

Modeling Adaptive Search in Human Decision Making

Michael D. Lee

Department of Cognitive Sciences
University of California, Irvine

Ben R. Newell

School of Psychology
University of New South Wales

Joachim Vandekerckhove

Department of Cognitive Sciences
University of California, Irvine

Abstract

We study how people terminate their search for information when making decisions in a changing task environment. In three experiments, differing in the cost of search, participants made a sequence of two-alternative decisions, based on the information provided by binary cues they could search. Whether limited search was sufficient, or extensive search was required, changed across the course of the experiment, but was not indicated to participants. We find large individual differences but that, in general, the extent of search is changed in response to environmental change, and does not necessarily involve a reduction in accuracy. We then examine the ability of four models to account for individual participant behavior, using a generalization measure that tests model predictions. Two of the models use reinforcement learning, and differ in whether they use error or both error and effort signals to control how many cues are searched. Two of the models use sequential sampling processes, and differ in the regulatory mechanisms they use to adjust the decision thresholds that control the extent of search. We find that error-based reinforcement learning is usually an inadequate account of behavior, especially when search is costly. We also find evidence in the model predictions for the use of confidence as a regulatory variable. This provides an alternative theoretical approach to balancing error and effort, and highlights the possibility of hierarchical regulatory mechanisms that lead to delayed and abrupt changes in the extent of search.

Introduction

Search is a basic cognitive ability. People and animals have always needed to search their environment for basic needs like food, mates, and safety. More recently, people have needed to search their external information environment, to find out whether employers are recruiting new workers, if a curtain is available in an acceptable color, how humid potential holiday destinations are forecast to be, and what features are available on a smartphone. In addition, people have always needed to search their internal environments, containing their knowledge and memories. Retrieving, recalling and reconstructing information from memory is a basic pre-cursor to much of human thinking, decision-making, and action. There is a large literature on human and animal behavior in foraging, mate search, visual search, information search, and memory retrieval in addressing how people and animals search their external and internal environments. A recent overview of these disparate but related areas is provided by Todd, Hills, and Robbins (2012).

A pervasive characteristic of both external physical and information environments, and internal knowledge and memory environments, is that they are changeable. Seasons change, adversaries move, companies start and stop recruiting, features are added to phones, new facts are learned, and experienced events are forgotten. Finding relevant information is challenging enough in static environments, but becomes much harder in dynamically changing environments. There is a large literature on animal and human behavior in dynamic environments, measuring experimentally how changes are detected and environments monitored, and attempting to understand the learning and adaptation processes underlying decision making. (e.g., Brown & Steyvers, 2005; Gallistel, Fairhurst, & Balsam, 2004; Nassar et al., 2012; Otto, Gureckis, Markman, & Love, 2010; Speekenbrink & Shanks, 2010).

Understanding the intersection of these two areas—the basic cognitive capability of search, and the inherent changeability of search environments—is a difficult problem in cognitive modeling. It addresses the question of how people learn, adapt and regulate the way in which they search, simultaneously finding the information they need to make decisions and act, while evaluating and modifying the effectiveness of those search processes as the environment shifts. This raises questions of how people monitor changing environments, how they make predictions about the information they will find, how they make inferences about what information they could find, and how they adapt their search processes accordingly. Put simply, the general

challenge is to understand how people adapt search in dynamic environments.

In this paper, we consider one limited but important part of the general challenge. We focus on how people adapt the *termination* of their search. It is almost always possible to continue to search, and it is often not obvious when enough information, or the right sort of information, has already been found. When people make decisions about how many job listings to check, how many curtains to look at, how many holiday destinations to consider, and how many features to evaluate on a phone, they are making decisions about when to terminate their search. As an environment changes, different termination decisions may be needed. Job seeking is different depending on whether many or a few jobs are available, and if makers of smart phones introduce a few new “must have” features, there may be no need to evaluate the other features of a phone that is missing these crucial elements.

The cognitive modeling problem of understanding how people adapt their termination of search can be approached from a number of theoretical perspectives. One approach involves assuming there is an internal competition between multiple decision processes. Many models of adaptive decision making adopt this general and powerful approach (e.g., Busemeyer & Myung, 1992; Erev & Barron, 2005). The approach can be naturally applied to the problem of modeling the adaptive termination of search, by considering competing heuristic decision strategies with different search termination properties. A good example of this approach is provided by the Strategy Selection Learning (SSL: Rieskamp & Otto, 2006) model, in which reinforcement learning controls which of two different heuristics are applied over a sequence of decision making trials. One of these heuristics is take-the-best (TTB: Gigerenzer & Goldstein, 1996), which terminates search as soon as evidence that discriminates between the choice alternatives is found. The other heuristic is the Weighted Additive heuristic (WADD: Payne, Bettman, & Johnson, 1990), which assumes all of the available information is searched. Thus, the SSL model provides an account of when and why people switch between limited ‘one-reason’ search and exhaustive search of their decision making environment.

Another theoretical perspective is provided by sequential sampling models of the time course of decision-making. These models have their origins in stimulus sampling processes that naturally correspond to information search. The basic assumption is that people sample information from stimuli, accumulating the evidence they provide for alternative decisions, until a sufficient threshold evidence has been reached to

make one of these decisions. There are many sorts of sequential sampling models, making different assumptions about how information is sampled and accrued, and their development and evaluation is a long-standing and currently active sub-field of cognitive modeling (e.g., Bogacz, Brown, Moehlis, Holmes, & Cohen, 2006; Brown & Heathcote, 2008; Link & Heath, 1975; Ratcliff, 1978; Ratcliff & McKoon, 2008; Vickers, 1979).

The termination of search is naturally addressed within the sequential sampling framework; it is controlled by how thresholds are set, used, and adapted. The speed-accuracy tradeoff, perhaps the most basic issue in search termination, is controlled by setting thresholds, and empirical success in capturing this behavior is seen throughout the modeling literature. It has also been shown that setting thresholds in sequential sampling models can make them behaviorally equivalent to some of the most important heuristic models (Lee & Cummins, 2004; Newell, 2005; Newell & Lee, 2011). When the threshold of evidence needed to make a decision is low, sequential sampling models mimic one-reason decision-making heuristics like TTB. When the threshold is very high, sequential sampling models conduct exhaustive search, and mimic tallying heuristics, like WADD, that use all the available features, cues or information.

There is far less work on developing sequential sampling models that allow for the adaptation of search. This extension requires learning rules that specify how thresholds are adjusted within or between decision-making trials, and the ability to detect changes in the environment. Simen, Cohen, and Holmes (2006) develop a candidate model that uses reinforcement learning methods to change the thresholds on a drift-diffusion sequential sampling model, in a way that is sensitive to the reward rates and payoff structures in the decision making environment. Busemeyer and Rapoport (1988) present a series of experiments and models that also address the issue of the extent of search within a sequential sampling framework. They also consider different payoff structures, and compare various sequential sampling models of evidence accumulation, as well as different heuristic strategies, as accounts of human and optimal behavior. The changes in the environment demand different levels in the extent of search, directly addressing the problem we tackle. The models evaluated by Busemeyer and Rapoport (1988) do not, however, extend as far as providing a learning or adaptation mechanism for the trial-by-trial change in thresholds of evidence accumulation. And, as with Simen et al. (2006), the focus remains on accuracy, reward and

payoff as the determinants of when and why people should terminate search.

It seems unlikely, however, that animals and people adapt how much they search based solely on externally observable accuracy and immediate reward and payoff (Kheifets & Gallistel, 2012). Many decisions involving search, and requiring the termination of search, do not provide immediate corrective feedback, and so the accuracy of decision-making is not available. The environment does not usually send a signal indicating whether a better food source could have been found by more extensive foraging. Relatedly, many decisions regarding search, even if good ones, do not necessarily provide clear immediate reward. Changing how extensively applicants are vetted for recruiting may have non-obvious consequences that are not realized for many years.

One way to deal with the inadequacy of observed accuracy or reward as a signal for adapting search is to introduce additional useful signals. The SSL model (Rieskamp & Otto, 2006) considers signals based on effort, motivated by the idea that people are motivated not just to be accurate and obtain rewards, but also to expend as little effort as possible. The use of an effort signal marks an important theoretical shift towards allowing internally generated, rather than environmentally provided, measures of performance to guide the adaptation of search. The sequential sampling framework allows for learning, adaptation or self-regulation mechanisms for thresholds that are based on internally generated measures of performance. The best developed model using this approach is the Self-Regulating Accumulator (SRA: Vickers, 1979), which relies on internally-generated measures of confidence, rather than external measures of accuracy, to adapt thresholds and control the extent of search.

Our goal in this paper is to examine how people adapt when they terminate search in changing environments. Empirically, we are interested both in situations where it is likely changes in accuracy will signal the change in environment, and in situations where it is unlikely accuracy will be affected by environmental change. Theoretically, we are interested in both simple heuristic decision rules and sequential sampling models, and in adaptive mechanism based on both external signals like error-correction and internal signals like confidence.

The paper is organized as follows. In the next section, we describe a series of three experiments measuring how people terminate their search, and change their termination of search, in changing environments. The different experiments use the

same task and environments, but manipulate the costs and incentives involved in limited and expansive search strategies. We report basic empirical results for all of these experiments, including examining individual differences. We then develop a sequence of four models that make a natural theoretical progression, and evaluate them against the empirical data. The first model is a basic reinforcement learning heuristic that only adapts when errors are made; the second model is a reinforcement learning heuristic that is sensitive to both errors and effort; the third model is a sequential sampling model that is sensitive to both error and effort through the unifying mechanism provided by confidence; and the final model is a hierarchical extension of this sequential sampling model closely related to the established SRA model. We evaluate the usefulness of these models in describing, explaining and predicting the empirical data, and discuss the implications of our findings for future empirical and theoretical development.

Experiments

The experimental task was designed to meet two guiding principles. First, we wanted to include two kinds of changes in the statistical structure across trials, both requiring changes in how people terminated search, but with only one likely to be signaled by a change in accuracy. Secondly, we wanted to be able to make fine-grained trial-by-trial measures of the extent to which people searched, so that we could quantify how they terminated search. To achieve these aims we designed a task in which participants answered a series of 200 two-alternative forced choice multi-attribute decision problems. Using a standard approach, the task involved deciding which of two objects was higher on a particular criterion. To make a decision participants could access information from a minimum of one and a maximum of nine cues on each trial. These cues provided binary information about each object in the current trial pair. The dependent measure of interest, assessing when participants terminated their search, is quantified in terms of how many cues participants examine before making a decision on each trial.

Method

Participants. A total of 94 undergraduate students (Experiment 1 $N = 30$; Experiment 2, $N = 36$, Experiment 3 $N = 28$) from the University of New South Wales participated in return for course credit and, in Experiment 3 only, performance

related pay.

Stimuli. The experimental environment was created by selecting pairs of objects from the widely-studied German cities environment developed by Gigerenzer and Goldstein (1996). In the original environment, the objects correspond to German cities, the criterion is the population of each city, and the nine cues correspond to features like “is a state capital” and “has a team in the Bundesliga”. We re-described the objects as soil samples, the criterion as the energy efficiency of the samples, and the cues as the binary results of tests of the soil samples like “contains Actinium” and “seismic analysis is positive”. Each cue is naturally associated with a cue validity, which is the proportion of times it indicates the correct decision when it discriminates between objects. In our energy task re-description, the cue validities of 99%, 91%, 87%, 78%, 77%, 75%, 71%, 56% and 51% correspond to success rates for the nine tests.

Procedure. A schematic presentation of the experimental task is shown in Figure 1. Each trial requires the participant to make a decision about which of two soil samples is the more energy efficient fuel source. The stimuli are always represented in terms of the same cues, with the same validities, but whether each cue is positive or negative for each stimulus varies trial to trial. To search the soil test cue information, participants could click a “RUN TEST” button, and reveal the binary information. Participants had to run at least one test per trial, but were free to choose as many as they liked after that, before making their decision. The order in which tests could be run was fixed according to the cue validities, which were displayed beside the cue names.

The success rate for each test was presented on screen and was described to participants as follows: if a test has a success of 75% this means that if there were 100 trials in which one sample had a positive result (YES) for that test and the other sample had a negative result (NO) for that test, then the sample with the positive result would be the correct choice (be richer in the energy source) on 75 of those 100 cases, whereas for the remaining 25 cases the other sample would have been richer in the energy source (cf. Rieskamp & Otto, 2006). Following each decision feedback was provided, and a record of how many correct decisions had been made was shown on the screen throughout the experiment.

The experiment had a total of 200 trials which were subdivided into three blocks

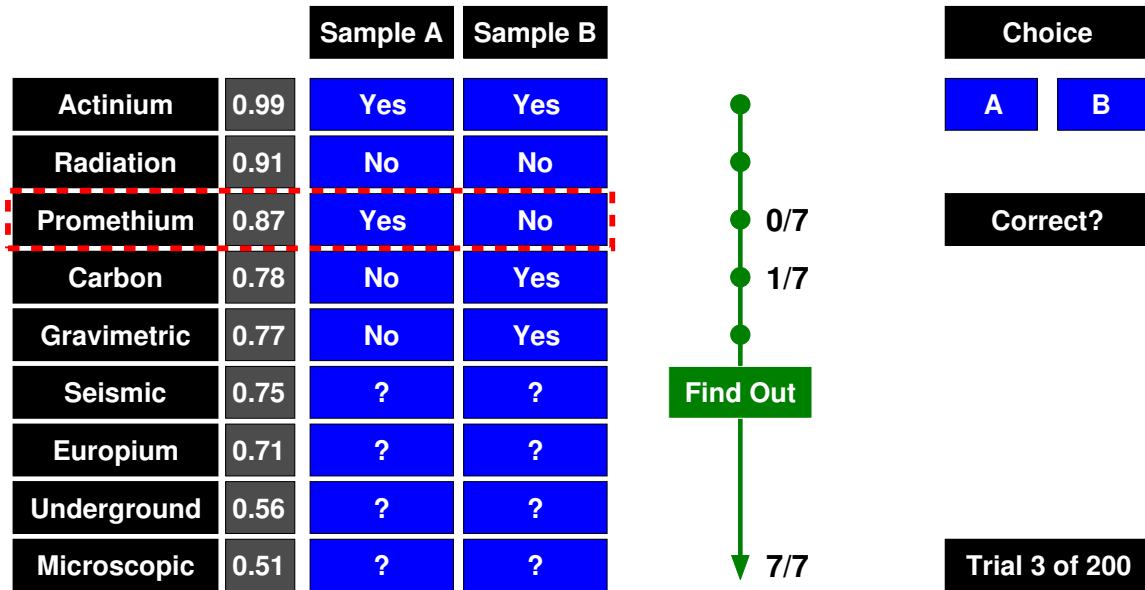


Figure 1. Schematic presentation of the experimental task, and the assessment of participant’s performance. On each trial, a sequence of binary cues can be searched in a fixed order. When search is terminated, a decision is made, and feedback is provided. The accuracy of decisions, and the extent of search—measured in terms of the proportion of cues beyond the first discriminating cue—are taken as measures of decision-making and search behavior.

of 50, 100 and 50 trials respectively. These subdivisions were not made explicit to the participants—from their perspective the experiment ran continuously from trial 1 to 200—but corresponded to change points in the statistical structure of the underlying environment. The experimental design is shown schematically in Figure 2. In the first block of 50 trials, participants learned in an environment in which the first discriminating cue always gave the correct answer. In the second block from trials 51 to 150, the first discriminating cue provided no information, corresponding to the correct answer on exactly half of the trials. In the third block from trials 151 to 200, as for the first block, the first discriminating cue always corresponded to the correct answer. Throughout all 200 trials, exhaustive search of all the cues, and the rational combination of the evidence they provided, always corresponded to the correct answer.¹ Thus, overall, the three blocks correspond to a first stage in which either limited or exhaustive search will lead to accurate decisions, a second block

¹By rational combination of the evidence, we mean the sum of the log odds defined by the cue validities (see Lee & Cummins, 2004; Lee & Zhang, 2012).

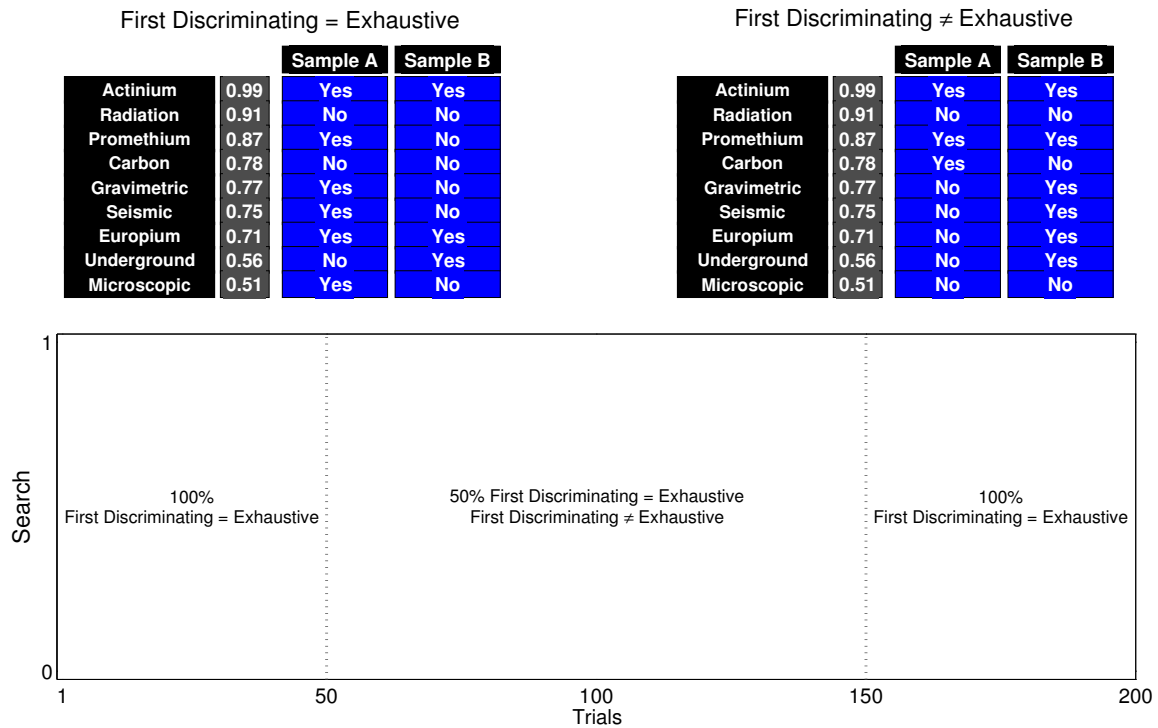


Figure 2. Schematic presentation of the experimental design, including example of the two problem types used to construct the environmental changes . In the first 50 trials, problems were presented for which searching to the first discriminating cue, as well as searching all cues, provided the correct answer. In the second block, the first discriminating cue gave the same answer as searching all cues exactly half the time, and searching all cues always provided the correct answer. The third block had the same properties as first block.

in which only exhaustive search is required, and a final block in which again either limited or exhaustive search will be effective.

The same 200-trial task was used in three separate experiments, which manipulated the costs and benefits related to information search. In Experiment 1 there was no time cost, and participants could run the tests to reveal information as quickly as they liked. In Experiment 2 there was a time cost to running each test. Specifically, participants had to wait for 3 seconds for the result of each test to be displayed on the screen. During this time a message with the words “Computer now running test” appeared on the screen. In Experiment 3, participants played for points that could be converted to dollars at the end of the experiment (100 points = AUD\$0.05). Participants were given 2000 points at the start of the experiment and could earn 70

points on each trial for a correct answer. However, each decision incurred a handling fee cost of 34.5 points (cf. Rieskamp & Otto, 2006). In addition, in the second block, clicking on each information button incurred a cost of 3 points. This cost structure combined with the changes in the statistical structure of the underlying environment encouraged limited search in the first and third blocks, and more extensive search in the second block. A reminder of the handling fee and reward was displayed on-screen throughout the experiment. Participants were told at the start of the experiment that on some trials they might need to pay for information. At the conclusion of the experiment, a final score was displayed, and participants were debriefed and paid a cash reward based on their score.

Empirical Results

The structure of the experimental trials, with three blocks, is designed to study how people change the extent of their search faced with two different sorts of environmental change. The first change occurs after trial 50, and requires that limited search be extended. This change is associated with an accuracy signal, because the use of limited search will lead to errors in decision-making. The second change occurs after trial 150, and allows for a return to limited search. This change is not associated with an accuracy signal, because both limited and exhaustive search are effective in the third block of trials.

In the language of fast and frugal heuristics, this design means that a one-reason decision-rule like TTB, which relies on the first discriminating cue to make a decision, is effective in the first and third blocks, but not the second. An exhaustive search heuristic like WADD is effective for all of the trials, and required in the second block. In the language of sequential sampling models, low thresholds on required evidence are effective in the first and third blocks, but larger thresholds are required in the second block.

The basic empirical interest is on whether and how people would be sensitive to the two types of changes in the task environment, and adapt the decisions they made, and how many cues they searched. The dependent measure of decision-making is straightforward, and corresponds to whether or not each decision was correct or an error. The dependent measure of the extent of search is more complicated. Following Newell and Lee (2009), we used the Proportion of Extra Cues (PEC) measure, which is illustrated in Figure 1. The PEC measure gives the proportion of cues beyond the

first discriminating cue that a participant chose to use. Recall that participants are required to run tests until discriminating information is found. At this point, there is some number n of remaining tests that could be run. Stopping search, and making a decision at this point, corresponds to a PEC measure of zero. Continuing search by running a further k tests corresponds to a PEC measure of k/n . Running all of the possible tests corresponds to a PEC measure of one. In this way, the PEC measure provides a normalized index of the extent of search for each participant on each trial.

Figure 3 provides visual summaries of the error measure of decisions and PEC measure of search for all of the participants in all three experiments. The same raw information—the PEC measure and accuracy of every participant on every trial in every experiment—is presented in two different ways. The first presentation format, in the upper column, focuses on individual behavior. The error counts and PEC measure are shown for each participant as lines. In the lower columns, the focus is on the distribution of behavior over individuals. The squares correspond to histograms, showing the distribution of error counts and PEC measures over the trial sequence. The change points between the three blocks at trials 50 and 150 are also indicated, and the overall average PEC behavior of all participants in the three blocks is shown by thick lines connecting cross markers.

The two different approaches to visual presentation are complementary, and together suggest a number of empirical findings. The most obvious one is the change in search behavior across blocks, involving an increase in the extent of search moving from the first to second block of trials, and a similarly-sized decrease in the extent of search moving from the second to third block of trials. This pattern of change is clear in the aggregated behavior across all participants, and in many of the individual participants. The same qualitative pattern is evident in all three experiments, although there are quantitative differences. The overall extent of search is higher in the first experiment, where there was no explicit cost for search. The increase and decrease in the extent of search is most marked in the third experiment, where the structure of the environment is aligned with the change in search strategies (Otto et al., 2010). The other obvious conclusion relates to the accuracy of decision making, and mirrors the changes in the patterns of search. Decisions made in the first and third blocks are relatively accurate, but there is a spike in errors at the beginning of the second block.

Overall, the empirical findings in Figure 3 are that people adapt their search to

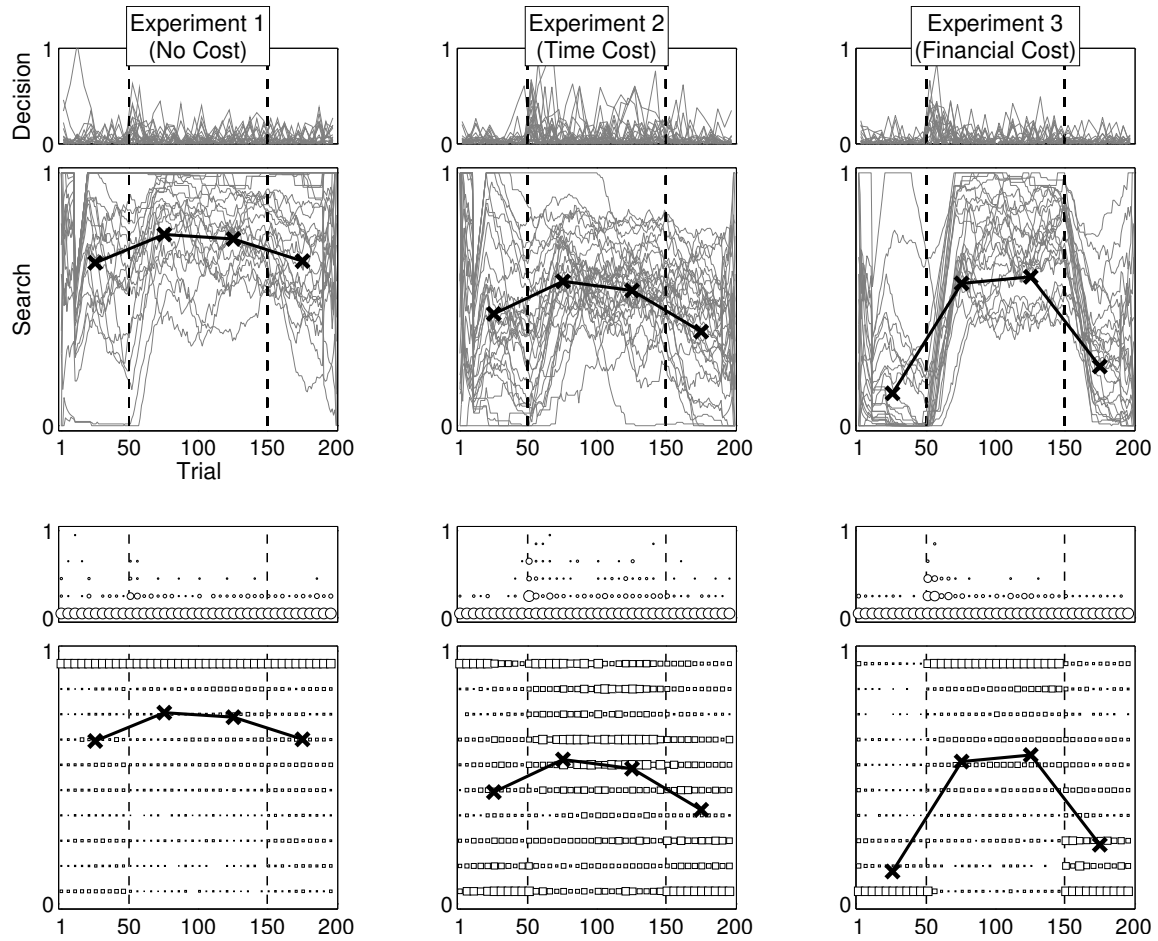


Figure 3. Summary of the search and decision-making behavior of all the participants in all three experiments. The top row of panels show the pattern of the extent of search and errors for each individual participant. The bottom row of panels shows the same information, but aggregated over participants, to make clear the distribution of search and errors. Mean behavior for each of the three experimental blocks is overlaid in both displays

changes in the task environment, including both increasing and decreasing their extent of search, with and without changes in accuracy. The patterns of these changes are sensitive to the cost of search, and, while there are clearly large individual differences in the detailed changes to searches and decisions, there are also broad regularities at the group level.

An Evaluation of Four Models

In this section, we develop and evaluate four models against the experimental data. Because the nature of the models and the data present significant challenges for evaluation, we first motivate our approach to evaluation. We then present the four models in a logical theoretical sequence, evaluating each model against its predecessor. Finally, we present some overall evaluation of the four models together.

Evaluation Method

Model evaluation for the models and task we consider presents a number of significant challenges. The first challenge is that the evaluation must simultaneously account for the decisions participants made, and how many cues they searched. Evaluating the fit of a model to bivariate behavioral data always requires making assumptions about the relative importance of capturing both sorts of behavior.

The second challenge is that the models have different levels of complexity, and these differences must be taken into account with goodness-of-fit (Myung, Forster, & Browne, 2000). In general, the models we consider do not have simple nested relationships to each other, with one being a special case of the other. The types of parameters and processes they use also vary widely, and make widely-used (and abused) model selection criteria like the AIC and BIC problematic. There is no reason to believe for most of the model comparisons we consider that a count of the number of parameters is a good approximation to model complexity.

The third challenge is that the models are not naturally probabilistic. As is often the case with heuristic models of decision making, the standard implementations of some models are deterministic. This means they do not automatically have a likelihood function, and so state-of-the-art Bayesian methods of evaluation cannot be applied.

There are difficult but principled ways in which these challenges could be addressed. For example, deterministic models can be made amenable to computational

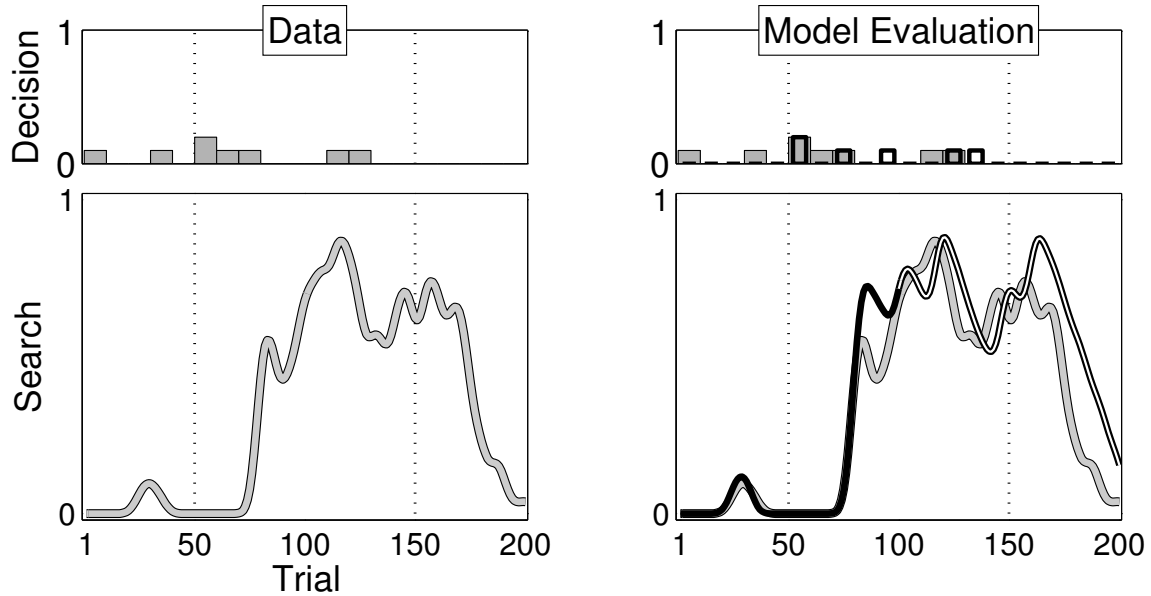


Figure 4. The basic approach to model evaluation, based on generalization performance. The decision and search behavior of a participant is summarized in the left panel. The fit of a model on the first 100 trials, based on best-fitting parameters is shown in the right panel, using a solid line. The generalization performance of the model on the second 100 trials, on which it was not trained, is also shown, using an open line. The agreement between participant behavior and model generalization predictions provide the assessment measure we use to evaluate models.

Bayesian inference by using synthetic likelihood methods, or theoretical extensions based on principles like entropification (Grünwald, 2000; Lee, 2004). Alternatively, it might be possible to define reasonable probabilistic versions of the basic models.

In this paper, however, we adopt a practical approach to model evaluation, summarized in Figure 4, based on generalization (Busemeyer & Wang, 2000). Generalization tests evaluate models by training under one set of circumstances, and testing how well they perform in different but related circumstances. Generalization tests differ from more routinely used prediction tests, like cross-validation, because the former require that testing is done on data coming from an environment that is different from the one used for training.

In some ways, our experimental design is well suited to this approach, because it is founded on a sequence of different environmental changes. A powerful test of a model on our task is to train it on the first 100 trials, so that it experiences the first type of environmental change, and then test the model on its ability to predict

behavior on the final 100 trials, which includes the second type of environmental change. Figure 4 provides a summary of our approach. The left panel shows the search and decision behavior of a participant, using the same approach to visual display as Figure 3. The right panel shows the behavior of a model that has been trained on the first 100 trials. This training results in parameter values being inferred that can then be used to generate model predictions for the final 100 trials. The search behavior of the model in the test trials is presented by the filled line, while the generalization test behavior is presented by the open line. The errors predicted by the model are shown as an open histogram overlayed on the gray histogram representing participant behavior.

The generalization measure we use to evaluate a model against a participant corresponds to the agreement of the predicted behavior on the last 100 trials. This approach is shown in the concrete example in Figure 4. The solid model line for the first 100 corresponds to its training performance in matching the human behavior, while the open line in the last 100 trials corresponds to its generalization performance. Only the agreement between the open line and human behavior is used to measure generalization, consistent with the logic of assessing prediction to control model complexity.

This generalization measure provides a useful practical assessment of a model. Because it relies on data that not only were not used to train the model, but involve a different environment change from the training data, an overly-complicated model cannot perform well simply by describing the training data. Good generalization performance on the test data provides compelling evidence that a model is capturing basic aspects of the way a participant is choosing to terminate their search and make decisions.

More formally, a participant’s behavior is represented by the decision d_i and number of cues searched t_i on the i th trial. A model with a set of parameters θ makes predictions $\hat{d}_i(\theta)$ and $\hat{t}_i(\theta)$ about these behaviors. To find the best-fitting values of these parameters on trials 1 to 100 in the training, we present the model with *exactly* the same sequence of trials received by the participant, and consider the sum-squared error measure

$$\text{SSE}_{\text{train}}(\theta) = \sum_{i \in \text{train}} \left\{ w(\hat{d}_i(\theta) - d_i)^2 + (\hat{t}_i(\theta) - t_i)^2 \right\}. \quad (1)$$

We found the parameter combination θ^* that minimizes this sum-squared error measure using a direct search optimization method known as iterated grid search (Kim, 1997; Thisted, 1988)

We then used θ^* to allow the model to make predictions about the test trials 101 to 200. The performance of the model follows that used in training, so that

$$\text{SSE}_{\text{test}}(\theta^*) = \sum_{i \in \text{test}} \left\{ w(\hat{d}_i(\theta^*) - d_i)^2 + (\hat{t}_i(\theta^*) - t_i)^2 \right\}. \quad (2)$$

A conveniently scaled final measure of generalization error performance is then the root-mean-squared-error $\sqrt{\text{SSE}_{\text{test}}(\theta^*)}$. This evaluation procedure was used for every model and every participant, based on the model receiving exactly the same sequence of problems seen by that participant.²

The value of w determines the relative weight given to capturing decisions. The results reported in this paper are based on $w = 1$ to give equal weight to both components of task behavior. We also examined results using $w = 5$ and $w = 20$, giving greater weights to decisions. These error measures are worth considering, because it is hard to evaluate a model that predicts the correct number of search cues but leads to the wrong decision. Higher values of w attempt to ensure the decisions of participants and the models agreed. In practice, we found that participants made relatively few errors, and most parameterizations of the models resulted in relatively few errors, and so there was good agreement on decisions, and it was not sensitive to the value of w . Hence we settled on $w = 1$ giving equal weights to both decision and search in evaluating models.

Modeling Search and Decisions

All four of the models we consider are placed on equal footing, by using the same process to accumulate evidence as cues are searched, and make decisions once search is terminated. The models vary in how they terminate search and, most interesting, in how they adapt their termination of search over the sequence of trials in the experiment.

²Because the PEC measure jumps between discrete values from trial-to-trial, it is difficult to present in an informative way without some sort of smoothing. Figure 4, and the remaining visual displays of search behavior in this paper, rely on a exponentially decaying smoothing filter. This approach to smoothing gives greatest emphasis to the trial being displayed, but averages it over adjacent trials, giving progressively less weight to trials further from the one being displayed.

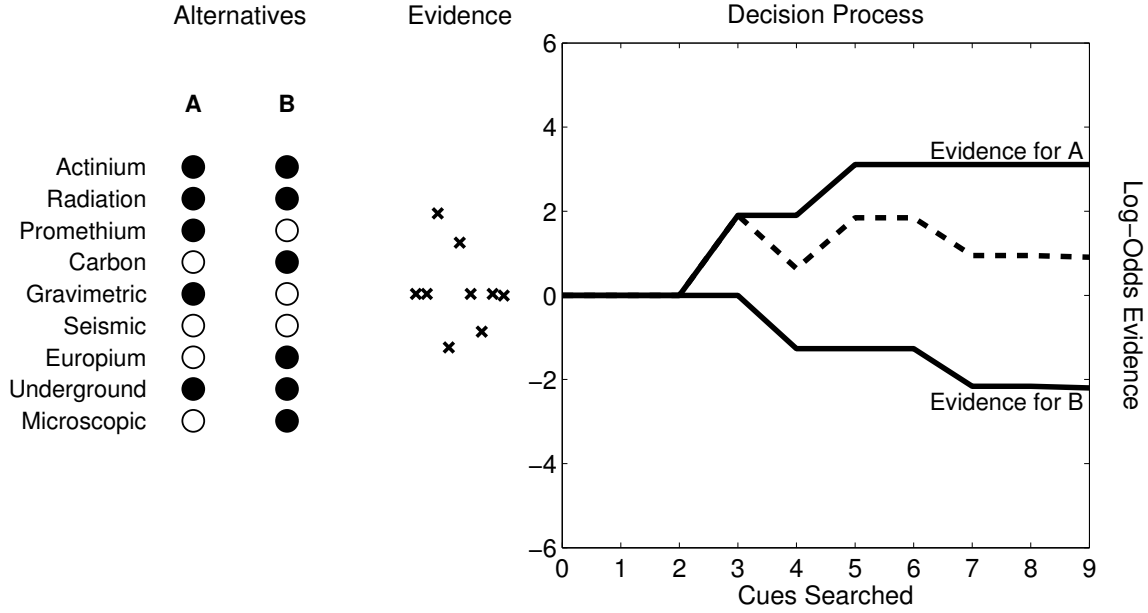


Figure 5. The basic search and decision processes used by all of the evaluated models. On a trial, the cues provide a sequence of evidence values, on a log-odds scale, that are accumulated. Both the tallies specific to each alternative, and the difference between these tallies, are accumulated. When search is terminated, the alternative with the greatest evidence is chosen.

Figure 5 shows the basic search and decision processes that are common to all of the models, at the level of a single trial. On the far left are the two alternatives presented on the trial, represented by two columns of circles that correspond to their cues. Cues that are present for an alternative are shown by black dots, and cues that are absent are shown by white dots.

The cues are searched from the top to bottom, and generate evidence in favor of one or other alternative when they discriminate. The sequence of these evidence values is shown by the crosses in Figure 5, which progress from left to right as cues are searched. The evidence values are calculated on a log-odds scale, since this is the natural additive scale for aggregating evidence (e.g., Cover & Thomas, 2006). Formally, If the k th cue discriminates in favor of the first alternative, the definition of cue validity means that the evidence it provides in favor of that choice is $\log(v_k / (1 - v_k))$ (Bergert & Nosofsky, 2007; Lee & Cummins, 2004; Lee & Zhang, 2012). Evidence in favor of the first (left-hand) alternative corresponds to the positive values, while evidence in favor of the second (right-hand) alternative corresponds to negative values.

These evidence values can be accumulated in a number of ways. Figure 5 shows the two approaches used in the models we consider. The solid lines labeled as “Evidence for A” and “Evidence for B” correspond to the cumulative sum of the evidence values that favor the alternative. Thus, the upwards-moving tally line is the accumulated evidence for the Alternative A as cues are searched, and the downwards-moving tally line is the evidence for Alternative B. The difference between these two tallies is shown by the broken line, and measures the signed evidence in favor of one or other of the alternatives as evidence is accumulated.

Both of these approaches to accumulating evidence are widely used in cognitive modeling, especially in sequential sampling models, where the first approach corresponds to race or accumulator models (Vickers, 1970, 1979), and the second corresponds to random-walk or diffusion models (Ratcliff & McKoon, 2008). Obviously, the two approaches lead to the same decision, choosing the alternative with the most evidence, once search is terminated after some number of cues. Thus, the differences between the four models we consider involve *when* search is terminated within a trial, and *how* the termination of search is adapted over a series of trials.

Error-Based and Effort-Based Learning

As mentioned in the introduction, a simple and widely-used account of how people adapt and learn is that they try to avoid making errors. Error-based learning is a cornerstone of psychological theorizing, and is naturally applied to model the regulation of search in our task. The basic idea is that, if people choose the wrong alternative, they will increase the extent of their search on future trials, aiming to collect more evidence, and so make more accurate decisions.

An extension of this approach is to consider not only decision accuracy, but also the extent of search, as signals or inputs into a learning process. The idea is that people are sensitive not only to errors, but also to the effort required to make a decision, and seek to minimize both. A simple way to implement this more general set of goals is to increase search if an incorrect decision is made, trying to minimize errors, but decrease search when a correct decision is made, trying to minimize effort.

Figure 6 shows how these mechanisms for regulating search are implemented in the models we consider. On the t th trial, a proportion of cues γ_t is searched, and the decision corresponds to the alternative with the greatest evidence is chosen. In the *error* model, search is increased by a learning parameter α following an error.

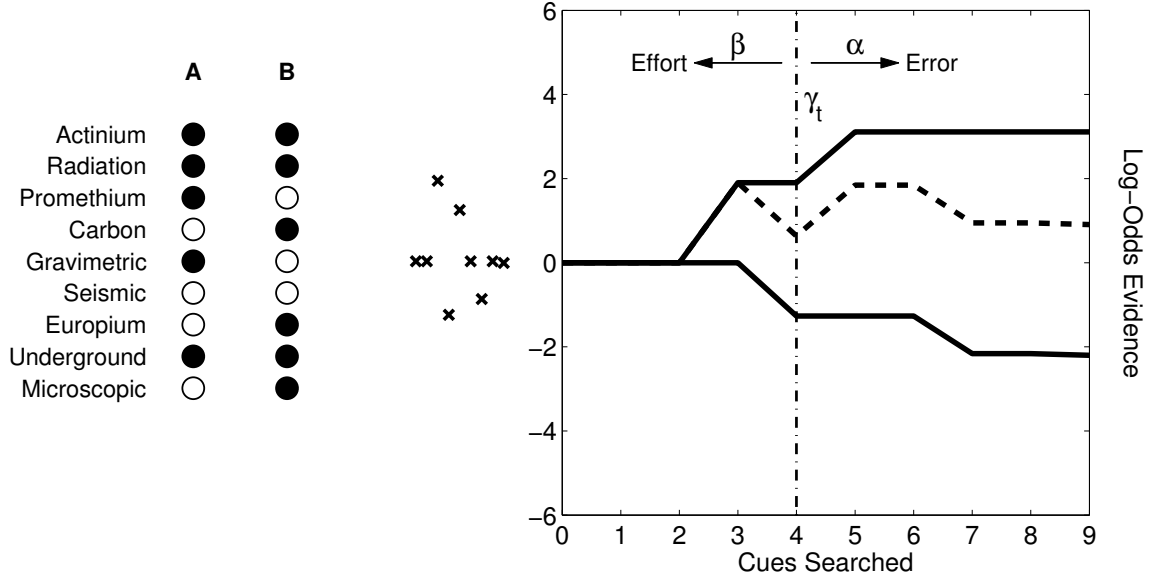


Figure 6. Overview of the error model, and error-effort model. Both models terminate search after searching γ_t proportion of cues on the t th trial. The error-based model increases the extent of search by a factor α if an incorrect decision is made. The error-effort model additionally decreases search by a factor β if a correct decision is made.

Formally, $\gamma_{t+1} \leftarrow (1 + \alpha) \gamma_t$. In the *error-effort* model, search is increased in the same way following an error, but search is also decreased by a learning parameter β following a correct decision. Formally, $\gamma_{t+1} \leftarrow (1 - \beta) \gamma_t$. The learning rates α , for both models, and β , for the error-effort model, are free parameters. For both models, the initial proportion of cues to be searched γ_1 is an additional free parameter.

Comparing Error-Based and Effort-Based Learning

Figure 7 presents the evaluation of the error model, and the error-effort model for two illustrative participants. Individual participants are organized in rows, and the models in columns. The first of these two participants was chosen because their behavior was highly representative of the regularities seen at the group level, as shown in Figure 3. This participant is considered in all of the model comparisons we report. The second participant in Figure 7 was chosen because they provide a good example of the important insights we observed based on examining all of the participants in all three experiments. In particular, they provide a clear example of how the simpler error model can be inadequate, and the more complicated error-effort model provides

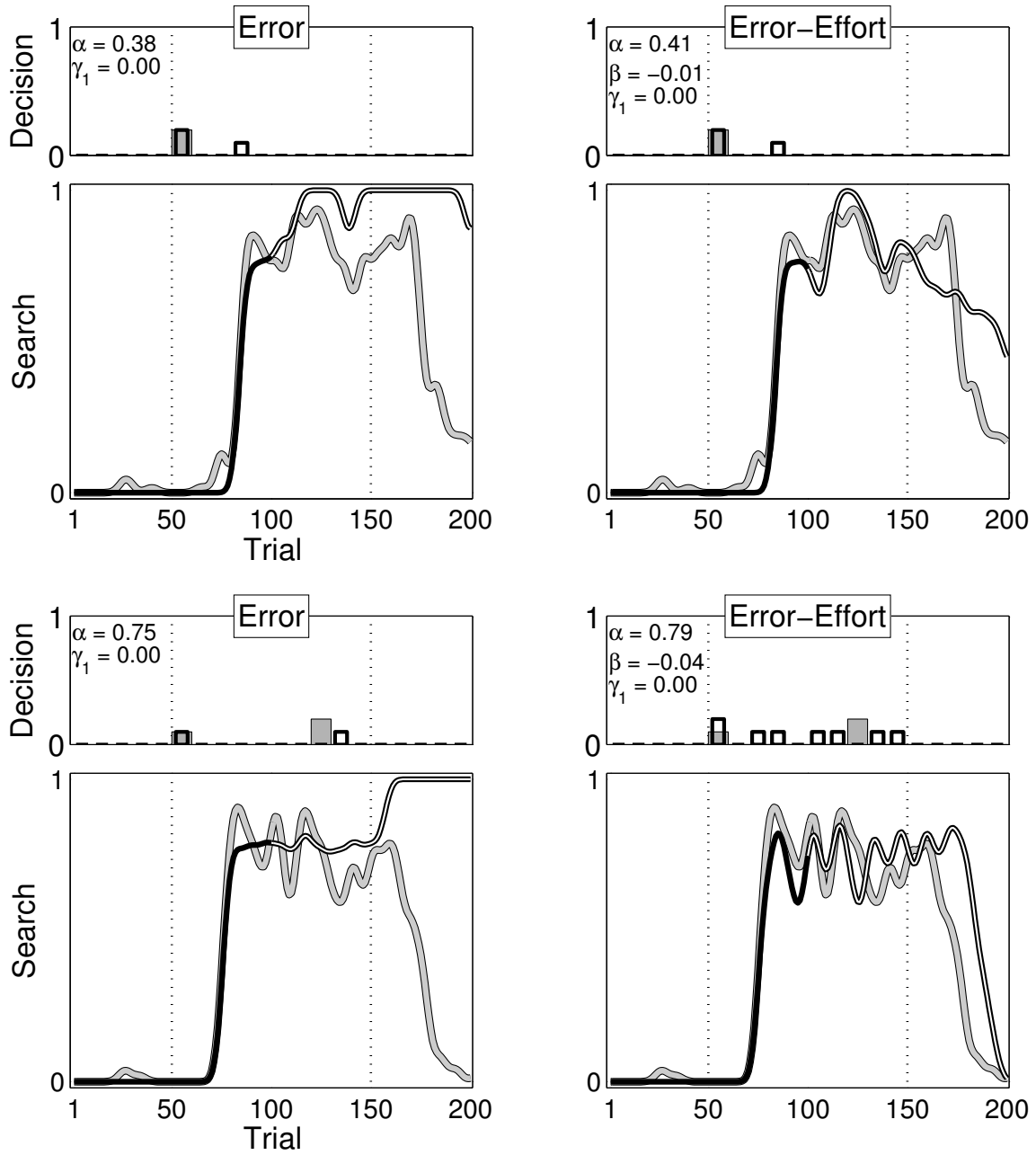


Figure 7. Evaluation of the error model, and the error-effort model, for two participants.

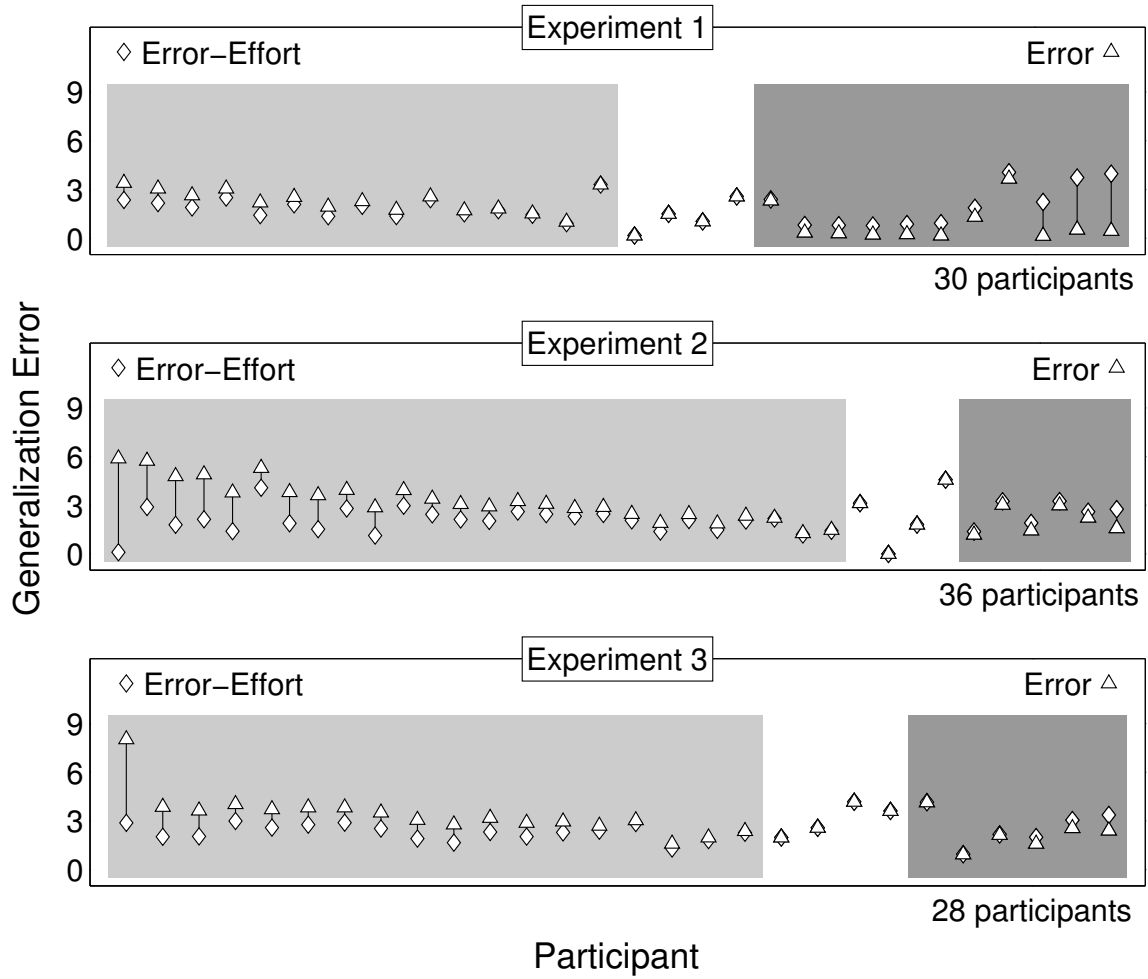


Figure 8. The generalization errors for all participants in all three experiments, comparing the error model and the error-effort model.

a better account of how people regulate search.

In Figure 7, both participants start with limited search, increase this search after making errors at the first environmental change, but then reduce their search, without making errors, after the second environmental change. For these sorts of participants, the error-effort model is able to capture the reduction of search, whereas the error model cannot. The summary of all of the empirical data in Figure 3 shows that the sort of decision and search behavior of the specific participants in Figure 7 was frequently observed. Thus, the apparent limitations of the error model are serious ones.

Figure 8 evaluates of the error and error-effort models on all of the participants in all of the experiments, using the root-mean-square generalization error measure. Each experiment corresponds to a panel, and the generalization error of each participant for each model is shown by markers. Within each experiment, participants are ordered from left to right in terms of the difference in the generalization errors for the two models. Those participants on the left, shaded in light gray, are those for whom the generalization error is lower for the error-effort model, while participants on the right, shaded in darker gray, are those for whom the generalization error is lower for the error model.³ Participants in the middle of panels, in the unshaded region, are those for whom the two models performed equally well.

Figure 8 shows that the majority of participants in all three experiments are better captured by the error-effort model. This is especially true in Experiments 2 and 3, where there are costs associated with search. In these circumstances, participants often reduced their search in the first and third blocks, and this behavior cannot be captured by the error model. It is clear that many participants adapt their search without being triggered by making errors.

Confidence-Based Learning

The first two models can be conceived as simple reinforcement learning heuristic accounts of searching and deciding. They adapt how many cues are searched, based upon feedback on the accuracy of decision making. A different approach to terminating search is to stop examining cues once sufficient evidence has been accumulated in favor of one or the other alternative. This is the approach taken by sequential sampling models of the time course of decision making, and the basis for the third and fourth models we consider.

Figure 9 shows a sequential sampling model approach to making a decision on a single trial. There are threshold levels of evidence δ_t^A and δ_t^B for the two alternatives. Following the standard assumptions of accumulator sequential sampling models (Vickers, 1970, 1979), the first alternative to accumulate a threshold level of evidence leads to search being terminated, and that alternative being chosen. In Figure 9, this means that Alternative A is chosen after the 5th cue is examined.

³The fact that sometimes the error model is preferred shows concretely that the generalization measure we use controls for model complexity. The error-effort model contains the error model as a special case, and could always fit any observed data at least as well. But, it is more complicated, which means it may not generalize as well when the error model is a better account of behavior.

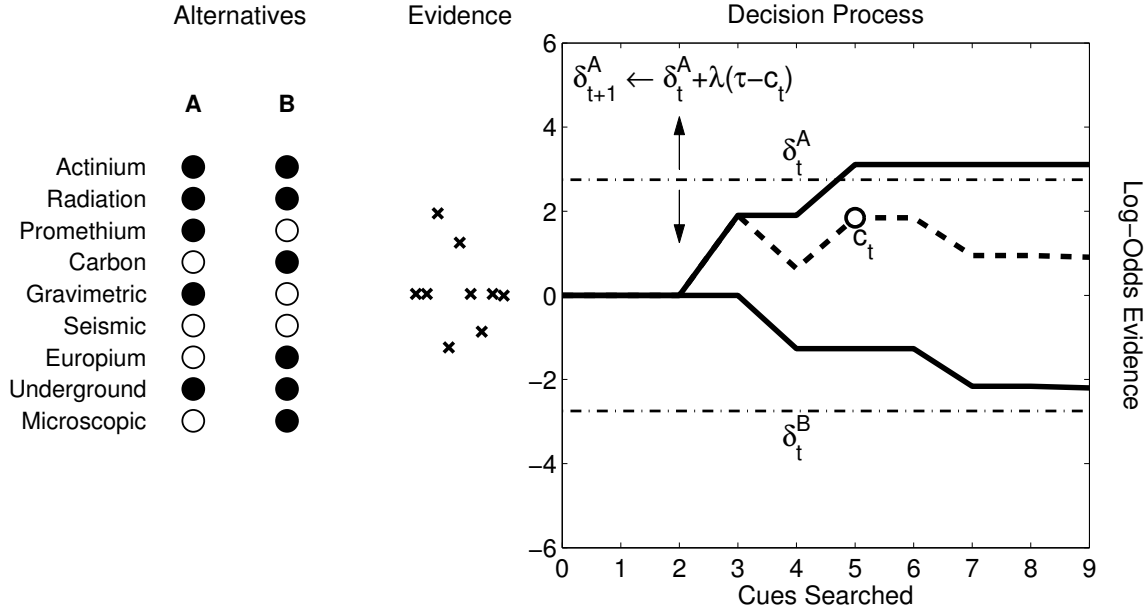


Figure 9. Overview of the confidence model. The model terminates search on the t th trial when one of the evidence thresholds δ_t^A or δ_t^B is exceeded. The threshold that was reached is then adjusted by comparing the achieved confidence c_t with the target level of confidence τ , using a learning rate λ .

The extent of search in the sequential sampling model is controlled by the thresholds. As has been emphasized previously (Lee & Cummins, 2004; Newell, 2005), small thresholds will lead to relatively few cues being searched, with one-reason decision making being a special case when the thresholds are near zero. The larger the thresholds, the more cues, in general, need to be searched to gather enough evidence to make a decision, and search becomes more exhaustive. This relationship between thresholds and search means that it is the adaptation of the boundaries that corresponds to how the extent of search is regulated.

In the model we consider, the adaptation of the thresholds is based on regulating confidence (Hausmann & Läge, 2008; Vickers, 1979). Accumulator sequential sampling models naturally provide a balance-of-evidence measure of confidence, given by the difference between the (magnitude of) the two tallies when a decision is made (Vickers, 2001), as represented by the broken line in Figure 9. Thus, when a decision is made after 5 cues, Alternative A has about 3 units of evidence, while Alternative B has about 1 unit of evidence. The difference between these two tallies thus gives 2

units of confidence on the log-odds scale for the chosen alternative. In the model we consider, this measure of confidence in each decision is compared to a target or desired level of confidence. Under-confidence corresponds to the case when the achieved confidence is less than the target level, and leads to the threshold that triggered the decision being increased. Over-confidence corresponds to the case when the achieved confidence is greater than the target level, and leads to the threshold being decreased.

Formally, the target level of confidence τ is a parameter of the model, as is a learning rate λ . If the balance-of-evidence measure of confidence in a decision is c_t , and is triggered by a threshold δ_t^A on the t th trial, the learning rule that updates the threshold for the next trial is $\delta_{t+1}^A \leftarrow \delta_t^A + \lambda(\tau - c_t)$. In this way, under-confidence leads to thresholds being raised, so that search is increased, while over-confidence leads to thresholds being lowered, so that search is decreased. The initial thresholds for both alternatives are assumed to be symmetric, and are given by a third model parameter δ_1 .

One way to think about the progression to the third model from the first two models is that it represents a change in basic modeling approach, from searching a proportion of cues and learning from accuracy, to searching until sufficient evidence is found and regulating based on confidence. Since both reinforcement learning and sequential sampling approaches are widely considered in modeling sequential decision-making, it is important to have examples from both. Another, complementary, perspective is that the sequential sampling model provides an alternative, and more parsimonious, account of how search can be increased and decreased than the reinforcement learning models. In the second error-effort model, for example, there are essentially separate processes, with separate parameters, for increasing and decreasing search. The use of confidence as a regulatory mechanism in the sequential sampling model naturally leads to both increases in search (when under-confident), and decreases in search (when over-confident) through the same psychological mechanism.

Figure 10 presents the evaluation of the error-effort model and the confidence model to two participants. The first is the same participant considered at the top of Figure 7, and their behavior is captured well by the confidence model, and reasonably well by the error-effort model. The behavior of the second participant in the final 100 trials, however, is better predicted by the confidence model. The error-effort model does not observe decreases in search following correct decisions during the training

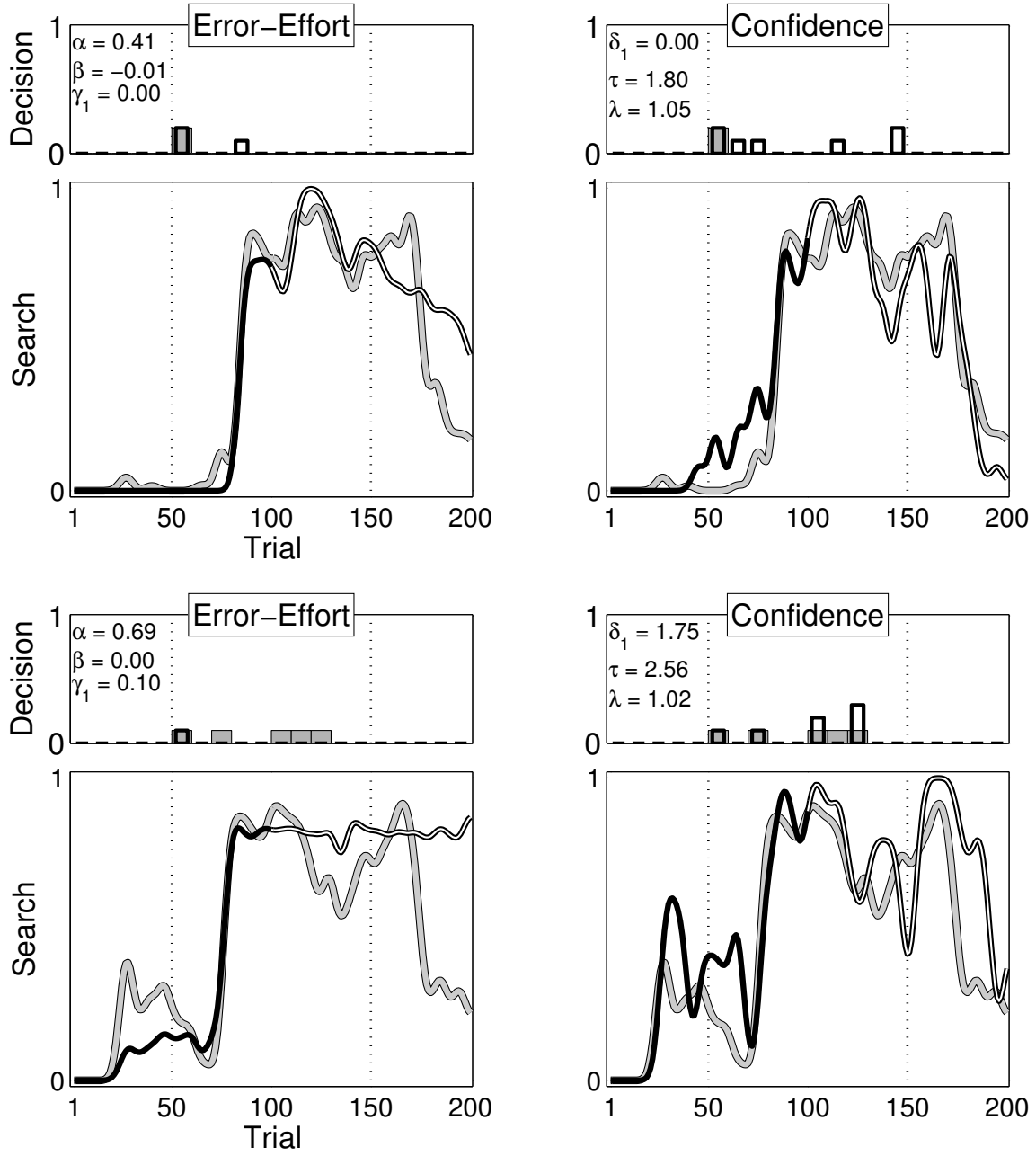


Figure 10. Evaluation of the error-effort model, and the confidence model, for two participants.

block, leading to a parameter estimate $\beta = 0$, and so incorrectly predicts a near-constant extent of search. The confidence model, however, tries more to maintain the target level of confidence $\tau = 2.56$ estimated from the training data in making its predictions. A natural property of the trials in the final block, where one-reason and exhaustive search both lead to the same answer, is that stronger evidence in favor of the correct alternative is found earlier in search. This property of the environment leads to over-confidence when using the large thresholds needed for extensive search in the second block, and so the confidence model predicts a lowering of thresholds, and subsequent reduction in the extent of search, as observed in the behavior of the participant.

Figure 11 summarizes the performance of the error-effort model and the confidence model on all of the participants in all the experimental conditions. Both of the models account of a significant proportion of participants in all of the experimental conditions. An examination of the difference in the generalization errors also suggests that, when there is a large difference in the predictive accuracy of the two models, it is almost always in favor of the confidence model. But, the overall conclusion is that both models provide useful accounts for many participants.

Hierarchical Confidence-Based Learning

The final model we consider is a hierarchical extension of the confidence model. The decision-process is the same, but the mechanism for adapting the threshold is extended. Rather than using a learning rule that adjusts the thresholds after every decision is made, internal accumulator processes are used to aggregate evidence of over- and under-confidence. As shown in Figure 12, there is now an internal accumulator for each of the two decision boundaries. These accumulators operate in the same way as the one that makes the overt decisions, except that they are driven by evidence of over- and under-confidence provided by the difference between the achieved and target level of confidence. When the over- or under-confidence tally reaches a critical level, a learning rule is applied to the decision threshold associated with the overt decision process. The model is naturally hierarchical, in the sense that the same sequential sampling process used to make overt decisions, as in the original confidence-based model, are now applied internally to make regulatory decisions about increasing and decreasing thresholds.

Figure 12 provides a concrete example that helps explain the formal notation of

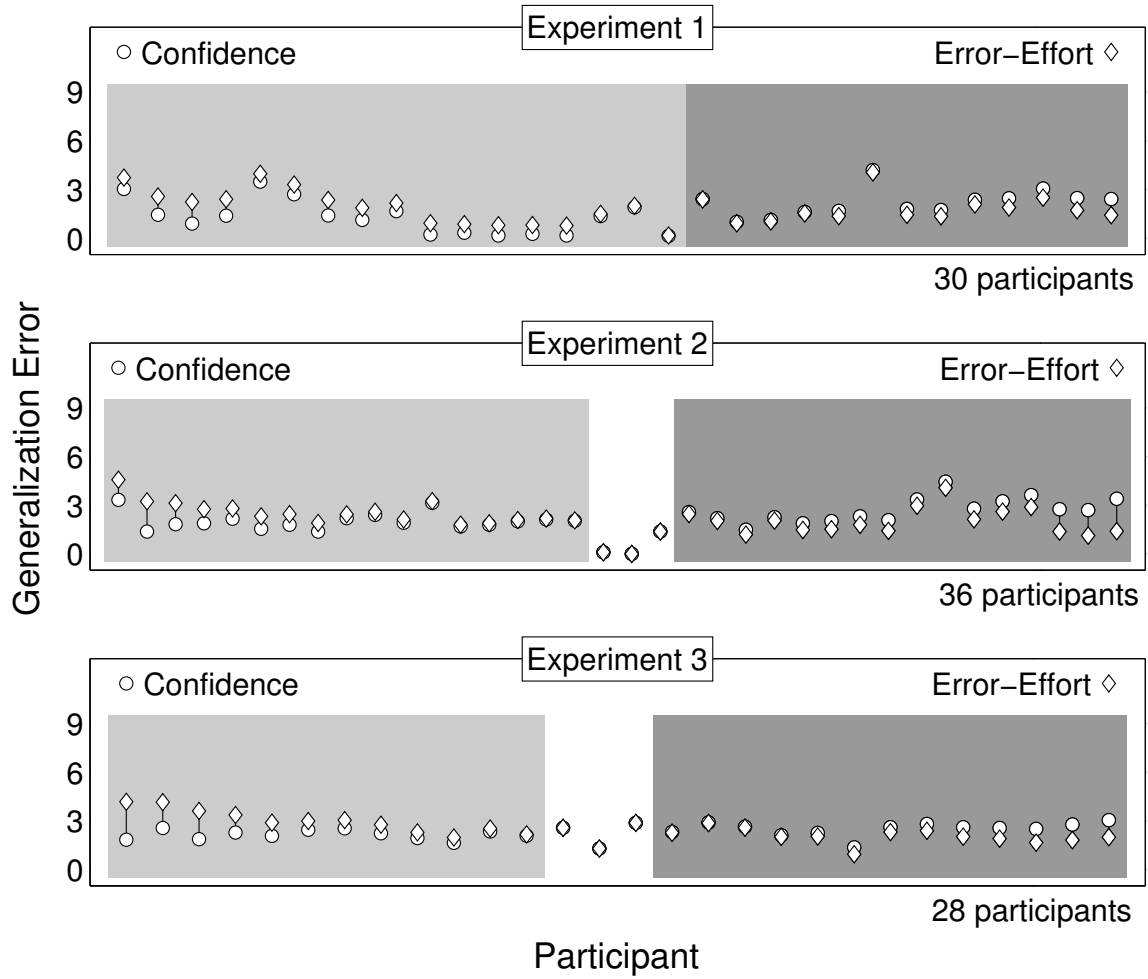


Figure 11. The generalization errors for all participants in all three experiments, comparing the error-effort model and the confidence model.

the hierarchical model. The over- and under-confidence tallies for the two alternatives are given by o_t^A , o_t^B , u_t^A and u_t^B . In Figure 12, Alternative B is chosen after 5 cues have been examined, but there is almost equal evidence for both alternatives at this point, and so the confidence in the decision c_t is near zero. For any reasonable target level of confidence τ significantly greater than zero, this will register as an under-confident decision, quantified by the difference $\tau - c_t$. This under-confidence will thus be added to the accumulator u^B , so that $u_{t+1}^B \leftarrow u_{t+1}^B + (\tau - c_t)$.

Figure 12 shows the history of the hierarchical model over a sequence of trials and, in particular, highlight a trial t_{adapt} around trial 5 where a threshold adjustment

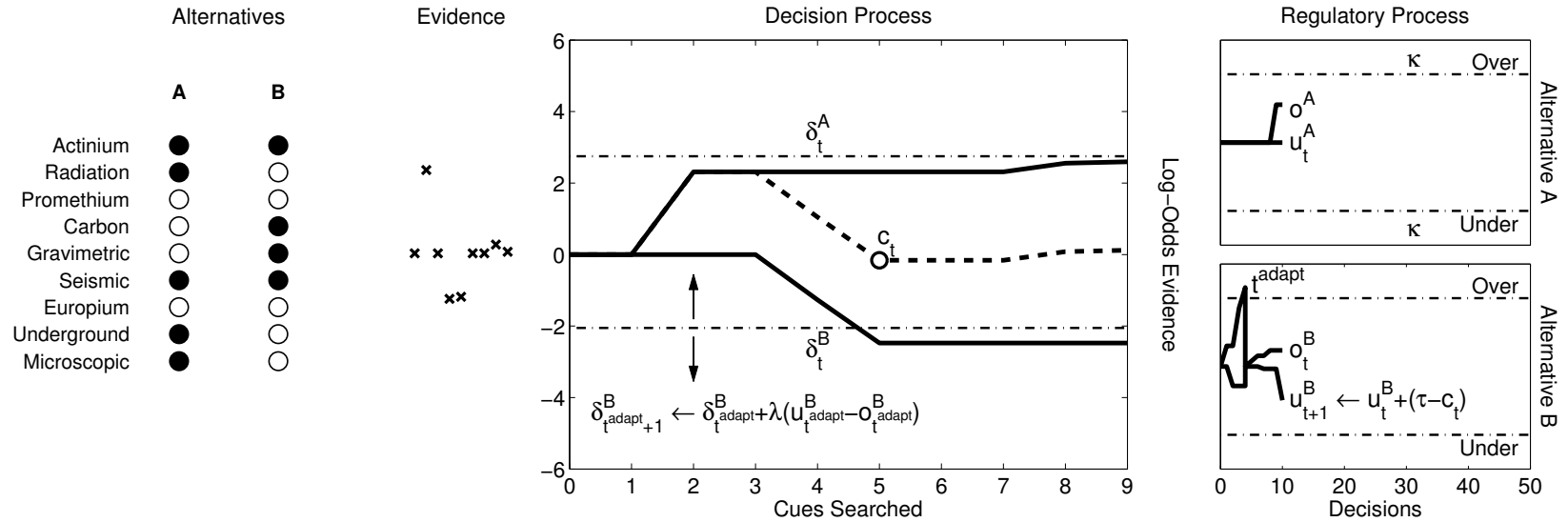


Figure 12. Overview of the hierarchical confidence model. The model terminates search on the t th trial when one of the decision thresholds δ_t^A or δ_t^B is exceeded. After each decision the achieved confidence c_t with the target level of confidence τ , and accumulated as evidence for over- or under-confidence in tallies associated with the threshold for the chosen alternative. When these tallies reach a critical level κ , the difference between the over- and under-confidence tallies is used to adjust the decision threshold using a learning rate λ .

was made. At this trial, the accumulator for over-confidence in Alternative B, o^B reached the critical threshold κ , resulting in the threshold for decision making δ^B being reduced. This previous adjustment is clear in Figure 12 from the smaller threshold for Alternative B than Alternative A, coming from the application of the learning rule $\delta_{\text{adapt}+1}^B \leftarrow \delta_{\text{adapt}}^B + \lambda(u_t^B - o_t^B)$ at the adaptation trial. The nature of this learning rule means that the extent of the adjustment in the threshold depends on the learning parameter λ , and on the difference between the evidence for under- and over-confidence at the time of adaptation. Note that, after adaptation, both the over- and under-confidence accumulators for Alternative B were reset, and began collecting evidence for the need for further adaptation afresh.

The hierarchical confidence model requires one additional parameter, κ , corresponding to the (fixed) thresholds for the internal regulatory accumulator processes. This parameter can be interpreted as measuring how delayed, lagged, or “twitchy” the adaptation of search is. If $\kappa = 0$ the hierarchical model will reduce to the confidence-based model, and adapt decision-making thresholds on every trial. Larger values of κ mean more evidence of over- or under-confidence is needed to trigger threshold adaptation, corresponding to greater delays or lags between adjustments.

The hierarchical confidence-based learning model is a natural extension of the SRA model developed by Vickers (1979), which is sometimes also called the PAGAN model in the literature (e.g., Vickers & Lee, 1998). While originally developed for simple perceptual decision-making tasks, the model is naturally adapted to the sorts of cue-based evidence involved in our task, and has previously been considered in the same form used here (Lee & Dry, 2006).

Figure 13 presents the evaluation of the confidence model and the hierarchical model to two participants. The first is again the same participant considered at the top of Figures 7 and 10, and their behavior is again captured well by both models. The second participant provides an example where the hierarchical model makes better predictions, because it predicts a delay between the change in the environment, and the adaptation of search. The confidence model predicts a decrease in the extent of search that is larger and more sudden than observed in the behavior of the participant. The hierarchical model, however, estimates a large critical threshold on adaptation $\kappa = 9.75$ from the way the participant changed their search during the first environmental change. Here, they made a sequence of errors, coming from too limited search, before adapting to more extensive search. A similar pattern is seen

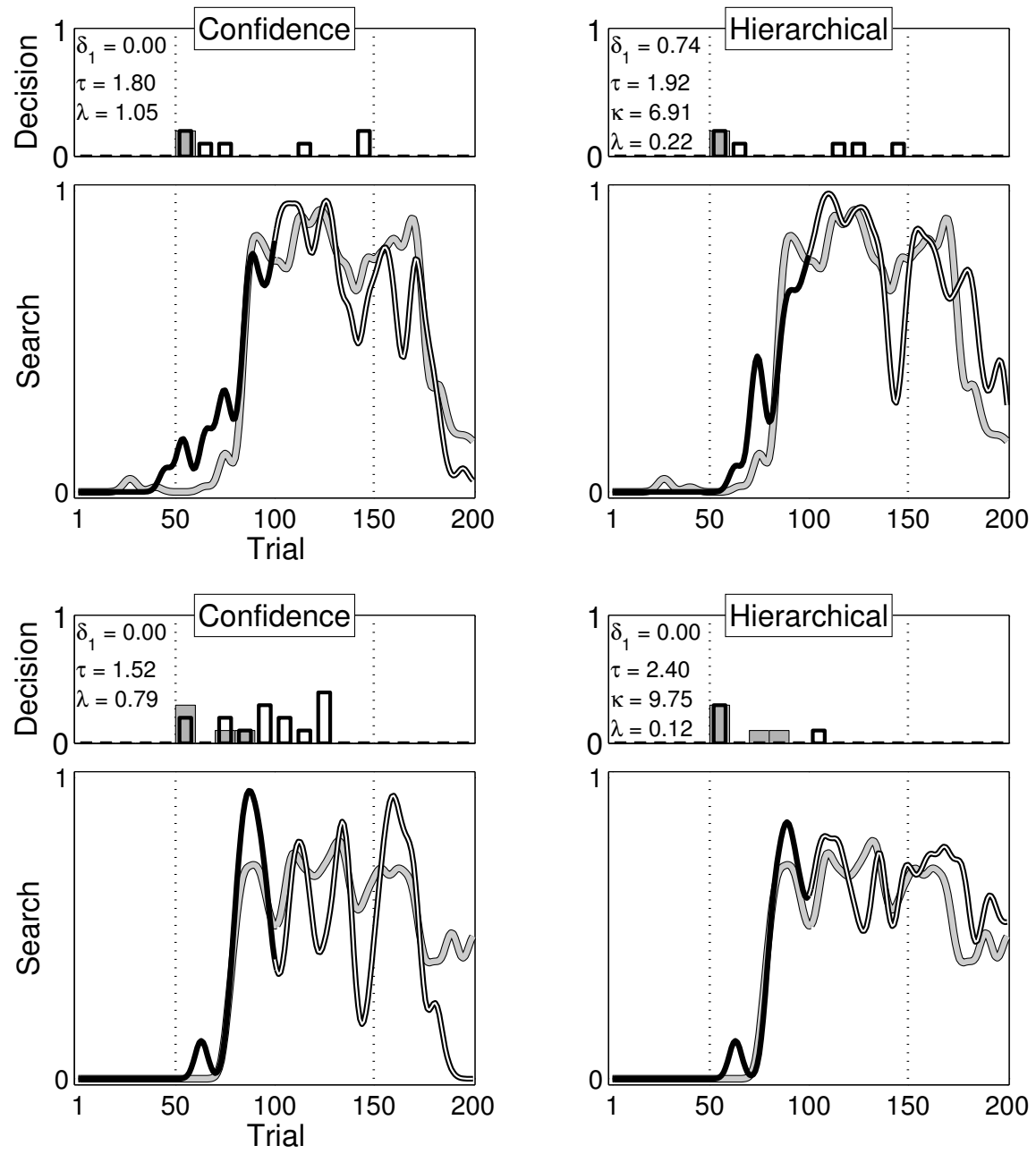


Figure 13. Evaluation of the confidence model, and the hierarchical confidence model, for two participants.

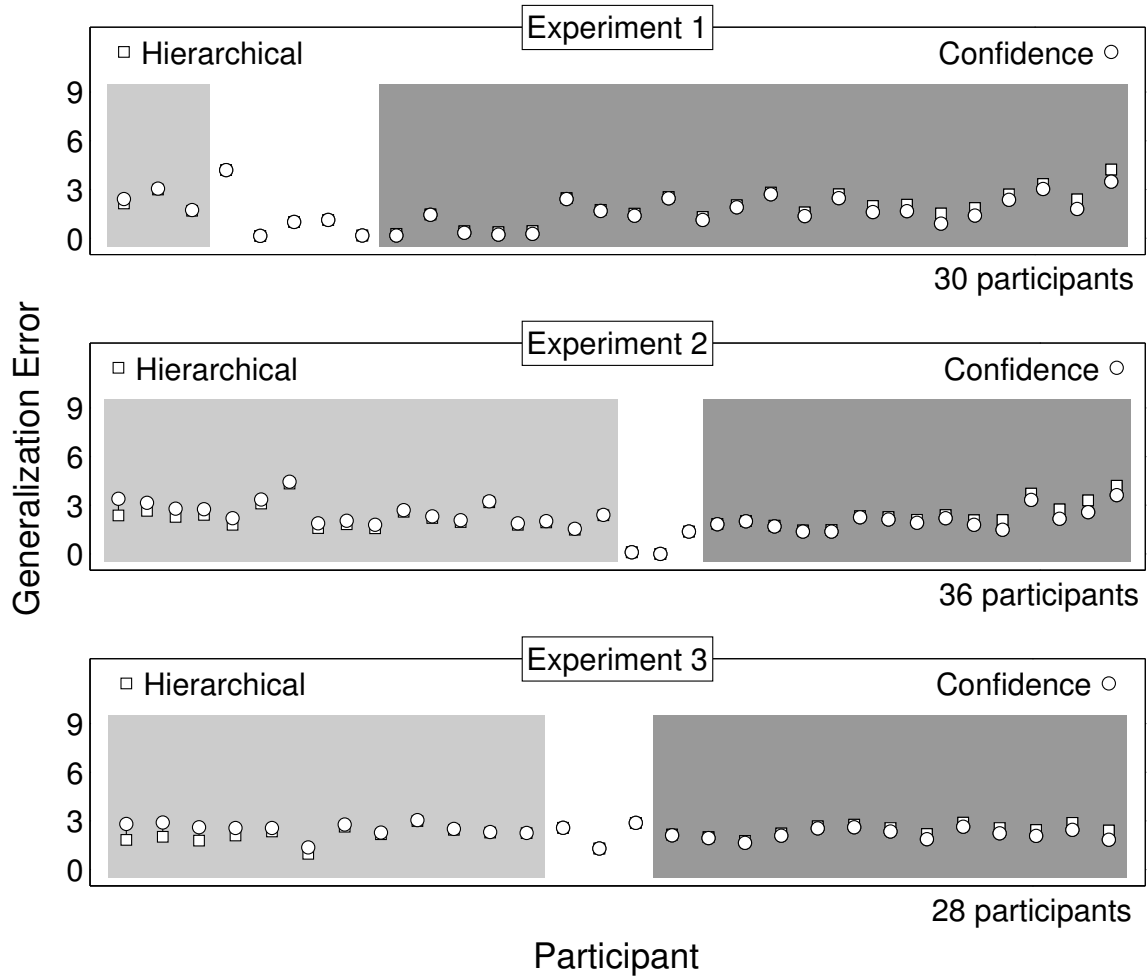


Figure 14. The generalization errors for all participants in all three experiments, comparing the confidence model and the hierarchical confidence model.

during the second environmental change, with a lack of immediate adjustment of the extent of search, and this is naturally predicted by the hierarchical model.

Figure 14 summarizes the performance of the confidence model and the hierarchical confidence model on all of the participants in all the experimental conditions. In the second and third experiments, where search is costly, there are many participants better accounted for by the hierarchical model, although the difference in generalization error between the models is small for almost all of the participants in all of the experiments.

Overall Modeling Results

The model comparisons presented in Figures 8, 11 and 14 evaluated pairs of logically related models. These comparisons are useful to understand when and why the additional elements of one of the models—such as including effort as well as error signals in learning, or making threshold adjustment hierarchical—are important for predicting how people search. It is also worthwhile, however, to compare all four models on all the participants in all of the experiments.

Figure 15 presents this comparison, in two ways. The left column of panels provide a relative measure of which model had the best generalization error for each participant, while the right column of panels show the distribution of generalization errors across participants. In the left column, each panel corresponds to an experiment, and the height of the bars for each model indicate how many participants in that experiment were best predicted by each model. In the right columns, the distributions of generalization errors are shown for each of the models

The results in the left column of Figure 15 suggest that all of the models are useful for explaining the behavior of at least some participants. The error-effort model and the confidence model consistently are the best accounts of the majority of participants. The error model is useful in the first experiment, but less so in the second and third experiments where search is more costly. In contrast, the hierarchical model is most useful when search is costly. The distribution of generalization errors in the right column of Figure 15 are consistent with these conclusions. Especially in the third experiment, there are some participants poorly described by the error and error-effort models, while the confidence and hierarchical confidence models always fare relatively well.

Discussion

Search is a fundamental cognitive ability, because it provides the mechanism for gathering information from the world or the mind on which decisions and actions can be based. Inherent in having a capability to search is having a capability to terminate search, and an ability to learn, adapt and regulate the termination of search. A search process that does not terminate is not useful, and termination strategies that are unable to learn from experience, or adapt to changes in the environment or the goals of a decision maker will usually be inefficient and limiting.

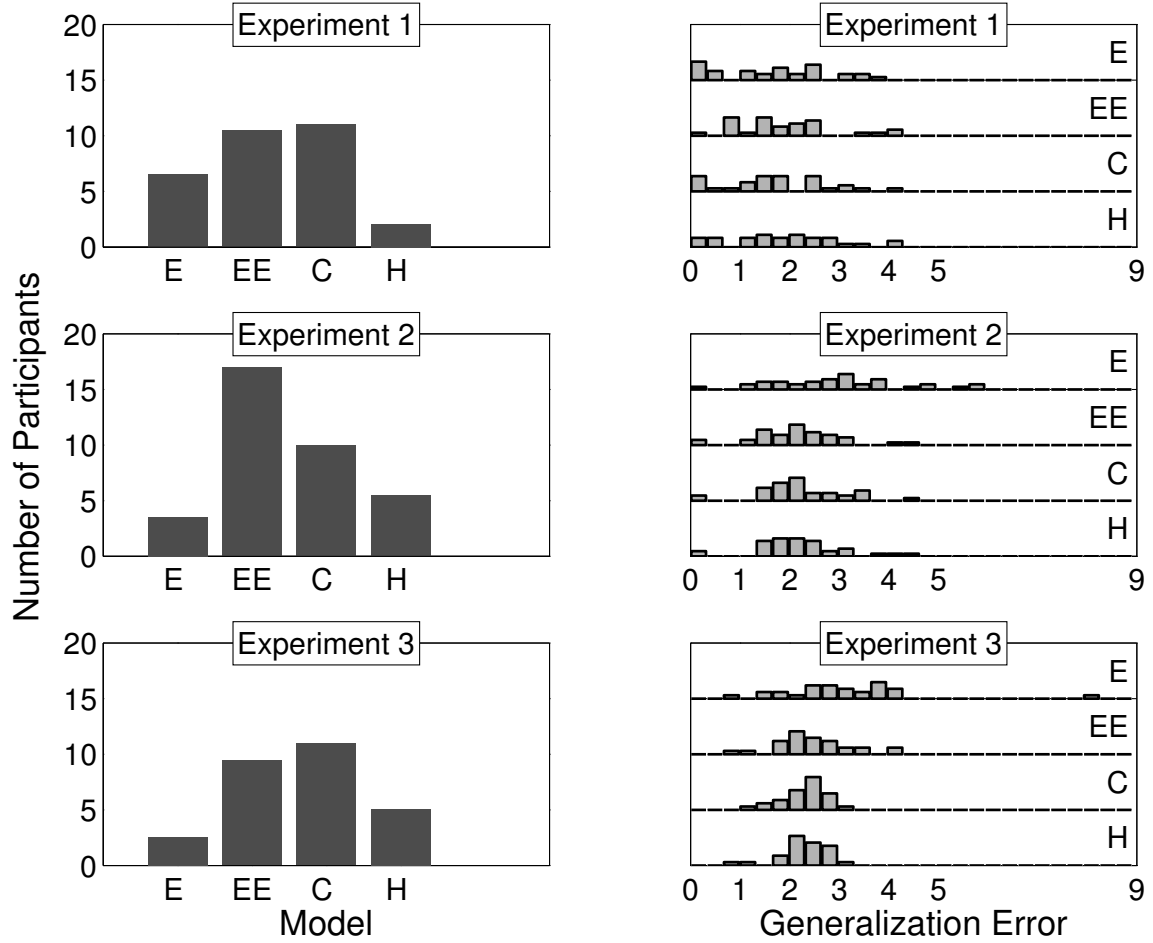


Figure 15. Overall evaluation of the four models for all participants in all three experiments. The panels in the left column show the numbers of participants best accounted for, in terms of lowest generalization error, for each model. The panels in the right column show the distributions of generalization errors. (E = error model, EE = error-effort model, C = confidence model, H = hierarchical confidence model).

In this study, we examined how people terminate their search in a dynamically changing environment. Unlike many previous studies of dynamic decision making, the environmental changes in our task were not always signaled by changes in accuracy. We also collected detailed information about the termination of search. By constraining the order in which cues could be examined, we gained the ability to have a fine-grained but simple measure of the extent of search. This allowed our analysis and modeling to move beyond the comparison of extreme heuristics that either rely on one-reason or exhaustive search. The best current models of the regulation of search in cue-based decision-making—like the SSL model (Rieskamp & Otto, 2006)—do not make detailed predictions about our behavioral data, because they assume either one discriminating cue or all cues are searched. The models we have developed and evaluated, in contrast, make predictions about the intermediate levels of search produced by people.

Many cognitive models that incorporate learning rely heavily on the availability of corrective feedback, the use of error-driven learning mechanisms. One general conclusion justified by our results is that error-based learning mechanisms are, by themselves, incomplete as accounts of how people regulate the extent of their search (Kheifets & Gallistel, 2012). We found clear empirical evidence that people change their search behavior without having made errors. Additional learning mechanisms—such as reducing effort when possible, or regulating confidence, as used in our models—are important elements in a full account of how people control their search behavior.

Of course, the set of four models we considered could and should be broadened greatly. Our models were chosen because they covered both the major reinforcement learning and sequential sampling approaches to modeling dynamic decision making. They also form a logical progression. The error-effort model extends the error model. The confidence model provides an alternative psychological mechanism to the error-effort model for increasing and decreasing search. The hierarchical confidence model extends the confidence model to allow for delayed and abrupt adaptation.

But other existing models and approaches are worth exploring. For example, we considered only accumulator (or race) processes for the tallying of evidence, as made clear in Figure 5. In this approach, the difference between tallies provides a measure of confidence. An alternative approach, widely used in the empirically successful class of drift-diffusion (random-walk) sequential sampling models (Bogacz

et al., 2006; Ratcliff, 1978; Ratcliff & McKoon, 2008; Ratcliff & Rouder, 1998), accumulates evidence as the difference between totals, and terminates search once this difference reaches a threshold. This approach necessitates alternative mechanisms for modeling confidence, which have recently been developed and evaluated (e.g., Pleskac & Busemeyer, 2010; Ratcliff & Starns, 2009). This theoretical development has not yet extended to the problem of adapting thresholds over sequences of trials—as required to make predictions in our task—but it is clear the building blocks needed for this development are in place. Thus, although we used the accumulator approach, because self-regulating models are well-developed (Vickers, 1979), and previously used for cue-based decision making (Lee & Dry, 2006), it is possible the diffusion approach could generate models worth evaluating.

The task we considered involved two (latent) sudden and large changes in the underlying decision environment. This is a useful design given our interest in the limits of accuracy as a signal for adaptation. It also provides an effective way to study to what extent the adaptation of search lags behind environmental change, and whether it involves sudden or gradual shifts. The ability of the hierarchical confidence model to provide the best predictions for some participants, particularly in direct competition with a reduced model that does not include delays, but is otherwise identical—suggests it is an important phenomenon. Delayed sequential effects in decision making are difficult to study and model (e.g., Gao, Wong-Lin, Holmes, Simen, & Cohen, 2009), and the extent of delay is sometimes captured in dynamic decision making models simply by the inclusion of free parameters (Brown & Steyvers, 2005). The hierarchical model, as with the original SRA model on which it is based, provides a viable alternative. In particular, this model has the attraction of modeling delay psychologically, rather than parameterizing it statistically, by making the elegant theoretical assumption that sequential sampling processes can embed hierarchically, and regulate one another based on confidence. Another class of models worth considering in this regard are extended reinforcement learning models known as actor-critic models (Barto, Sutton, & Anderson, 1983; Konda & Tsitsiklis, 2000). These models also have a hierarchical structure, with an overt decision making process being monitored and regulated by a latent control process.

Just as the models we considered are a small sample of possible relevant models, the types of dynamic environmental change we considered are only a small sample. As Speekenbrink and Shanks (2010) point out in their empirical and modeling inves-

tigation of cue-based categorization, real-world environments change in many ways. Previous work in modeling human decision making has used environments involving gradual drift (e.g., Otto et al., 2010; Rakow & Miller, 2009), patterns consistent with cyclical change (e.g., Yi, Steyvers, & Lee, 2009), step change jumps (e.g., Brown & Steyvers, 2005), and combinations of all of these sort of dynamics (e.g., Speekenbrink & Shanks, 2010). Our experimental design considered only large discrete changes of two types, encouraging a transition from limited to extensive search, and back again. It is not unreasonable to consider dynamics like this, involving a change to the original state, given the prevalence of cyclic patterns of change in real environments, and our design was appropriate for our interest in whether error signals are necessary for adaptation.

Obviously, however, there are combinatorially many sensible experimental designs that could be evaluated, in a large possible program of empirical work. One way to manage the scale of such an undertaking might be to characterize the dynamic patterns seen in real-world environments in an organizing taxonomy. Much as our use of the German cities data set gave the cues an environmental structure, it should be possible to match the dynamics of environmental change to real-world sequences. All of the models we have considered are immediately applicable to any sequence or structure of environmental change, and testing their ability to predict how people search and decide in those environments should provide stringent evaluations of the psychological assumptions and mechanisms on which the models are based.

We think the experimental and modeling evidence we have presented make a clear case for the insufficiency of error as a means of adapting search, for the sensitivity of people’s adaptation to the costs of search, for the potential role of confidence as a unifying regulatory variable, and for the usefulness of considering hierarchical and latent adaptation in human search. But there are many more environments, task conditions, and models that could and should be considered to understand how people decide how extensively they should search before making decisions, and how they adapt the extent of their search to changing goals and environments.

Acknowledgements

MDL and BRN acknowledge support from Australian Research Council Grant DP110100797, BRN acknowledge support from Australian Research Council Grant FT110100151, and MDL acknowledges support from the Air Force Office of Scientific

References

- Barto, A., Sutton, R., & Anderson, C. (1983). Neuron-like elements that can solve difficult learning control problems. *IEEE Transactions on Systems, Man and Cybernetics*, 13, 835–846.
- Bergert, F. B., & Nosofsky, R. M. (2007). A response-time approach to comparing generalized rational and take-the-best models of decision making. *Journal of Experimental Psychology: Learning, Memory & Cognition*, 33(1), 107–129.
- Bogacz, R., Brown, E., Moehlis, J., Holmes, P., & Cohen, J. D. (2006). The physics of optimal decision making: A formal analysis of models of performance in two-alternative forced choice tasks. *Psychological Review*, 113, 700–765.
- Brown, S., & Heathcote, A. (2008). The simplest complete model of choice response time: Linear ballistic accumulation. *Cognitive Psychology*, 57, 153–178.
- Brown, S., & Steyvers, M. (2005). The dynamics of experimentally induced criterion shifts. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, 31, 587–599.
- Busmeyer, J. R., & Myung, I.-J. (1992). An adaptive approach to human decision making: Learning theory, decision theory, and human performance. *Journal of Experimental Psychology: General*, 121, 177–194.
- Busmeyer, J. R., & Rapoport, A. (1988). Psychological models of deferred decision making. *Journal of Mathematical Psychology*, 32, 91–134.
- Busmeyer, J. R., & Wang, Y.-M. (2000). Model comparisons and model selections based on generalization criterion methodology. *Journal of Mathematical Psychology*, 44, 171–189.
- Cover, T. M., & Thomas, J. A. (2006). *Elements of information theory*. New York, NY: Wiley.
- Erev, I., & Barron, G. (2005). On adaptation, maximization, and reinforcement learning among cognitive strategies. *Psychological Review*, 112, 912–931.
- Gallistel, C. R., Fairhurst, S., & Balsam, P. (2004). The learning curve: Implications of a quantitative analysis. *Proceedings of the National Academy of Sciences*, 101, 13124–13131.

- Gao, J., Wong-Lin, K., Holmes, P., Simen, P., & Cohen, J. D. (2009). Sequential effects in two-choice reaction time tasks: Decomposition and synthesis of mechanisms. *Neural Computation*, *21*, 2407–2436.
- Gigerenzer, G., & Goldstein, D. G. (1996). Reasoning the fast and frugal way: Models of bounded rationality. *Psychological Review*, *103*(4), 650–669.
- Grünwald, P. (2000). Model selection based on minimum description length. *Journal of Mathematical Psychology*, *44*, 133–152.
- Hausmann, D., & Läge, D. (2008). Sequential evidence accumulation in decision making: The individual desired level of confidence can explain the extent of information acquisition. *Judgment and Decision Making*, *3*, 229–243.
- Kheifets, A., & Gallistel, C. (2012). Adapting without reinforcement. *Communicative & Integrative Biology*, *5*, 531–533.
- Kim, J. (1997). *Iterated grid search algorithm on unimodal criteria*. Unpublished doctoral dissertation, Virginia Polytechnic Institute and State University.
- Konda, V., & Tsitsiklis, J. (2000). Actor-critic algorithms. In *Siam journal on control and optimization* (pp. 1008–1014). MIT Press.
- Lee, M. D. (2004). An efficient method for the minimum description length evaluation of cognitivemodels. In K. Forbus, D. Gentner, & T. Regier (Eds.), *Proceedings of the 26th annual conference of the cognitive science society* (pp. 807–812). Mahwah, NJ: Erlbaum.
- Lee, M. D., & Cummins, T. D. R. (2004). Evidence accumulation in decision making: Unifying the “take the best” and “rational” models. *Psychonomic Bulletin & Review*, *11*, 343–352.
- Lee, M. D., & Dry, M. J. (2006). Decision-making and confidence given uncertain advice. *Cognitive Science*, *30*, 1081–1095.
- Lee, M. D., & Zhang, S. (2012). Evaluating the process coherence of take-the-best in structured environments. *Judgment and Decision Making*, *7*, 360–372.
- Link, S. W., & Heath, R. A. (1975). A sequential theory of psychological discrimination. *Psychometrika*, *40*, 77–105.
- Myung, I. J., Forster, M. R., & Browne, M. W. (2000). A special issue on model selection. *Journal of Mathematical Psychology*, *44*.

- Nassar, M. R., Rumsey, K. M., Wlson, R. C., Parikh, K., Heasly, B., & Gold, J. I. (2012). Rational regulation of learning dynamics by pupil-linked arousal systems. *Nature Neuroscience*, *15*, 1040–1046.
- Newell, B. R. (2005). Re-visions of rationality. *Trends in Cognitive Sciences*, *9*(1), 11–15.
- Newell, B. R., & Lee, M. D. (2009). Learning to adapt evidence thresholds in decision making. In N. Taatgen, H. van Rijn, J. Nerbonne, & L. Shönmaker (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 473–478). Austin, TX: Cognitive Science Society.
- Newell, B. R., & Lee, M. D. (2011). The right tool for the job? comparing evidence accumulation and a naive strategy selection model of decision making. *Journal of Behavioral Decision Making*, *24*, 456–481.
- Otto, A. R., Gureckis, T. M., Markman, A. B., & Love, B. C. (2010). Regulatory fit and systematic exploration in a dynamic decision-making environment. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *36*, 797–804.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1990). *The adaptive decision maker*. New York: Cambridge University Press.
- Pleskac, T. J., & Busemeyer, J. (2010). Two-stage dynamic signal detection: A theory of confidence, choice, and response time. *Psychological Review*, *117*, 864–901.
- Rakow, T., & Miller, K. (2009). Doomed to repeat the successes of the past: History is best forgotten for repeated choices with nonstationary payoffs. *Memory & Cognition*, *37*, 985–1000.
- Ratcliff, R. (1978). A theory of memory retrieval. *Psychological Review*, *85*, 59–108.
- Ratcliff, R., & McKoon, G. (2008). The diffusion decision model: Theory and data for two-choice decision tasks. *Neural Computation*, *20*, 873–922.
- Ratcliff, R., & Rouder, J. N. (1998). Modeling response times for two-choice decisions. *Psychological Science*, *9*, 347–356.
- Ratcliff, R., & Starns, J. (2009). Modeling confidence and response time in recognition memory. *Psychological Review*, *116*, 59–83.
- Rieskamp, J., & Otto, P. (2006). SSL: A theory of how people learn to select strategies. *Journal of Experimental Psychology: General*, *135*, 207–236.

- Simen, P., Cohen, J. D., & Holmes, P. (2006). Rapid decision threshold modulation by reward rate in a neural network. *Neural Networks*, 19, 1013-1026.
- Speekenbrink, M., & Shanks, D. R. (2010). Learning in a changing environment. *Journal of Experimental Psychology: General*, 139, 266–298.
- Thisted, R. A. (1988). *Elements of statistical computing: Numerical computation* (Vol. 1). Chapman & Hall/CRC.
- Todd, P. M., Hills, T. T., & Robbins, T. W. (Eds.). (2012). *Cognitive search: Evolution, algorithms, and the brain*. MIT Press.
- Vickers, D. (1970). Evidence for an accumulator model of psychophysical discrimination. *Ergonomics*, 13, 37–58.
- Vickers, D. (1979). *Decision processes in visual perception*. New York, NY: Academic Press.
- Vickers, D. (2001). Where does the balance of evidence lie with respect to confidence? In E. Sommerfeld, R. Kompass, & T. Lachmann (Eds.), *Proceedings of the seventeenth annual meeting of the international society for psychophysics* (pp. 148–153.). Lengerich: Pabst.
- Vickers, D., & Lee, M. D. (1998). Dynamic models of simple judgments: I. Properties of a self-regulating accumulator module. *Nonlinear Dynamics, Psychology, and Life Sciences*, 2, 169–194.
- Yi, S. K. M., Steyvers, M., & Lee, M. D. (2009). Modeling human performance in restless bandits using particle filters. *Journal of Problem Solving*, 2, 33–53.