shows the overall performance and learning curves for the 52 non-contaminant participants. As has been emphasized in previous studies (e.g., Lee, 2006) optimal stopping problems afford two complementary ways to measure the accuracy of decision making. One is how often the correct maximum number in a problem sequence was chosen. The other is how often the number chosen was consistent with following the optimal decision process. The first is a measure of *correspondence*, based on matching the environmental truth, while the second is a measure of *coherence*, based on following rationally the available information to make a decision (see Dunwoody, 2009). The left panel of Figure 1 shows the percentage of problems for which individual participants were accurate in terms of both of these measures. In both environments, participants adhere to the optimal decision process about 60–80% of the time (with a few participants performing worse in the sparse environment). This leads to the maximum value being chosen about 40–70% of the time. Both of these findings are consistent with what has previously been found in environments where values are uniformly distributed (see Lee, 2006, Figure 2). The right panel of Figure 1 shows the learning curves, averaged over all participants, for both these measures. Perhaps surprisingly, but consistent with previous literature (e.g., Campbell & Lee, 2006; Lee, 2006) there is little evidence of learning. The curves are noisy, especially for the less stable maximum value measure, but, at least after 8 trials, there is no evidence of consistently improving performance.

Sensitivity to Position in Sequence   To examine how the position in the sequence affected people's decision making, we looked at the distribution of values that participants chose and did not choose at every position. Figure 2 shows these distributions for all five positions,
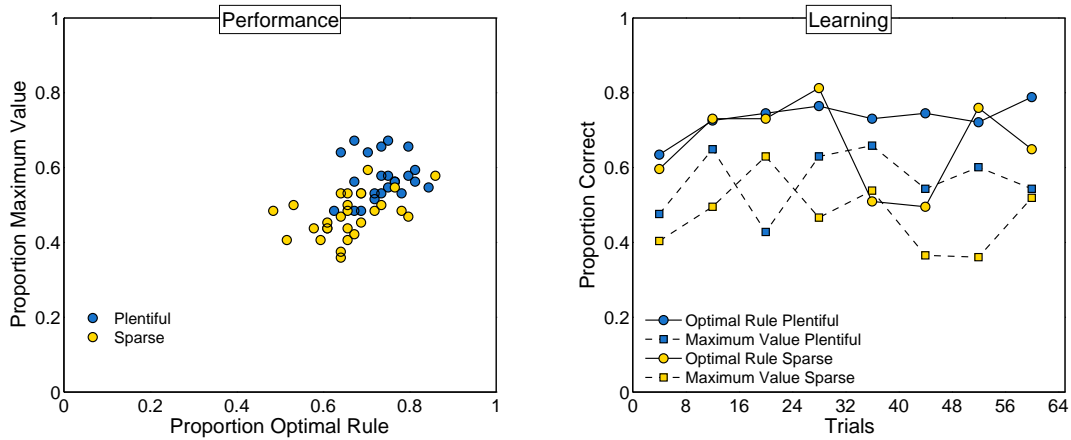
Figure 1: The left panel shows, for both plentiful and sparse environments, the proportion of trials on which the optimal decision process was followed, and the maximum was chosen, for each participant. The right panels shows the learning curves across trials for the average of these measures over all participants.
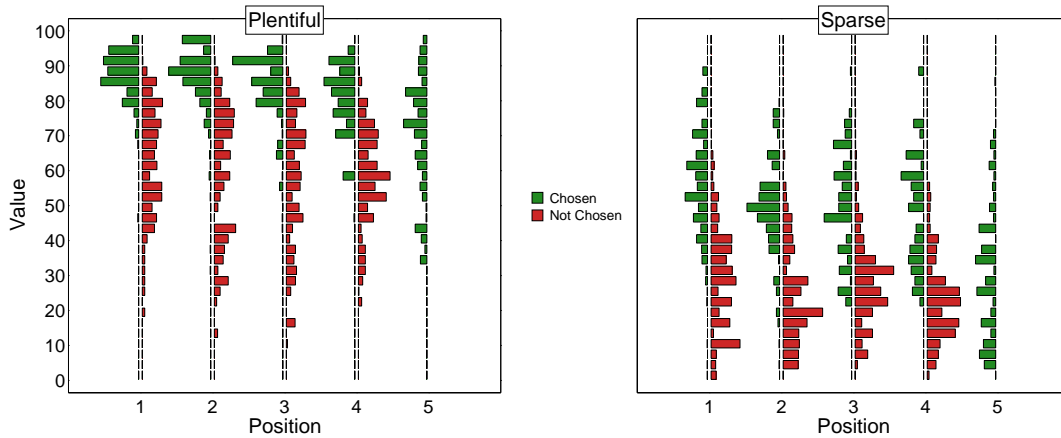

Figure 2: The distribution of values that were chosen and not chosen, as a function of position in the sequence, for both the plentiful (left panel) and sparse (right panel) environment conditions.
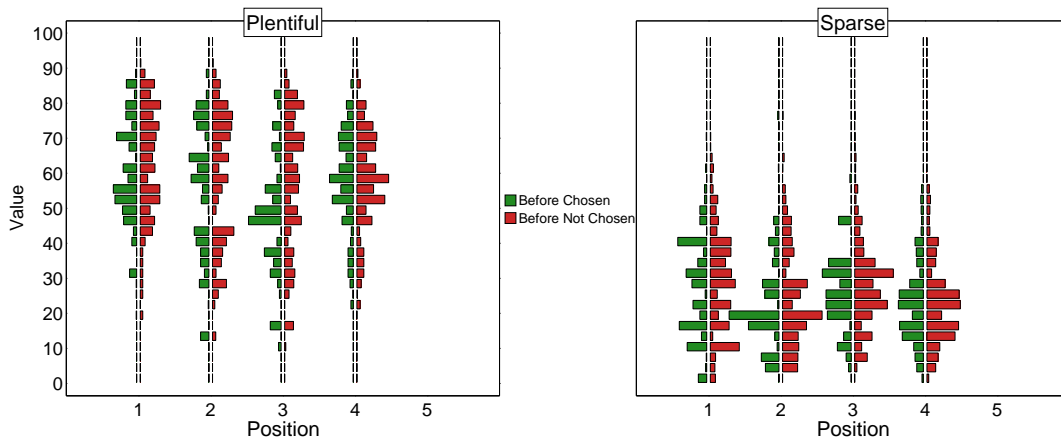

Figure 3: Histograms of immediate preceding stimuli depending on whether the next alternative in the sequence was chosen or not chosen. Note that Position 5 is empty because histograms show frequency of immediate preceding values.

collapsed across all participants, but separated into the plentiful and scarce environment conditions. It is evident that for every position, excluding values in the fifth position that participants were forced to take, the distribution of chosen values is higher than the distribution of values that were not chosen. Figure 2 also shows that the chosen values tend to be smaller in later positions in the sequence, and that values not chosen early in the sequence are chosen later. These two empirical regularities are consistent with the idea that people compare each option to a series of decreasing internal thresholds, as in the optimal decision making process.

Sensitivity to Preceding Value   Making optimal decisions on optimal stopping problems requires ignoring previous values in the sequence. Whether or not to choose the value 80 in second position should not be affected by whether the first value was 79 or 10. But people often make decisions sensitive to the context provided by earlier stimuli. To examine this possibility, Figure 3 shows the distribution of values in each position, separated by whether they immediately preceded a decision to choose or not chose the next presented value. For example, if the first two values in a problem sequence were 67 and 72, the value 67 would be part of the "before not chosen" distribution in position 1 if the participant did not choose the subsequent value 72, but part of the "before chosen" distribution in position 1 if the participant did chose the subsequent value 72. Visually, the distributions for "before not chosen" and "before chosen" in Figure 3 seem similar in each position, and for both environments. This suggests that the decisions made by participants are not strongly influenced by the preceding value in a problem. In our modeling analysis presented later, we provide a stronger test of this claim, using Bayesian model comparison

Threshold Model Analysis

The basic empirical results are consistent with a model in which people use a fixed sequence of potentially decreasing thresholds to decide whether to accept or reject presented values. There is, however, evidence of individual differences in performance in Figure 1, and so it is possible different people use different thresholds.

Model Definition and Implementation

Accordingly, we implemented a model based on a sequence of latent thresholds $\tau_{i1}, \ldots, \tau_{i4}$ for the ith participant in each of the first four positions where a choice must be made. We place the order constraints $\tau_{i1} \geq \tau_{i2} \geq \tau_{i3} \geq \tau_{i4}$ on these thresholds, so that they (non-strictly) decrease, and place uniform prior probability on the subspace of $(0, 100)^4$ these constraints define.[1]

---

[1]The modeling results change little if this order constraint is removed, but it captures relevant theory, and so should be included in the model (Vanpaemel & Lee, 2012).

Following the logic of the threshold model, the probability the ith participant will choose the value they are presented in the kth position on their jth problem is

$$\theta_{ijk} = \begin{cases} \alpha_i & \text{if } v_{ijk} > \tau_{ik} \ \& \ v_{ijk} = \max\{v_{ij1}, \ldots, v_{ijk}\} \\ \frac{1-\alpha_i}{4} & \text{otherwise} \end{cases}$$

for the first four positions and $\theta_{ij5} = 1 - \sum_{k=1}^{4} \theta_{ijk}$ for the last position. In these definitions, $\alpha_i \sim \text{Uniform}(0, 1)$ is a "probability of execution" parameter that measures how often the deterministic threshold model is followed by the ith participant (Lee & Newell, 2011).

The observed data are the positions chosen by each participants on each problem. Denoting by d_{ij} the position chosen by the ith participant on the jth, our generative probabilistic model is completed by $d_{ij} \sim \text{Discrete}(\theta_{ij1}, \ldots, \theta_{ij5})$.

We implemented this model as a graphical model using JAGS (Plummer, 2003), which is software that facilitates MCMC-based computational Bayesian inference (Lee & Wagenmakers, 2013). Our results are based on 4 chains of 1000 samples each, collected after 1000 discarded burn-in samples, and with the chains checked for convergence using the standard $\hat{R}$ statistic (Brooks & Gelman, 1997).

Model Results

We first examined the ability of the model to fit the behavioral data, using a standard Bayesian approach based on the posterior predictive distribution (Gelman, Carlin, Stern, & Rubin, 2004). This is the distribution of choices the model expects, based on the inferred joint posterior distribution over the model parameters τ_{ik} and α. Specifically, we found the mode of posterior predictive distribution for each participants on each problem, as a summary of the decision the model expects the participant to have made. The top panel of Figure 4 shows how often this decision agreed with the one the participant actually made, for all of the participants in both environments. Given the base-rate or chance level of agreement is 20%, the fact that the model generally captures 70–90% of the decisions a participant makes suggest it provides a reasonable account of people's behavior.

The two bottom panels of Figure 4 shows the marginal posterior expectations for all the inferred thresholds for all participants. Also shown, by the solid line, is the optimal threshold, based on the information provided by Gilbert and Mosteller (1966, Tables 7 and 8).[2] It is clear

---

[2]Gilbert and Mosteller (1966) provide thresholds for a uniformly distributed environment, which give the appropriate thresholds for our plentiful and scarce environments in terms of the percentiles that match the thresholds for the uniform distribution. For example, the threshold for the second-last position is 50 in a uniform distribution, which means the threshold for any other distribution in the second-last position is the median of that distribution.
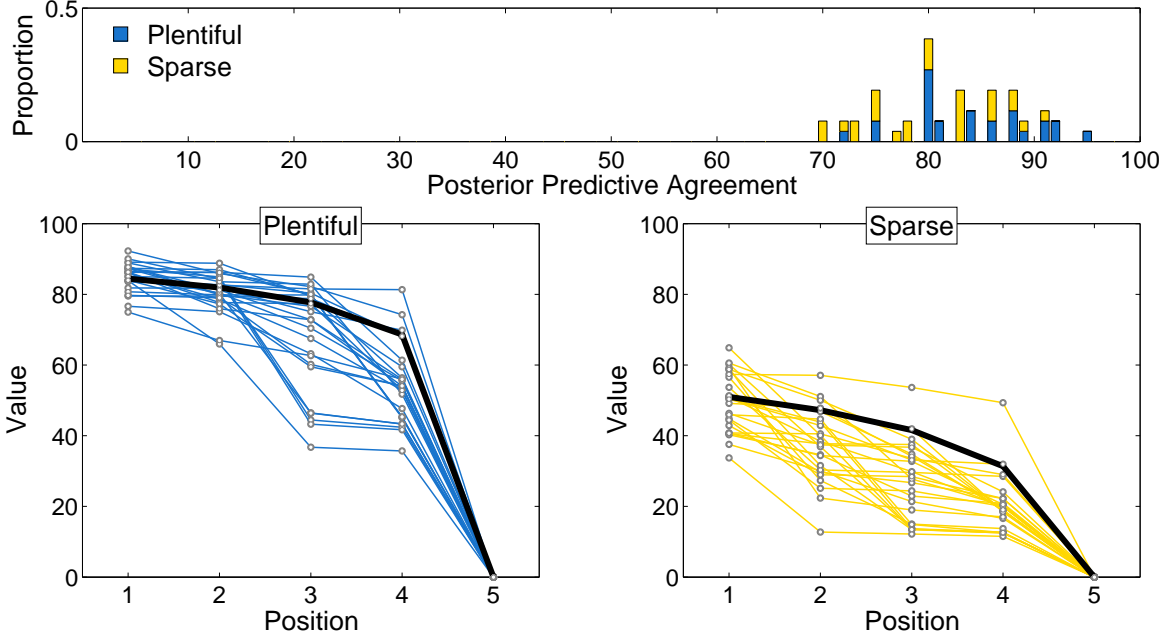
Figure 4: The top panel shows the distribution of posterior predictive agreement measure of fit for the threshold model for each participant, broken down by environment condition. the two bottom panels show the inferred thresholds for each participant in both the plentiful (left) and scare (right) environment conditions. The optimal thresholds for each condition are shown by solid black lines.

that participant performance is sensitive to environment, since thresholds in the plentiful environment are much higher than those in the scarce. It is also clear that there is individual variation in thresholds used across participants within both environments. Comparing the inferred participant thresholds to the optimal threshold, the majority of participants used lower thresholds than they should, in both environments.

Dependent Threshold Model Comparison

In the current model, the τ_{ik} thresholds vary by participant and position, but are insensitive to preceding values, capturing the assumption that participants do not adjust their threshold based on the context provided by this earlier information. Figure 2 presented some basic empirical evidence for this assumption. As a model-based test, we developed an extended threshold model in which the thresholds can be affected by the preceding value in a problem sequence. Formally, the affected thresholds are given by $\tau'_{ijk} = \tau_k + w_i \left(v_{ijk} - v_{ij(k-1)} \right)$, where $w_i \sim \text{Gaussian} \left(0, 0.01 \right)$ is a parameter measuring how the preceding value affects thresholds for the ith participant. Intuitively, the w_i acts to increase or decrease a threshold in proportion to the difference between the current and immediately preceeding value.

We compared this model to the original model using Bayes factors (Kass & Raftery, 1995), which is a standard Bayesian approach to comparing models that con-

trols for goodness-of-fit and model complexity. We estimated the Bayes factors using a latent mixture procedure based on model-indicator variables (Lee & Wagenmakers, 2013, Ch. 6). Figure 5 shows the distribution of log Bayes Factors for each participant. Also shown are standard interpretative boundaries at log-odds of 2, 6, and 10 corresponding to "moderate", "strong", and "very strong" evidence (Kass & Raftery, 1995, p. 777). It is evident that there is moderate to strong evidence in favor of the original model that assumes thresholds are independent of the preceding value.

Conclusion

Optimal stopping problems provide an interesting sequential decision making task that formalize two properties often found in real-world situation: once an option has been rejected it is no longer available, and only the best option is a correct choice. In an extension of most previous work, we studied how people solve short optimal stopping problems in environments where the available values are non-uniformly distributed. Our empirical results show that people still perform well, in terms of agreeing with the optimal decision making process as well as achieving the correct outcomes, in both plentiful and scarce environments. These results suggest that people are capable of identifying at least basic distributional properties of the environment, and tuning their decision making to match these properties.
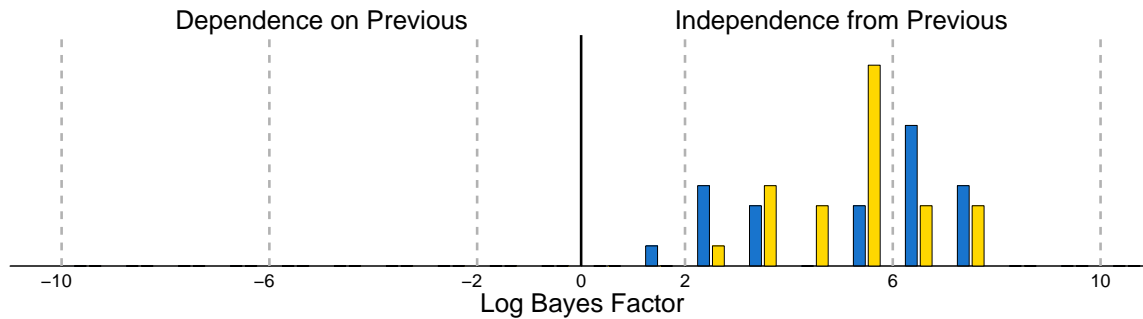
Figure 5: The distribution of log Bayes factors for all participants, broken down by environment condition, comparing threshold models assuming independence versus dependence on the preceding value in a problem sequence.

Our behavioral data also suggested that people may use threshold-based models to solve optimal stopping problems, maintaining a decreasing sequence of thresholds over the positions in the sequence. We presented empirical and model-based analyses that suggest these thresholds are subject to individual differences, often lie below the optimal thresholds, and are not affected by earlier values in particular problem sequences. Obvious directions for future work include understanding the basis of these deviations from optimality, the causes of the individual differences, and the relationship between human decision-making on this task and other sequential tasks involving risk and uncertainty.

References

Brickman, P. (1972). Optional stopping on ascending and descending series. *Organizational Behavior and Human Performance*, *7*, 53–62.

Brooks, S. P., & Gelman, A. (1997). General methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, *7*, 434–455.

Campbell, J., & Lee, M. D. (2006). The effect of feedback and financial reward on human performance solving 'secretary' problems. In R. Sun (Ed.), *Proceedings of the 28th Annual Conference of the Cognitive Science Society* (pp. 1068–1073). Mahwah, NJ: Erlbaum.

Dunwoody, P. T. (2009). Introduction to the special issue: Coherence and correspondence in judgment and decision making. *Judgment and Decision Making*, *4*, 113–115.

Ferguson, T. S. (1989). Who solved the secretary problem? *Statistical Science*, *4*, 282–296.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian Data Analysis* (Second ed.). Boca Raton, FL: Chapman & Hall/CRC.

Gilbert, J. P., & Mosteller, F. (1966). Recognizing the maximum of a sequence. *American Statistical Association Journal*, *61*, 35–73.

Hey, J. D. (1982). Search for rules for search. *Journal of Economic Behavior and Organization*, *3*, 65–81.

Kahan, J. P., Rapoport, A., & Jones, L. V. (1967). Decision making in a sequential search task. *Perception & Psychophysics*, *2*, 374–376.

Kass, R. E., & Raftery, A. E. (1995). Bayes factors. *Journal of the American Statistical Association*, *90*, 377–395.

Kogut, C. A. (1990). Consumer search behavior and sunk costs. *Journal of Economic Behavior and Organization*, *14*, 381–392.

Lee, M. D. (2006). A hierarchical Bayesian model of human decision-making on an optimal stopping problem. *Cognitive Science*, *30*, 555–580.

Lee, M. D., & Newell, B. R. (2011). Using hierarchical Bayesian methods to examine the tools of decision-making. *Judgment and Decision Making*, *6*, 832–842.

Lee, M. D., & Wagenmakers, E.-J. (2013). *Bayesian Cognitive Modeling: A Practical Course*. Cambridge University Press.

Plummer, M. (2003). JAGS: A program for analysis of Bayesian graphical models using Gibbs sampling. In K. Hornik, F. Leisch, & A. Zeileis (Eds.), *Proceedings of the 3rd international workshop on distributed statistical computing*. Vienna, Austria.

Seale, D. A., & Rapoport, A. (1997). Sequential decision making with relative ranks: An experimental investigation of the "Secretary Problem". *Organizational Behavior and Human Decision Processes*, *69*, 221–236.

Seale, D. A., & Rapoport, A. (2000). Optimal stopping behavior with relative ranks. *Journal of Behavioral Decision Making*, *13*, 391–411.

Sonnemans, J. (1998). Strategies of search. *Journal of Economic Behavior and Organization*, *35*, 309–332.

Vanpaemel, W., & Lee, M. D. (2012). Using priors to formalize theory: Optimal attention and the Generalized Context Model. *Psychonomic Bulletin & Review*, *19*, 1047–1056.