



Suboptimal tradeoffs in information seeking

Wai-Tat Fu^{a,*}, Wayne D. Gray^b

^a *Carnegie Mellon University, USA*

^b *Rensselaer Polytechnic Institute, USA*

Accepted 12 August 2005

Available online 13 December 2005

Abstract

Explicit information-seeking actions are needed to evaluate alternative actions in problem-solving tasks. Information-seeking costs are often traded off against the utility of information. We present three experiments that show how subjects adapt to the cost and information structures of environments in a map-navigation task. We found that subjects often stabilize at suboptimal levels of performance. A Bayesian satisficing model (BSM) is proposed and implemented in the ACT-R architecture to predict information-seeking behavior. The BSM uses a local decision rule and a global Bayesian learning mechanism to decide when to stop seeking information. The model matched the human data well, suggesting that adaptation to cost and information structures can be achieved by a simple local decision rule. The local decision rule, however, often limits exploration of the environment and leads to suboptimal performance. We propose that suboptimal performance is an emergent property of the dynamic interactions between cognition and the environment.

© 2005 Elsevier Inc. All rights reserved.

Keywords: Problem solving; Adaptive search; Information seeking; Suboptimal tradeoffs; Satisficing, Bayesian learning; ACT-R; Cognitive modeling; Sequential decision making

1. Introduction

Consider a person deciding which route to take to go from one city to another. The person may be seeking information about traffic conditions of various routes. Since each information-seeking action takes time, tradeoffs are often made as exhaustive information seeking may be too costly to be justified. The person may decide to stop seeking

* Corresponding author.

E-mail address: wfu@cmu.edu (W.-T. Fu).

information when a reasonably good route is found, and he believes that the utility (i.e., usefulness) of checking the traffic condition of the next route is unlikely to justify the cost of the next information-seeking action (Stigler, 1961). The problem of when to stop seeking information is quite common. For example, doctors need to know when to stop performing diagnostic tests to decide on a treatment, as testing cannot be carried out indefinitely. Similarly, chess players need to decide when to stop evaluating alternative moves as it is almost impossible to evaluate all possible moves. These are but instances of a general decision problem, in which the decision-maker has to decide when to stop seeking information by trading off the cost of information-seeking actions against the utility of the information being sought. In the domain of Artificial Intelligence, numerous algorithms have been proposed that optimize the solution path by minimizing search costs. Such algorithms, however, often require extensive computations that make them psychological implausible. The goal of this article is to study the degree to which people are able to adapt their information seeking to the cost-structure of their task environment; and to understand the circumstances under which this adaptation may plateau at suboptimal levels.

2. Information seeking as the interface between cognition and the environment

Most distal properties of the environment (such as the utility of information) cannot be directly perceived (Brunswik, 1952; Fiedler, 2000). Rather, these distal properties have to be inferred from proximal information obtained from dynamic interactions of the person and the environment. Information-seeking actions can be considered a major form of these cognitive-environment interactions, as the purpose of these information-seeking actions is to obtain samples from the environment so that certain distal properties of the environment can be inferred. Actions are then selected based on the person's cognitive representation of these distal properties. In the route-finding example above, each information-seeking action allows the person to update his or her knowledge of the traffic conditions of possible routes, which allows the person to make a better decision on which route to take. The amount of information obtained therefore indirectly influences performance. Ideally, perfect performance can be attained when the person has complete knowledge of the environment. But in most situations, the information-seeking costs prevent exhaustive search of information. In this article, we study the adaptiveness of this kind of information-seeking behavior. In particular, we focus on how people are able to tradeoff the cost against the utility of information adaptively. The tradeoff may not be fully under the person's cognitive control, and our goal is to characterize the cognitive processes underlying the tradeoff. To preview our conclusions, we find that suboptimal tradeoffs are often a natural consequence of the dynamic interactions between cognition and the environment.

3. Tradeoffs between information-seeking costs and utility of information

In finding a fast route to another city, both costs and utility of information can be cast in the dimension of time: information-seeking costs will be high when they take more time, and information utility will be high when the information is able to lead to a route that takes less time. One can then measure performance by the total costs required to finish the task (the lower the cost the better the performance). The total costs can be calculated as

Total costs = Execution costs + Information-Seeking costs.

To illustrate the tradeoff between execution and information-seeking costs, consider the three episodes in Fig. 1. The lengths of the arrows represent the amount of time to finish the task. Episode B shows how the task is finished without information seeking, hence the information-seeking cost is zero and the total costs equal the execution costs. In episode A, information was sought and the time spent in executing the information-seeking actions is the information-seeking cost. Based on the information obtained, the execution costs are much lower in episode A than in episode B. The difference in the execution costs between episode A and episode B can be used as a measure of the utility of the information sought in episode A. In this case, the utility of information is higher than the information-seeking costs, and as a result, the total costs in episode A are lower. In other words, the information obtained has led to performance improvement in episode A. On the other hand, it is possible that additional information seeking does not always lead to the same reduction in execution costs, as, for example, information obtained could be redundant. As a result, too much information seeking may not be justified by the utility of information obtained. This is the case in episode C, where the total costs are larger than that in both episode A and episode B. The three episodes in Fig. 1 illustrate the point that optimal performance depends critically on the decision on when to stop seeking information.

If we assume the simplistic view that each information-seeking action incurs a constant cost, and each piece of information obtained reduces the execution cost required to finish a task, we can calculate the relationship among the number of information-seeking actions (n), the information-seeking costs ($n * C$), the execution costs ($f(n)$), and the total costs ($f(n) + n * C$). As shown in Fig. 2, the positively sloped straight line represents the increase of information-seeking costs with the number of information-seeking actions. The curve $f(n)$ represents the execution costs as a function of the number of information-seeking actions. The function $f(n)$ has the characteristic of diminishing return, so that more information-seeking actions will lead to smaller savings per information-seeking action. The U-shape curve is the total costs, which equals the sum of information-seeking costs and execution costs. The U-shape curve implies that optimal performance is associated with a moderate number of information-seeking actions. In other words, too much or too little information seeking may lead to suboptimal performance. How are people able to adapt to different numbers of information-seeking actions in different environments?

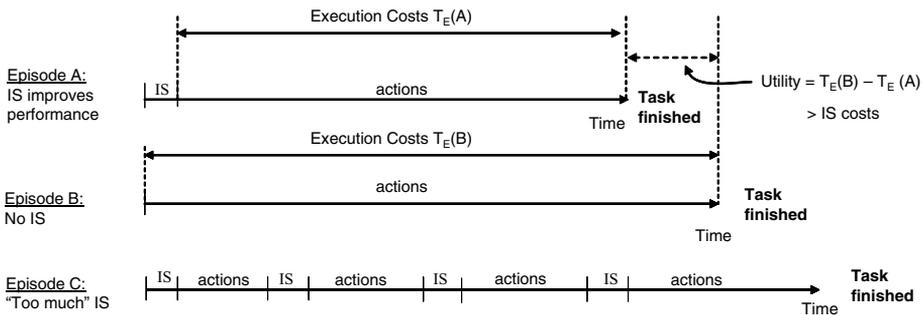


Fig. 1. Three problem-solving episodes showing how cost and utility of information are measured in the same dimension. The current definitions of cost and utility of information are measured by time. (IS = information-seeking actions.)

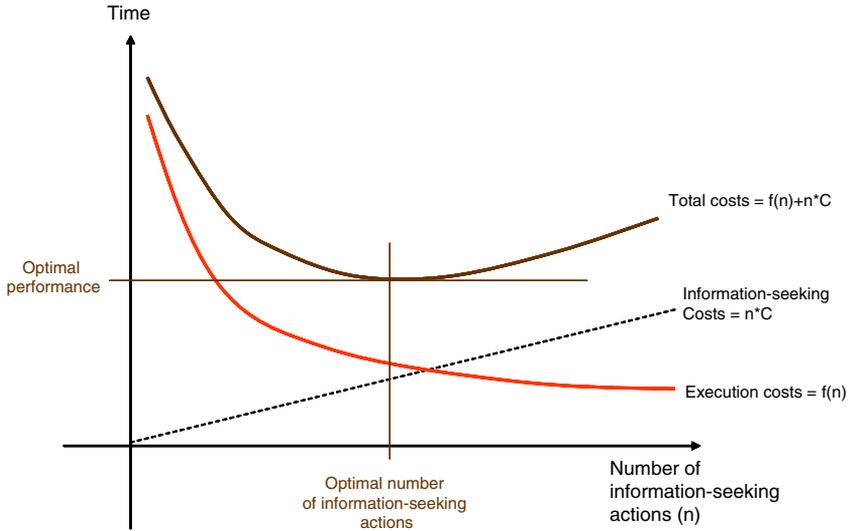


Fig. 2. The relationship between information-seeking costs, execution costs, and total time spent to finish the task. Execution costs (i.e., time to finish the task excluding all information-seeking costs) is represented as $f(n)$, a decreasing function of the number of information-seeking actions (n). Assuming a constant cost C for each information-seeking action, total information-seeking costs increase linearly with the number of information-seeking actions (i.e., total costs = $n * C$). The total time to finish the task is the sum of information-seeking costs and execution costs. Optimal performance, defined as the minimum total time to finish the task, is thus associated with a particular number of information-seeking actions. From the figure, it is clear that the optimal number of information-seeking actions depends on both C and $f(n)$.

What are the cognitive processes underlying the adaptation? Before we present our approach to answer these questions, we will first briefly review the relevant research on information-seeking behavior.

3.1. Sensitivity to information-seeking costs

Information-seeking costs have been manipulated to study problem-solving processes (Ballard, Hayhoe, & Pelz, 1995; Fu & Gray, 2000; Gray & Fu, 2004) and decision-making processes (Beach & Mitchell, 1978; Payne, Bettman, & Johnson, 1993). In general, behavior is found to be sensitive to even small changes in information-seeking costs. For example, in the experiments by Ballard et al., subjects were asked to copy some configurations of colored blocks presented on one computer screen to another computer screen. Ballard et al. found fewer information-seeking actions when the cost of information seeking increased from an eye-movement to a head-movement. Fu and Gray replicated and extended the findings by manipulating the information-seeking costs by either a single key press on the keyboard (low cost), a single mouse movement (medium cost), or a single mouse movement and a 1-s lockout time (high cost). The results suggest that people are sensitive to increases in information-seeking costs as low as 50 ms and as great as 1100 ms: people reduce the number of information-seeking actions when their costs are increased.

Although previous research has shown that people are sensitive to information-seeking costs, two questions remain unanswered by these studies. First, since only information-

seeking costs were manipulated in these studies, the results were not sufficient to distinguish between the case where people simply minimized the information-seeking costs and the case where people were trading off information-seeking costs against the utility of information. To make this distinction, one needs to explicitly manipulate the utility of information. If people are simply minimizing information-seeking costs, the level of information seeking should not change when the utility of information is changed. On the other hand, if people are engaged in tradeoffs, we should see a continuum in the number of information seeking actions when the utility of information is increased or decreased for a given level of information-seeking cost. Second, it is still not clear whether information-seeking behavior is truly *adaptive* in the sense that the increase or decrease of information-seeking actions will lead to better global performance (i.e., reduction in total costs in the example shown in Fig. 2). In other words, it is not clear whether people are able to adapt to the optimal number of information seeking actions in different environments.

3.2. *Bounded rationality and the local adaptive process*

The argument for the adaptiveness of behavior is most often discussed in the context of human rationality. In its strong form, behavior is assumed to be normatively rational. Normative rationality implies that one should always adopt the level of information seeking that leads to globally optimal performance. The normative standard, however, often requires the assumption of perfect information and infinite computational resources. A more successful hypothesis about adaptive behavior is that it exhibits *bounded rationality* or makes choices based on *satisficing* (Simon, 1956, 1996). Following Simon, many researchers (Anderson, 1990; Gigerenzer & Todd, 1999; Oaksford & Chater, 1994, 1996) argue that the rationality of the decisions made by humans and animals about their world are always within the bounds of limited time, knowledge, and computational power. In the case of a person exploring an unfamiliar task environment, the best that one can do is to make decisions based on local information obtained from direct experience with the environment. The concept of satisficing refers to the process that searches for and evaluates different options until an option is found that suffices to satisfy the goal. Satisficing can therefore be characterized as a local adaptive process that relies on limited knowledge and does not require exhaustive search of available options. However, the local process does not necessarily lead to the globally optimal solution. Rather, it describes a possible advantageous adaptation process that may or may not lead to the global optimum, depending on the particular characteristics of the environment.

There has been a long history of empirical studies on the use of local information as a variable that controls behavior (Herrnstein, 1991; Herrnstein & Prelec, 1991; Mazur, 1981; Vaughan, 1981, 1985). Impressive evidence of the use of local information in the underlying mechanism for tradeoffs between two options has been provided by Vaughan (1981). In his experiment, pigeons were first trained to distribute their pecks on the two available buttons so that the peck ratio roughly matched the reinforcement ratio on the two buttons. After performance stabilized, the reinforcement rates of the two buttons were changed. Pigeons, as expected, changed the distribution of pecks on the two buttons. The cleverness of the experimental design lies in the different predictions derived from whether the pigeons used local or global reinforcement rates to adapt to the new environment. Specifically, if global reinforcement rates were used, pigeons would stabilize in a region where

global reinforcement rates were highest. Results showed that pigeons did not stabilize on regions with the highest global reinforcement rates, but on regions with the highest *local* reinforcement rates (i.e., slightly increasing or decreasing the choice of a particular button would decrease the reinforcement rates). The results provide strong evidence that learning is a local adaptive process that relies on limited information obtained from direct experience, which does not necessarily lead to the globally optimal level of performance.

4. Plan of the article

The goal of this article is to study the degree to which information-seeking behavior is adaptive and how well people are able to attain the optimal level of information-seeking in environments with different cost and information structures. In the next section, we present our approach to studying how people adapt to different environments with different information structures. Specifically, we present a model that integrates the learning of information utility with the decision on when to stop seeking information. The model generates a set of predictions that guide the designs of three experiments. We implemented the model in the ACT-R architecture so that the predictions can be directly tested against the empirical data.

Fig. 3 shows the structure of the article and how the model and the three experiments are related. The central theme of this paper rests on the Bayesian satisficing model (BSM), which will be discussed in the next section. The behavior of the model leads to three

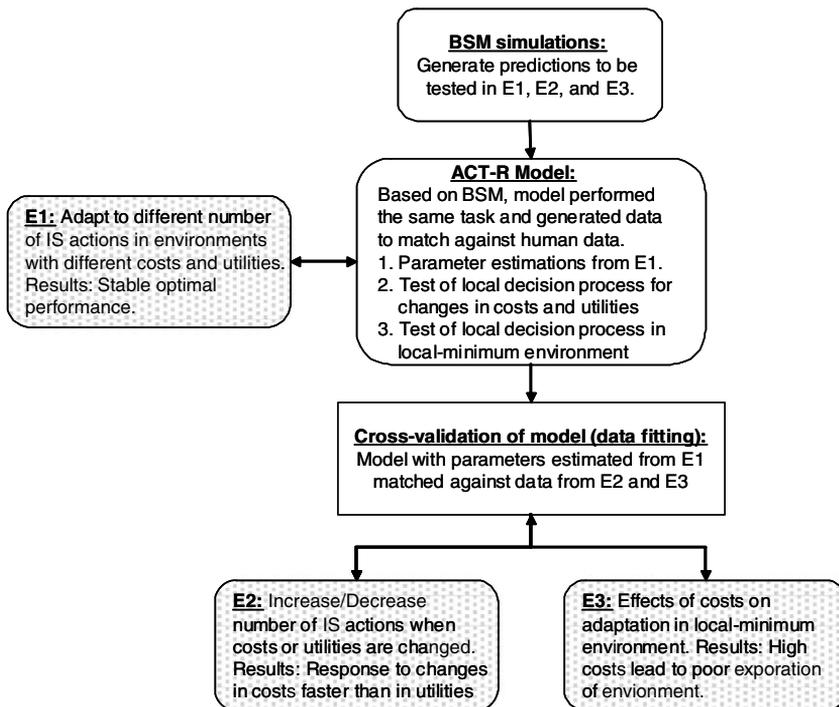


Fig. 3. The structure of the article relating the Bayesian satisficing model (BSM), the ACT-R model, the three experiments, and the approach to validate the model.

hypotheses to be tested in each of the three experiments. In E1, we test whether subjects will adapt their information-seeking behavior in environments with different costs and utilities. In E2, we directly test whether subjects respond faster to local than to global feedback information. In E3, we test whether suboptimal performance is related to the local nature of the adaptive process. An ACT-R model, constructed based on the BSM, performed the same task as the subjects and generated data that allow us to match the behavior of the model to human performance. The ACT-R model has a number of free parameters, and the values of these parameters were set to best fit the E1 human data. The same model, with its parameters set to the E1 data, performed the tasks in E2 and E3. We will elaborate on the validity of this model fitting process in Section 11.

5. The Bayesian satisficing model

As discussed in the last section, satisficing is a process based on bounded rationality that does not require perfect information and unlimited cognitive resources. Satisficing does not look for the globally optimal solution, but searches for solutions based on simple heuristics (e.g., hill-climbing) utilizing imperfect local information as included constraints. The overall BSM replaces complex computations by a simple effective mechanism that often performs reasonably well by exploiting the general structure of the environment. In this section, we will first describe the BSM; and then we will show how the BSM behaves in different environments. Finally, we will generate predictions on when the BSM is likely to settle at suboptimal levels of performance.

5.1. Structure of the BSM

As illustrated in Fig. 2, to model how well people are able to adapt to the optimal level of information-seeking requires the specification of two processes: (1) the estimation of the function $f(n)$, and (2) the decision on when to stop seeking information. The first process requires the understanding of how people estimate information utility based on experience. The second process requires the understanding of how information-seeking behavior is sensitive to the cost and utility of information. These two processes are integrated into the BSM (Fig. 4; details of the model are presented in Appendix A). In the global learning

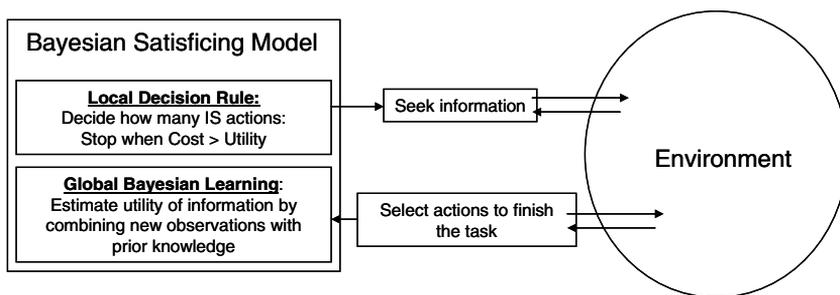


Fig. 4. The structure of the Bayesian satisficing model (BSM): (1) the local decision rule decides when to stop seeking information, and (2) the global Bayesian learning mechanism updates the knowledge of the environment (utility of information) after actions are executed. Information-seeking actions act as the interface between the model and the environment.

process, the model assumes that execution costs can be described by a diminishing-return function of the number of information-seeking actions (i.e., $f(n)$ in Fig. 2). A local decision rule is used to decide when to stop seeking information (see Fig. 5) based on the existing estimation of $f(n)$. Specifically, when the estimated utility of the information (i.e., $f(N) - f(N + N')$) is lower than the information-seeking cost, the model will stop seeking information. This local decision rule decides how many information-seeking actions will be done. The time spent to finish the task given the chosen number of information-seeking actions is then used to update the existing knowledge of $f(n)$ based on Bayes' theorem (see Appendix A for details). Since we are interested in the general behavior of the model, we have simply chosen a set of parameters and environments to study how they affect performance. The prior distribution of the mean execution costs (B) to solve a problem is assumed to be a gamma distribution, and the likelihood function is assumed to be an exponential distribution. Using Bayes' theorem, the model updates its own representation of the relationship between information-seeking costs and utility of information. As we show below, the simulations of BSM will be useful for generating specific predictions about behavior of the subjects.

5.2. Behavior of the BSM in diminishing-return environments

Two categories of diminishing-return environments were created to verify the behavior of the model. Each category assumed a different diminishing-return relationship between the amount of information-seeking and the amount of effort required to solve the problem. For the exponential environment the relationship is expressed by $C = C_{np} * e^{-k*n}$, where C is the amount of effort required, n is the amount of information-seeking, C_{np} is the amount of effort required when no information-seeking is done, and k is the parameter controlling the rate of change. In the power environment, the relationship is expressed by $C = C_{np} * n^{-k}$. In general, the higher the value of k , the lower will be the optimal amount of information-seeking for a given value of information-seeking cost.

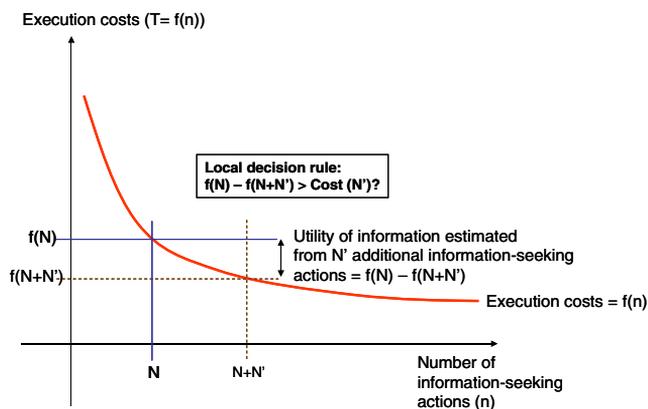


Fig. 5. The local decision rule that decides when to stop seeking information. In the figure, the relationship between the number of information-seeking actions (n) and the execution costs (i.e., time to finish the task excluding all time spent on seeking information) is represented by the function $f(n)$. The utility of information associated with an additional N' number of information-seeking actions is $f(N) - f(N + N')$. The local decision rule is to stop seeking information when n is such that the marginal utility of information is smaller than the cost.

Fig. 6 shows the simulation results of the BSM with different parameters values for the prior distribution in different environments. The figure also shows the optimal number of information-seeking actions based on the local decision rule as shown in Fig. 5. The figure shows that when the BSM first approaches the problem, the number of information-seeking actions chosen by the model is close to the mean (calculated as the $a * b$) of the prior distribution. With sufficient observations from the environment, the number of information-seeking actions either decreases or increases and eventually converges to the optimal level specific to characteristics of each environment (the dashed line in Fig. 6). The convergence to the optimal levels in both the exponential

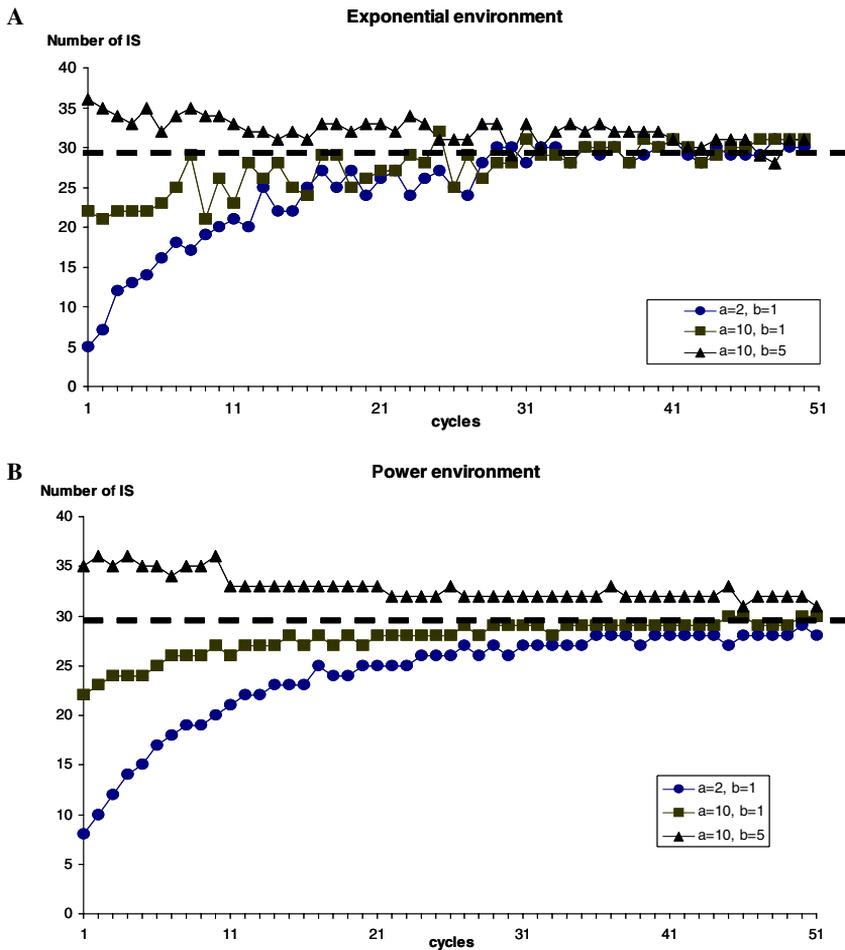


Fig. 6. The number of information-seeking (IS) actions plotted against the number of new observations sampled (cycles) from the (A) exponential and (B) power environments with different prior distributions of the amount of effort required to solve the problem. Dashed line shows the optimal number of IS for the environment. a, b are the parameters of the prior distribution of the mean savings of effort per information-seeking action. The mean of the distribution is given by ab . The equation relating the amount of effort and the number of IS is given by (A) $C_{np} * \exp(-k*n)$ in the exponential environment, and (B) C_{np}^{-nk} for the power environment. C_{np} is set to 100; k is set to 0.05, and n is the number of IS.

and power environments demonstrates that the model adapts to different diminishing-return environments even when the prior knowledge of the environment are very different. In fact, the simulations show the behavior of the model is quickly dominated by experience after a few cycles.

5.3. Behavior of the model in environments when cost and utility of information are changed

Two sets of simulations were conducted in a diminishing-return environment (exponential) with the same equation as the previous simulation (i.e., $C = C_{np} * e^{-k*n}$). For the *switch-utility* set, the cost of each information-seeking action was fixed at 1. Two switch-utility simulations were run. In the first simulation, k (which reflects utility of information in the environment, the lower the value of k , the higher the utility of information) was set to 0.2 for 25 learning cycles and then switched to 0.05 for another 25 learning cycles. The second simulation switched in the opposite direction; that is, the k was initially set to 0.05 for 25 learning cycles and then switched to 0.2 for another 25 cycles. The number of information-seeking actions chosen by the model is plotted in Fig. 7A. The values of k were chosen such that the optimal numbers of information-seeking actions were 30 and 14, as shown by the horizontal straight lines in Fig. 7A.

In the *switch-cost* set of simulations, the model learned in the same diminishing-return environment with k fixed at 0.05. In the first switch-cost simulation, the model first learned for 25 cycles with cost of information-seeking action set to 1, and then learned for another 25 cycles with the cost set to 2.5. The model was then reset and learned with cost set to 2.5, then learned for another 25 cycles with the cost set to 1. The results of these simulations are plotted in Fig. 7B. Similar to the values of k in the switch-utility set, the values of cost were chosen such that the optimal numbers of information-seeking actions were 30 and 14, as shown by the horizontal straight lines in Fig. 7B.

To summarize, in each of the two sets of simulations, one set of costs (C) and information-seeking returns (k) was identical ($C = 1$; $k = 0.05$; optimal IS actions = 30). In the second set of the switch-utility simulations, costs were held constant but k was set to 0.2 ($C = 1$; $k = 0.2$; optimal IS actions = 14). In the second set of the switch-costs simulations, utility was held constant, but costs were set to 2.5 ($C = .5$; $k = 0.05$; optimal IS actions = 14). In both simulations, when utilities or costs are switched at cycle 25, we can observe from Fig. 7 how the model adapts by changing the number of information seeking actions from 30 to 14, and vice versa. The different time courses of adaptations generated from the two simulations can therefore reveal the differences in how the model adapts to changes in utilities and costs.

Fig. 7 shows that the rates of approach to the new optimal values are faster in the switch-cost simulations (b) than in the switch-utility simulations (a). In the switch-utility case, the Bayesian learning mechanism uses new observations from the environment to update its prior knowledge of the environment. Since the Bayesian learning mechanism takes both the uncertainties of new observations and prior knowledge into account, the changes to the estimation of the parameters of the global environment are slow. On the other hand, since a local decision rule is used, when the cost is changed, the number of information-seeking actions will be affected directly, and the rates of approach to the new optimal values are therefore higher. The differential rates of approach is therefore an inherent property of the interaction of the global Bayesian learning and the local decision process.

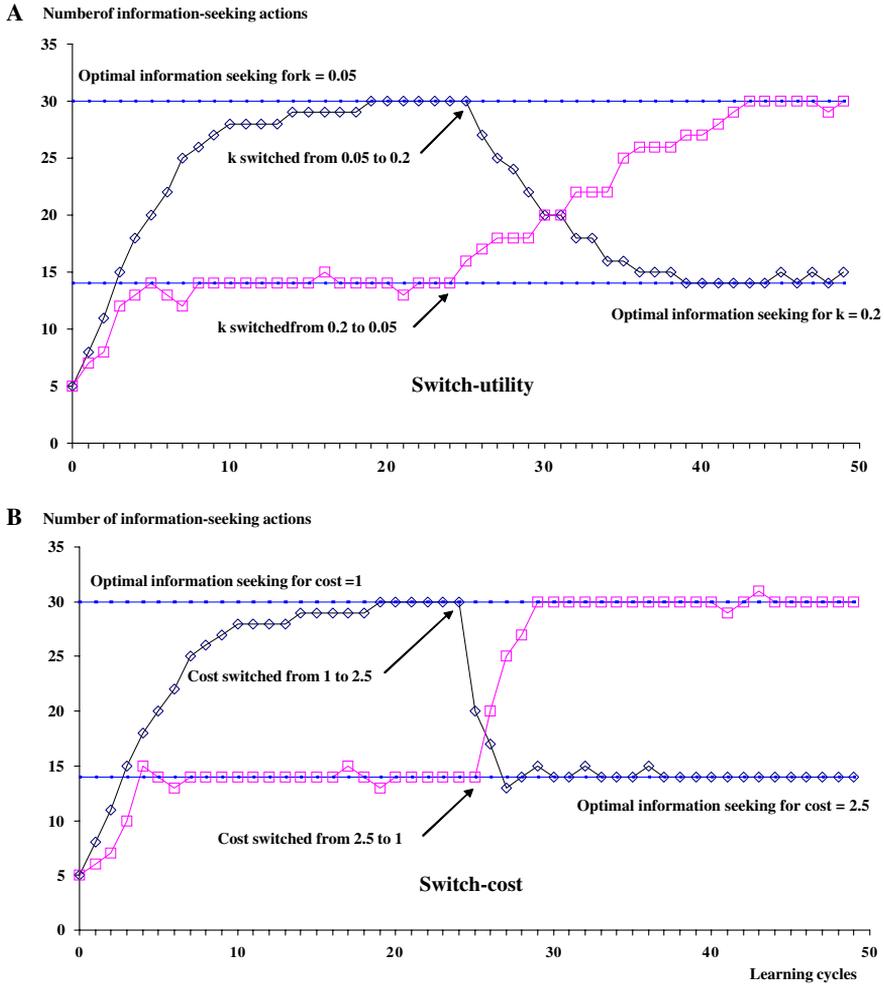


Fig. 7. Simulations of the BSM in the (A) switch-utility set (k was changed) and (B) switch-cost set (cost was changed). The model was switched to another environment at cycle 25. Note that in both cases, the optimal number of information-seeking actions are the same before and after the switch.

5.4. Behavior of the BSM in a “local-minimum” environment

The use of the local decision rule suggests that the decision to stop seeking information does not guarantee that it will lead to optimal performance. For example, Fig. 8 shows a “local-minimum” environment in which the marginal utility of information (i.e., the slope of $f(n)$) varies with the number of information-seeking actions. The marginal utility is high during initial information-seeking, becomes flat with intermediate amounts of information-seeking, but then becomes high again with greater amounts of information-seeking. The flat region in the curve (i.e., region B) is the local-minimum region. Using the local decision rule, information seeking is likely to stop at this region (i.e., when the marginal utility of information is lower than the cost), especially when the cost is high. From previous simulations, the global Bayesian learning of information utility is found to be slow.

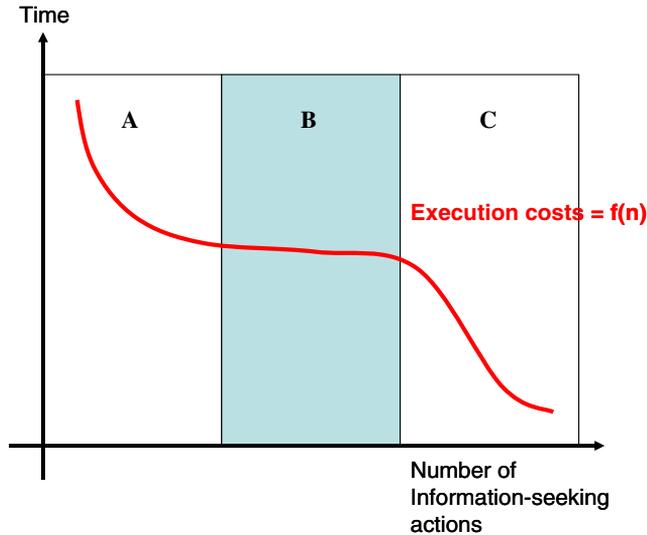


Fig. 8. A “local-minimum” environment—in region A and C, the marginal utility of information (i.e., the slope of $f(n)$) is high. However, in region B, the marginal utility of information is low. When the local decision rule is used, information-seeking may stop at region B, especially when the cost is high.

Therefore unless somehow people are forced to accumulate experiences in region C, the estimation of the marginal utility of information is likely to be dominated by experiences in regions A and B. When the cost is high, information-seeking will likely stop before region C is reached. We therefore predict that in a local-minimum environment, if a local decision rule is used, high cost may lead to poor exploration of the task space. In other words, suboptimal tradeoffs are more likely when the cost is high in a local-minimum environment.

5.5. Summary of predictions from the BSM

To summarize, simulation results of the BSM lead us to predict that (1) with sufficient experience, people make good tradeoffs between costs and utility of information and converge to a reasonably good level of performance in a general diminishing-return environment, (2) people may respond to changes in costs faster than changes in utility of information, and (3) in a local-minimum environment, high cost may lead to suboptimal tradeoffs and poor exploration of the problem space, thus worse performance. These three predictions were tested against human data collected from three experiments. To preview the results, human data supported these predictions, suggesting that the BSM was consistent with the underlying tradeoff mechanisms.

6. The task

Our three experimental studies used a map-navigation task to collect human and simulated human (i.e., ACT-R models) data in information seeking. In map navigation, a simple hill-climbing strategy (usually the shortest route) is always applicable and sufficient to accomplish the task (and any path can eventually lead to the goal), but the hill-climbing

strategy is not guaranteed to lead to the best (i.e., fastest) path. With sufficient experience, one learns the speeds of different routes and turns, and will be able to improve performance by a better choice of solution paths.

Subjects used the train map shown in Fig. 9. Subjects were given a start station and destination (end station), and were asked to travel from the start station to the end station. Subjects could choose to go to one of the adjacent stations on the same line (literally and figuratively). To go to an adjacent station, subjects had to point the mouse cursor to the station, press and hold down the mouse key. A red line would be drawn from the current station (the red dot) to the station clicked. The speeds of the train lines (and the transfers, described below) were reflected by the time it took for the red line to go from one station to the next. When the red line reached the station, the station turned red and became the current station.

Subjects could use the transfer at the intersection of the train lines to change direction. There were four transfer stations at every intersection of the train lines. When subjects were at a transfer station, they could go to another train line or stay on the same train line. Subjects were told that there were two kinds of transfers, pink transfers and orange transfers, and one of them was faster than the other. However, they were not told which one was faster. When the trial started, the colors of the transfers were covered (i.e., in black). The color of a transfer would be shown when the subject was at the transfer station or when the subject clicked on any of the transfer stations. As soon as the experiment started, the subject could check the color of any transfer (i.e., an information-seeking action) in the map anytime before they reached the end station. At any time during the experiment, the subject could see at most one transfer uncovered.

In this task, the hill-climbing heuristic (i.e., no information-seeking action) was always sufficient to finish the task, but was not guaranteed to yield the fastest path. With sufficient experience, subjects learned the speeds of different routes and turns, and were able to improve performance by a better choice of solution paths.

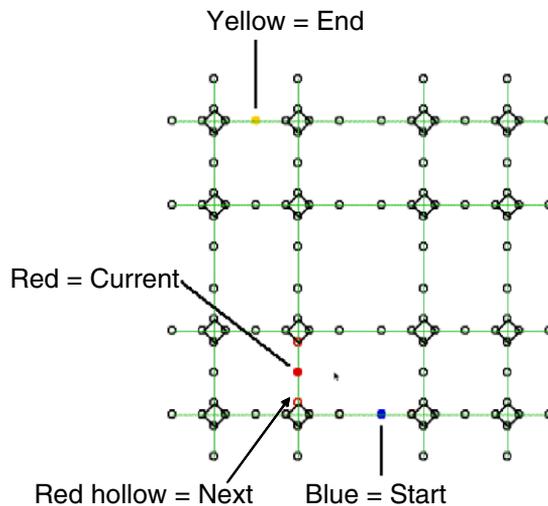


Fig. 9. The train map used in experiments 1 and 2. The blue dot is the start station, the yellow dot the end station, the red dot the current station. The red hollow circles are stations that can be reached next by clicking on one of them. Colors of all the transfers are covered (black) unless being clicked. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this paper.)

In each experiment, the information-seeking costs and the utility of information were manipulated. The task was constructed so that more information-seeking actions would decrease the execution costs. The number of information-seeking actions was measured by the number of transfers checked (and rechecked). The cost of each information-seeking action was increased by adding a 1-s lockout time after the transfer station was clicked. Utility of information (i.e., the color or speed of a transfer) was manipulated by varying the speed of the slow transfers; that is, the speed of the fast transfer was held constant. When the difference between the speeds of the slow and fast transfers was large, utility of information would be large because using slow transfers to finish the map would be much slower than the path that uses only fast transfers (which requires more information-seeking actions to discover).

The execution costs depended on the length of the path and the speed of the transfers used in the solution path. Most of the maps were designed so that the major factor affecting the execution costs was the speed of the transfers used, not the length of the path (see description of Experiment 1 later). This feature was necessary because the length of different possible paths could be judged perceptually. Besides, we thought it safe to assume that this kind of perceptual judgment would be well-learned before subjects came to the experiment. It was therefore unlikely that any significant learning effects could be observed (which was one of the goals of the current research). Hence, the maps were constructed so that explicit information-seeking actions (i.e., using the mouse to click on a transfer to uncover its color) were required to reduce the execution costs.

7. The ACT-R model of the task

The BSM aims at characterizing adaptive information-seeking at a fairly abstract level—no specific cognitive mechanisms have been specified. The model also does not assume any cognitive constraints such as memory or attention limitations. The BSM as presented above is therefore not precise enough to generate quantitative predictions that can be matched directly to actual human performance. To describe human performance at the mechanistic level, psychologically plausible mechanisms and constraints need to be imposed on the BSM so that precise predictions can be generated. ACT-R is a good candidate for this purpose. The ACT-R architecture consists of multiple mechanistic modules and a theory of how these modules are integrated to produce precise predictions of human behavior. This set of integrated mechanisms is able to explain a wide range of behavior (Anderson & Lebiere, 1998). By building models in ACT-R, we expect to obtain precise predictions that can be directly matched to human performance. A brief description of the ACT-R architecture and the relevant mechanisms can be found in Appendix B. Although ACT-R has 35 variables and parameters (see Table 12.3, p. 434 of Anderson & Lebiere, 1998), not all of these parameters are relevant to the current model. In fact, only 4 parameters¹ were set to fit the data from E1, 1 parameter was set according to the experimental conditions,² and the rest were all set at default values. These parameters were estimated from E1 and were used in the model for E2 and E3, as illustrated in Fig. 3.

¹ The 4 parameters were: the prior number of successes, the prior number of failures, the noise, and the procedural learning decay parameter.

² The cost parameter is set according to the condition.

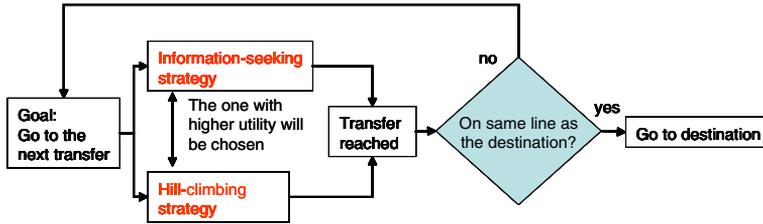


Fig. 10. The overall structure of the ACT-R model. The model assumes that the hill-climbing strategy will be used when the information-seeking strategy is not chosen. The decision on whether information will be sought is decided by the utility values of the two productions. Strategies compete only at transfers. Once a transfer decision is made the subject continues on the line to the next transfer station. At that point the strategies compete again. Details of the strategies are in Appendix C.

The main goal of the model is to go from the start station to the end station. The model assumes that an information-seeking action (i.e., moving the mouse cursor to a transfer station, clicking on it, and waiting for the color of the transfer to be shown) is initiated to decide which transfers to use. Once a transfer is chosen, the model will move to it and another cycle of decision will start. We also make the assumption that when subjects decide not to seek any information, they use hill-climbing to get to a particular transfer station. The model, however, does not assume that subjects will use a single strategy to finish the whole map; rather, in the course of completing a given map, the model can combine hill-climbing with information-seeking. Hence, the information-seeking versus hill-climbing (i.e., no information-seeking action) decision is a local choice, not a global disposition. This allows a series of dynamic choices of strategies that do not have rigid constraints on when information seeking initiates and stops. Therefore, although only two “strategies” are described below, the current model is able to capture relative differences in the number of information-seeking actions for different task environments.

Fig. 10 shows the overall structure of the model that finishes the task. When the task begins, a strategy is chosen based on ACT-R’s conflict resolution mechanism. The conflict resolution mechanism is a noisy process that selects a strategy (represented as a set of productions) based on the utility value of the first production in each strategy. Either the information-seeking or the hill-climbing strategy selects a transfer station, and the model moves to the transfer station and uses it to transfer to another line (i.e., change direction either from left–right to up–down). When the model reaches the line the end station is on, the model will move to the end station; otherwise another cycle of strategy selection will begin. Details of the model are described in Appendix C.

8. Experiment 1

Experiment 1 (E1) was designed to test the first prediction of the BSM, that behavior adapts to the costs and utility of information in the task environment. In E1, we varied the cost and utility of information to create six different environments. If people adapt to the cost and information structures of their environments, then we should see six different patterns of information-seeking behavior.

8.1. Design and procedures

E1 was a 2×3 between-subject design. The two between-subject independent variables were cost (with or without the 1-s lockout time before the color of the transfer is uncovered) and utility of information (transfer speed difference). The dependent variables were number of information-seeking actions, relative path time, and relative trial time (see Table 1). The independent and dependent variables were the same across the three experiments. The number of information-seeking actions simply measures the number of times subjects clicked on a transfer to uncover its color. Relative path time represents the execution costs spent on the path, excluding all information-seeking costs. It therefore represents the quality of the path chosen (i.e., relative path time should be described as a function of the number of information-seeking actions, similar to the function $f(n)$ in Fig. 2). Relative trial time measures the total time the subject actually spent on each trial. This includes both the execution costs and all information-seeking costs (including any lockout time).

Each subject solved 64 maps. The 64 maps were blocked into 8 groups with 8 maps in each group. The order and the types of maps within each block are shown in Table 2. In E1, 6 out of 8 maps in each group were round-about-fastest (RAF) maps. These maps had 12 slow transfers and 4 fast transfers. All transfers along the shortest path were slow transfers (3, 4.5, and 6 s). This made the shortest path not the fastest path. However, there was always one and only one path that contained only fast transfers, and it was always the fastest path. Therefore information-seeking was necessary to find the fastest path (which uses only fast transfers) on these maps. Under the assumption that people would use hill-climbing to solve the map initially, the organization of the fast and slow transfers in the RAF maps allowed the study of the transition of strategies with practice (i.e., from hill-climbing to information-seeking).

Table 1
The variables used in the three experiments

Variables	Explanation
<i>Independent variables</i>	
Information access cost (cost)	Between-subject, 2 levels Low: 0 s lockout time High: 1 s lockout time
Utility of information (utility)	Between-subject, 3 levels Low: fast transfer 1 s, slow transfer 3 s Med: fast transfer 1 s, slow transfer 4.5 s High: fast transfer 1 s, slow transfer 6 s
<i>Dependent variables</i>	
Information-seeking	The number of transfer clicks to check its color
Relative path time (i.e., execution costs—see the results section for details)	$RPT = (\text{chosen path time} - \text{fastest path time}) / (\text{hill-climbing path time} - \text{fastest path time})$ where path time is the “pure” time spent on the trains and transfers, excluding all information-seeking costs
Relative trial time (see the results section for details)	$RTT = (\text{trial time} - \text{fastest trial time}) / (\text{hill-climbing path time} - \text{fastest trial time})$, where trial time includes both information-seeking costs and execution costs

Table 2

For experiment 1, the types of map encountered by each subject in each group of 8 maps

Map 1	Map 2	Map 3	Map 4	Map 5	Map 6	Map 7	Map 8
RAF	RAF	RAF	AS	RAF	RAF	RAF	DF

Both the round-about-fastest (RAF) maps and the direct-fastest (DF) maps have 12 slow transfers and 4 fast transfers. However, all transfers on the shortest path are slow for round-about-fastest maps, but are fast for the direct-fastest maps. All-slow (AS) maps have 16 slow transfers and no fast transfers.

In one of the maps in each group, fast transfers (1 second) were located along the line of the shortest path. These direct-fastest (DF) maps had 12 slow transfers and 4 fast transfers. The purpose of these maps was to avoid the situation where subjects would never try to use the shortest path to finish the map. In one map in each group, there were no fast transfers on the map. These all-slow (AS) maps had 16 slow transfers. These maps were used to test how many transfers subjects would check before they stopped and chose a path. The AS maps served as catch trials to measure how much information seeking subjects were willing to do when no useful information could be found.

The locations of the start and end stations were randomized in all maps, with two constraints imposed to make sure that they were reasonably far apart. The first constraint was distance. Distance was measured by city block distance (CBD). Mathematically, the CBD of two points $A(x_a, y_a)$ and $B(x_b, y_b)$ on a 2-dimensional plane is simply the sum of the absolute distances in each dimension, i.e., $CBD(A, B) = |x_a - x_b| + |y_a - y_b|$. In the train map, CBD measures the minimum number of stations one needs to go through to reach the end station. In E1, the start and end stations were separated by at least 10 CBD (the maximum possible was 16 CBD in the map). The second constraint was the minimum number of transfers required to go from the start station to the end station. In E1, the start and end stations were always placed on parallel train lines (both vertical or both horizontal), which required the use of at least two transfers.

8.2. Procedure

Before the experiment began, each subject was given a practice trial. The map in the practice trial was a short-distance map (and was the same for all conditions). The cost was set according to the condition (i.e., 0- or 1-s lockout time). Subjects were told to go from the start station to the end station as fast as possible, and that they would be timed during each map. Subjects were told that there were two kinds of transfer, one was orange and the other was pink, and that one kind was faster than the other. Half of the transfers in the practice trial were orange and the other half were pink, but the actual speed of the two kinds of transfers was the same in the practice trial. They were shown how to go from one station to another, as well as how to uncover the color of the transfers. Subjects were then asked to solve the map by themselves. The experimenter answered any questions that the subjects had during the task. After they solved the map, feedback was given to let them know that they had finished one map. After the practice trial, subjects were given 64 maps.

8.3. Results

The effects of the two independent variables, cost and utility of information, on the dependent variables were analyzed by a series of ANOVAs (analysis of variance). The

dependent variables were (1) the number of information-seeking actions, (2) the relative path time (RPT), and (3) the relative trial time (RTT). The equations that calculate the RPT and the RTT will be presented below. Owing to programming errors, one direct-fastest and one slow map in all conditions were corrupted and not analyzed. There were therefore 7 direct-fastest and 7 all-slow maps in each of the 6 between-subject conditions. These direct-fastest maps and all-slow maps were identical across conditions. The round-about-fastest maps were, however, all randomly generated.

In presenting the data, we first present analyses of the empirical data and then compare it with the predictions of the model. As ACT-R models are probabilistic, we ran the same model 15 times so that the predictions we report are stable.³ The model was run with the goal value (G) set to the default value of 20, the expected gain noise (i.e., s in the conflict resolution equation) set to 0.3. The expected gain noise controls the proportion of times the production with the highest expected gain is picked in the conflict resolution mechanism.⁴

8.3.1. Number of information-seeking actions

Subjects had significantly more information-seeking actions in the all-slow maps (mean = 5.6, $SD = 0.5$) than in the round-about-fastest maps (mean = 4.5, $SD = 0.4$), and they had the least in the direct-fastest maps (mean = 2.7, $SD = 0.2$). Paired t tests show that all differences are significant ($p < 0.01$). Since different types of maps required different numbers of information-seeking actions to find a reasonably fast path, the results suggest that subjects stopped seeking information when a reasonably fast path was found in the direct-fastest and round-about-fastest maps.

Fig. 11 shows the number of information-seeking actions in the round-about-fastest, direct-fastest, and the all-slow maps respectively. The main effects of cost and utility of information, as well as their interaction were significant in the round-about-fastest maps ($F(1, 84) = 56.11$, $MSE = 12083.5$, $p < 0.01$, $F(2, 84) = 42.82$, $MSE = 9222.6$, $p < .01$, $F(2, 84) = 10.15$, $MSE = 2185.9$, $p < .01$), direct-fastest maps, ($F(1, 84) = 12.31$, $MSE = 201.168$, $p < .01$ and $F(2, 84) = 20.35$, $MSE = 332.63$, $p < .01$, $F(2, 84) = 5.00$, $MSE = 81.68$, $p < .01$), and all-slow maps ($F(1, 84) = 37.95$, $MSE = 2645.53$, $p < .01$ and $F(2, 84) = 34.39$, $MSE = 2397.386$, $p < .01$, $F(2, 84) = 6.70$, $MSE = 467.10$, $p < .01$). The same pattern of results is obtained from the all-slow maps as for the other two map types. This all-slow pattern shows that subjects' decision to stop seeking information was sensitive to the cost and information structures of the task environment and was not specific to the particular maps.

Fig. 11 also shows that the model fit the data well in the round-about-fastest maps ($R^2 = 0.93$, $RMSE = 0.07$), in the direct-fastest maps ($R^2 = 0.83$, and $RMSE = 0.05$), and the all-slow maps ($R^2 = 0.89$, and $RMSE = 0.12$). The model exhibited the same pattern of information-seeking behavior as human subjects. This shows that the model has shown the same sensitivity to costs and utility of information as subjects, and was able to adapt to different cost and information structures as subjects did.

³ We do not mean to imply that different runs of the model are equivalent to data from different human subjects. However, it is the case that after 15 runs, the variance from the model in each condition was within the range of the variance from the 15 human subjects in that condition.

⁴ For the sake of parsimony, the standard procedural learning mechanism was used (a refined procedural learning mechanism that incorporates time-based decay is available, see Appendix B).

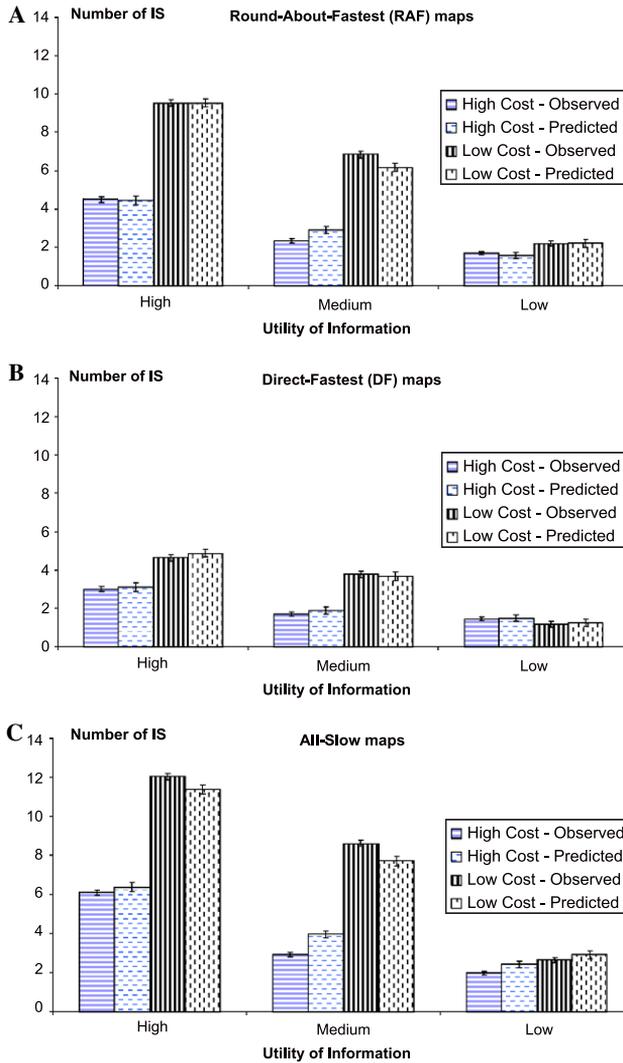


Fig. 11. The number of information-seeking actions (IS) for each of the between-subject conditions in (A) round-about-fastest, (B) direct-fastest, and (C) all-slow maps. Error bars represent standard errors of the means. The standard errors of the model are calculated from 15 iterations of model runs.

The results across different map types show a consistent pattern supporting the hypotheses that subjects had more information-seeking actions when the utility of information was high and had fewer information-seeking actions when the cost was high. The results support the notion that information-seeking is adaptive—people are sensitive to both costs and utility of information and they will adjust the number of information-seeking actions to tradeoff costs with utility of information. The results complement previous research that showed people are sensitive to information-seeking costs. Our results show that people are willing to spend more time in information-seeking when the utility of information is high, even when costs are high. The good match of the model data to human data in all three

types of maps shows that the three mechanisms in ACT-R capture the tradeoffs. The match provides at least partial support for the cognitive reality of BSM. To simplify our presentation of the results, from here on in we focus on performance in the round-about-fastest maps.

8.3.2. Relative path time (RPT)

Path time is defined as the execution costs taken by the chosen path from the start station to the end station, excluding all time spent in information seeking. Path times therefore depend on the paths chosen as well as the experimental condition the subjects were in. Since the RAF maps were designed so that the most direct hill-climbing paths (also the shortest distance) were always slower (thus lower quality) than the fastest path, the time spent in the shortest path in each round-about-fastest map can be used as the lower bound for the measure of quality of paths. The hill-climbing path can also be taken as the path chosen when no information seeking is done. The difference between the hill-climbing path time and the fastest path time therefore roughly reflects the maximum time savings the subjects could gain with information seeking. The quality of a path chosen in the round-about-fastest maps can be calculated as (see Table 1)

$$\text{RPT} = (\text{chosen path time} - \text{fastest path time}) / (\text{hill-climbing path time} - \text{fastest path time}).$$

Fig. 12 shows the RPT in the 6 between-subject conditions. Overall, subjects chose faster paths (by more information seeking, see Fig. 11) when the cost was low or when the utility of information was high. The main effects of cost and utility are significant ($F(1, 84) = 8.49$, $MSE = 7.16$, $p < .01$ and $F(1, 84) = 4.31$, $MSE = 3.63$, $p < .01$ respectively). The interaction between cost and utility is not significant ($F(2, 84) = 0.27$, $MSE = 0.228$, $p > .5$). Fig. 12 shows that the model fit the data well, $R^2 = 0.95$, $RMSE = 0.016$. The model not only had the similar number of information-seeking actions across different cost and information structures, it also found similar paths and finished the map in similar ways as subjects.

8.3.3. Trial time and relative trial time

Trial time measures the total time the subjects spent to finish each map (i.e., the sum of information-seeking costs and execution costs). Subjects were instructed to minimize the trial time in all conditions. Trial time depends on costs and utility in different experimental conditions, as well as the path chosen and the number of information-seeking actions. However, differences in trial times could be an artifact of the design of the task, as different experimental conditions inherently made the trial time different. To statistically compare across benefit conditions, a normalized trial time measure is needed. Similar to relative path time, relative trial time⁵ (RTT) can be calculated as:

$$\text{RTT} = (\text{trial time} - \text{fastest path time}) / (\text{hill-climbing path time} - \text{fastest path time}).$$

⁵ RTT were square-root transformed so they follow a normal distribution.

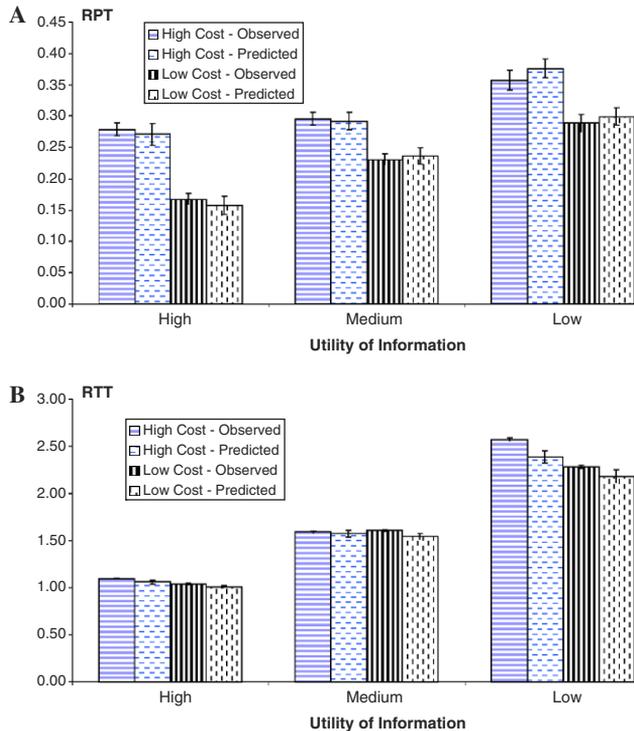


Fig. 12. (A) The relative path time (RPT) and (B) relative trial time (RTT) for the round-about-fastest maps in each of the between-subject conditions. The lower the relative path time, the higher the quality (or faster) of the path chosen. The error bars represent standard errors of the means. The standard errors of the model are calculated from 15 iterations of model runs.

Relative trial times are also useful in showing how well subjects traded costs for utility of information. Subjects in the high cost conditions had fewer information-seeking actions, but it is not clear whether the time savings from less information-seeking justified the use of slower paths. Fig. 12 shows the RTT in the 6 between-subject conditions for round-about-fastest maps. The main effects of cost and utility are significant ($F(1,84) = 6.07$, $MSE = 12.80$, $p < 0.01$ and $F(1,84) = 321.34$, $MSE = 677.791$, $p < .001$ respectively). The interaction between cost and utility is significant ($F(2,84) = 4.366$, $MSE = 9.2$, $p < .05$). The effects of utility are significant in all levels of cost. However, the effect of cost is significant only in the low utility condition ($F(1,28) = 5.84$, $MSE = 29.78$, $p < .05$). The effect of cost is not significant in the high utility condition ($F(1,28) = 2.46$, $MSE = 0.604$, $p > .1$), nor in the medium utility condition ($F(1,28) = 0.159$, $MSE = 0.156$, $p > .5$).

From previous analyses, when costs were high, subjects reduced the number of information-seeking actions. Although the main effect of cost is significant, the significant interaction implies that the effects of cost depend on the level of utility. Specifically, we found that the simple effect of cost is significant when the utility of information is low, but not when the utility of information is high. This suggests that subjects were adaptively trading off costs with utility of information—the slower path chosen was compensated by the reduction of information-seeking cost—thus there was no significant difference in the overall

performance. Note that Fig. 12 shows that relative to the best possible trial time the difference between costs for the high utility conditions was not significant. Although path time was higher when the utility is higher (see Fig. 12A), the relative trial time was actually lower. This suggests that subjects in the high utility conditions were performing close to the optimal level of performance.

As shown in Fig. 12, the fit of the model to these data was good, $R^2 = 0.88$, $RMSE = 0.21$. In general, however, the model slightly under-predicted the relative trial times—it performed faster than the subjects. This is not too surprising as the model assumed unrealistically that the subjects were completely focused on the task. The model can thus be taken as the ideal performance of the subjects. On the other hand, the model does duplicate the interaction of cost and utility as exhibited by the subjects. The model thus provided evidence that by trading off costs against utilities (by changing the expected values of the two strategies), the same interaction effects are obtained. This model has thus provided a reasonable account of the pattern of empirical results.

8.3.4. Summary and conclusions of E1

In general, E1 shows that subjects were sensitive to costs and utility of information and were able to perform good tradeoffs between costs and utility. Subjects had fewer information-seeking actions when costs were high, and had more when utility was high. With more information-seeking actions, subjects were able to find faster paths from the start station to the end station (as shown by relative path times). The tradeoffs between the costs incurred by information-seeking and the utility of information seem adaptive. In the low cost conditions, subjects had more information-seeking actions but as a consequence they found faster paths. In the high cost conditions, subjects had fewer information-seeking actions but as a consequence they found slower paths. Taken together, subjects in the low and high cost conditions had roughly the same level of overall performance (as measured by RTT).

The same pattern of information-seeking behavior in different cost and utility conditions in the all-slow maps suggests that the decision on when to stop seeking information did depend on past experiences, which apparently helped the subjects to decide how much time savings they could obtain from further information seeking. (If the decision on when to stop seeking information were independent of past experiences, there would have been little difference in the number of information-seeking actions across different conditions.)

Overall, the model provides a compelling account of the tradeoff between costs and utility of information from E1. Both the model and subjects adapt to different environments by using similar numbers of information-seeking actions. Likewise, the quality of paths chosen by the model is similar to the quality of paths chosen by the subjects. In common with our human subjects, the model subjects are sensitive to costs as well as to situations where information-seeking is not necessary to improve performance. The fit of the model to data also suggests that the model's stop-rule for information seeking is similar to that of the subjects. The model's ability to match the human data in all three dependent variables provides a solid ground for using the same model for matching data from E2 and E3.

Another promising aspect of the model is that only *environment parameters* (i.e., the two independent variables in the six between-subject conditions) were changed. All other parameters were held constant. Hence, the good match in the six conditions comes from running the same model in different environments. Different information-seeking behavior

of the model was caused by adaptation through interacting with the environments. The same approach will be adopted when matching the data from E2 and E3—i.e., only the environment parameters will be changed according to the different experimental conditions. This provides even stronger constraints to the model.

9. Experiment 2

E1 showed that subjects adaptively changed the number of information-seeking actions in response to different task environments. These results, however, were based on between-group differences. To further study the processes underlying the tradeoffs, we conducted a within-subject study. The study of the change of information-seeking behavior within an individual in response to changes in costs and utility of information will give a direct test as to how well the Bayesian learning mechanism of the model resembles the learning mechanisms of human subjects. This was the goal of Experiment 2 (E2).

As in E1, in E2 subjects were instructed to minimize trial time. To do that, subjects could either reduce information-seeking costs or execution costs. Our focus in E2 is on comparing changes in subjects' behavior when costs are varied to changes in behavior when utility of information is varied. We predict that subjects will be more sensitive to changes in costs than to changes in utility.

As discussed earlier, the use of the local decision rule (i.e., the stopping rule in the BSM) implies that when costs are varied, the perception of different costs will lead to an immediate change in the number of information-seeking actions. On the other hand, when utility of information is varied, the Bayesian learning equation requires a series of consistent observations to update the global estimation of the environment. In the map-navigation task, when the costs are varied (i.e., 0- or 1-s lock-out time when uncovering the color of the transfer), the decision on how many transfers to check will be changed once the cost is perceived. However, when the utility is varied, the decision on how many transfers to check will be changed when the end station was reached and the total trial time was compared to past experiences of the task. E2 focused on subjects' differential sensitivity to changes in costs and utility of information.

In E1 we used two levels of costs and three levels of utility. In E2 we used the same levels of low and high cost as in E1 and the medium and high levels of utility from E1. As there are only two levels of utility in E2, we refer to these levels as low and high. (Hence, the E2 low level of utility was the E1 medium level.) We define our E2 transfer conditions in terms of costs and utility. Transfer to the Lo–Hi (Low cost and High utility) condition from either the Hi–Hi condition or the Lo–Lo condition should increase the tendency to seek information (because for the Hi–Hi to Lo–Hi transfer, the cost is decreased; for the Lo–Lo to Lo–Hi transfer, the utility is increased). However, if subjects are more sensitive to cost than utility, the increase in information seeking should be faster when the cost is decreased (Hi–Hi to Lo–Hi) than when the utility is increased (Lo–Lo to Lo–Hi). On the other hand, the transfer to the Hi–Lo condition from either the Hi–Hi condition or the Lo–Lo condition should *decrease* the tendency to seek information (because for the Hi–Hi to Hi–Lo transfer, the utility is decreased; for the Lo–Lo to Hi–Lo transfer, the cost is increased). Again, we predict that subjects are more sensitive to costs than utility; hence, the decrease in information seeking should be faster when the cost is increased than when the utility is decreased.

9.1. Design and procedures

E2 was a within-subject design. The main manipulation was to transfer subjects from one set of costs and utility to another. Like E1, subjects were given instructions and a practice trial. Half of the subjects were assigned to the High Incentive group and the other half to the Low Incentive group. In the High Incentive group, the focus was on how subjects adapted to the increase in information-seeking incentive, i.e., either lower cost or higher utility. In the Low Incentive group, the focus was on how subjects adapted to the decrease in information-seeking incentive, i.e., either higher cost or lower utility.

After a base block of 8 trials, subjects in each group were given two sets of one switch-utility block and one switch-cost block (see Table 3). The order of the switch-utility and switch-cost blocks was counterbalanced. For the High Incentive group, the switch-cost block consisted of 8 training trials of the Hi–Hi condition and 8 transfer trials of the Lo–Hi condition (decreased cost); the switch-utility blocks consisted of 8 training trials of the Lo–Lo condition and 8 transfer trials of the Lo–Hi condition (increased utility). Since the 8 transfer trials in both the switch-cost and switch-utility blocks were in the same set of costs and utility (i.e., in the Lo–Hi condition), the behavior of the subjects in the transfer trials could be compared across blocks. If there was no difference between the sensitivity of costs and utility, then the responses of the subjects in the transfer trials will be the same in the switch-cost and the switch-utility block. However, if subjects were more sensitive to costs than to utility, responses (i.e., more information-seeking actions) in the transfer trials of the switch-cost block would be faster than those in the switch-utility block. Similar comparisons could be in the Low Incentive group, in which the information-seeking incentive was decreased.

To obtain a stable prediction of performance, the model was run 10 times in each condition of E2. Since E2 is a within-subject manipulation, the adaptation to the changing costs and utility are modeled by using ACT-R's time-based decay procedural learning mechanism that discounts past experiences. All parameters for the E1 model were reused in the E2 model. Indeed, the two models were identical except for changes in the expected gain noise parameter and the use of the time-based decay mechanism. For the sake of parsimony, the time-based decay mechanism had not been used by the model during E1.

Other than the use of time-based decay, the only difference between the E1 and E2 ACT-R models was the setting of expected gain noise (egs). The E1 value was 0.3. For E2, egs was set to 0.5 as a post hoc change designed to obtain a better fit to the data. In general, the higher the expected gain noise, the higher the chance that a production with a lower utility value will be chosen. The consequence is that the model will be more likely

Table 3
Table showing the design of E2

Group	Base	Block 1				Block 2			
		Switch-utility		Switch-cost		Switch-utility		Switch-cost	
		Training	Transfer	Training	Transfer	Training	Transfer	Training	Transfer
HI	Lo–Hi	Lo–Lo	Lo–Hi	Hi–Hi	Lo–Hi	Lo–Lo	Lo–Hi	Hi–Hi	Lo–Hi
LI	Hi–Lo	Hi–Hi	Hi–Lo	Lo–Lo	Hi–Lo	Hi–Hi	Hi–Lo	Lo–Lo	Hi–Lo

HI, high incentive; LI, low incentive. Lo–Lo, low cost and low utility; Lo–Hi, low cost and high utility, etc. The switch-cost and switch-utility blocks were counterbalanced within each incentive group.

to choose different strategies. Our interpretation of the higher expected gain noise is that in the within-subject manipulation, subjects were aware that the costs and utility were changing and, therefore, were more willing to try different strategies than in the static E1 environment. Note that we changed only 2 out of 35 parameters in E2 (one for the decay and the other for the noise parameter). To preview, in E3, we used exactly the same set of parameters as in E1 (i.e. the noise was set to 0.3 and the decay mechanism was turned off).

9.2. Empirical results

The same independent (cost and utility of information) and dependent variables (number of information-seeking actions, relative path time, and relative trial time) were used in E2 as in E1. Two sets of analyses were performed on each incentive group. For each group, a separate 2×2 (switch-cost and switch-utility \times first and second blocks) ANOVAs was conducted on the training and transfer sets. The transfer analyses attempted to determine if there were differential transfer effects between the switch-cost and switch-benefit blocks, and if so, whether these differential transfer effects interacted with practice. If differential transfer effects exist, the results of the training analyses will tell if the differences can be explained by performance in the training blocks (i.e., before the transfer). To preview our results, we found that subjects in general were more sensitive to changes in costs than changes in utility. Since the overall pattern of results from both Incentive groups were similar as far as the support for the hypothesis on the differential sensitivity was concerned, we chose to report only results from the High Incentive group below. As the effect of order of transfer (e.g., the switch-cost block presented first or second) was not significant, this comparison is not reported here.

9.2.1. Number of information-seeking actions

Fig. 13 shows the number of information-seeking actions in the transfer trials of the High Incentive group. An ANOVA on the number of information-seeking actions in the transfer trials (i.e., the Lo–Hi condition) showed that subjects had significantly more information-seeking actions in Block 1 than Block 2, $F(1, 75) = 9.64$, $MSE = 290.51$, $p < .01$. Subjects in the transfer trials also had significantly more information-seeking actions in the decreased-cost transfer than in the increased-utility transfer, $F(1, 75) = 7.02$, $MSE = 88.17$, $p < .01$. The kinds-of-transfers by Blocks interaction was not significant, $F(1, 75) = 2.75$, $MSE = 51.04$, $p > .1$. These analyses suggest that subjects were more sensitive to changes in costs than utility. Simple effect analyses showed that subjects had significantly more information-seeking actions in the transfer trials in the decreased-cost transfer than the increased-benefit transfer in Block 1, $F(1, 15) = 5.45$, $MSE = 143.521$, $p < .05$, but not in Block 2, $F(1, 15) = 0.243$, $MSE = 1.69$, $p > .5$. This suggests that the differential sensitivity to costs and benefits became not significant with practice.

ANOVA on the number of information-seeking actions in the training trials (i.e., the Lo–Lo versus Hi–Hi conditions) showed that the main effect in the number of information-seeking actions was not significant, $F(1, 75) = 0.0004$, $MSE = 0.1$, $p > .9$, but the main effect of Blocks was significant, $F(1, 75) = 14.43$, $MSE = 253.3$, $p < .01$ (see Table 4). Subjects had significantly more information-seeking actions in Block 1 than in Block 2. The lack of difference between the Lo–Lo and Hi–Hi condition (in Block 1) suggests that subjects had roughly the same number of information-seeking actions in the training

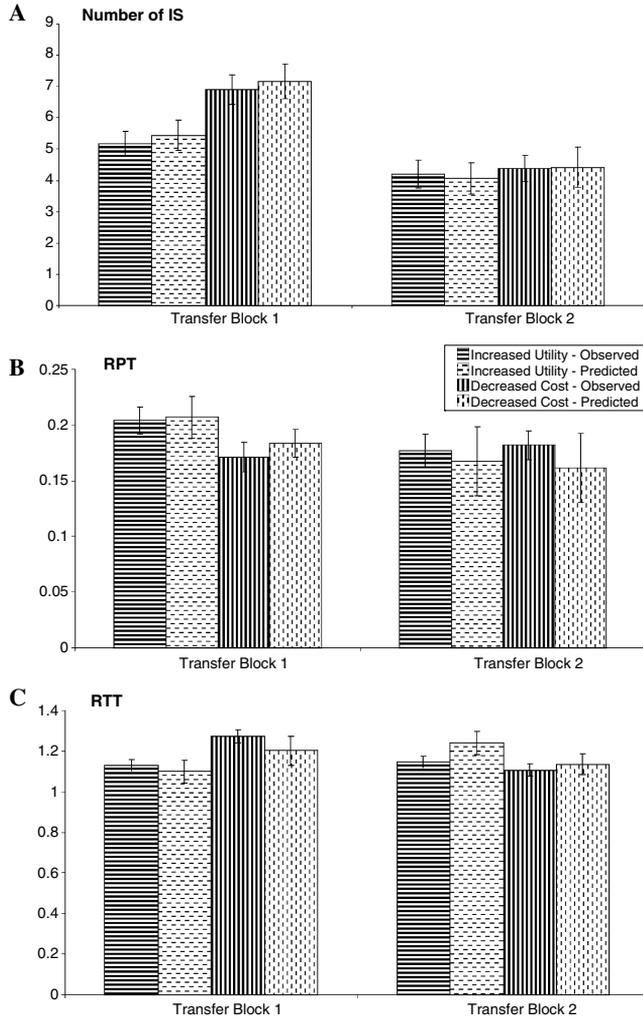


Fig. 13. (A) The number of information-seeking actions (IS), (B) relative path time (RPT), and (C) relative trial time (RTT) in the High Incentive group. The error bars of the model represent deviations in 10 simulations of the model.

blocks before they were transferred to the Lo–Hi condition, thus the differences found in the Lo–Hi condition cannot be simply attributed to inertia from the training blocks.

Fig. 13 and Table 4 also show the model fit of the number of information-seeking actions in E2. In general, the model fit the data well in both the training and transfer trials, $R^2 = 0.98$, $RMSE = 0.24$. Specifically, the model responded to changes in costs faster than changes in utility in Block 1, but not in Block 2. In Block 1, the model had relatively fewer past experiences to rely on than it did in Block 2. Thus, the decay-based procedural learning mechanism was able to respond faster to changes in the environment by putting higher weighting on recent experiences. However, in Block 2, the longer history of past experiences (even after the discounting by the decay-based mechanism) made the relative impor-

Table 4
Means of the training trials in E2

	Training Block 1		Training Block 2	
	Observed	Predicted	Observed	Predicted
<i>Increased utility</i>				
Number of IS	5.125 (0.487)	5.144 (0.562)	3.448 (0.335)	3.682 (0.446)
RPT	0.228 (0.021)	0.230 (0.032)	0.251 (0.024)	0.227 (0.027)
RTT	1.889 (0.040)	1.855 (0.086)	1.761 (0.037)	1.220 (0.042)
<i>Decreased cost</i>				
Number of IS	5.062 (0.555)	5.515 (0.539)	3.490 (0.336)	3.668 (0.455)
RPT	0.222 (0.019)	0.227 (0.031)	0.241 (0.019)	0.270 (0.030)
RTT	1.312 (0.040)	1.220 (0.073)	1.195 (0.028)	1.246 (0.046)

Numbers in parentheses represent standard deviations. Standard deviations for the model are calculated from 10 simulations of model runs.

tance of recent experiences lower, thus producing a similar rate of response to the changed costs and utility.

9.2.2. Relative path time

Fig. 13 shows the RPT (the same definition as that in E1) in the transfer trials of the High Incentive group. A 2 (switch-cost and switch-utility transfers) \times 2 (Blocks) ANOVA on the relative path times in the transfer trials showed that the main effect of Blocks was not significant, $F(1, 75) = 0.43$, $MSE = 0.006$, $p > .5$. The main effect of kinds-of-transfer was marginally significant $F(1, 75) = 3.68$, $MSE = 0.02$, $p = .07$. Subjects in the decreased-cost transfer found slightly faster paths than in the increased-utility transfer. The kinds-of-transfer by Blocks interaction was significant, $F(1, 75) = 4.27$, $MSE = 0.034$, $p < .05$.

Simple effect analyses showed that the difference in the transfer trials between the decreased-cost and increased-utility transfers was significant in Block 1, $F(1, 15) = 4.0$, $MSE = 0.304$, $p < .05$, but was not significant in Block 2, $F(1, 15) = 0.344$, $MSE = 0.011$, $p > .5$. The difference between the blocks suggests that when the cost was decreased, subjects had more information-seeking actions and found better paths, but that this greater sensitivity to cost than to utility diminished with practice.

ANOVA on the RPTs in the training trials showed that the main effect of kinds-of-transfer in the RPTs was not significant, $F(1, 75) = 0.158$, $MSE = 0.006$, $p > .6$, nor was the main effect of Blocks, $F(1, 75) = 1.51$, $MSE = 0.042$, $p > 0.2$ (see Table 4). None of the differences in the training trials (i.e., between the Lo–Lo and Hi–Hi conditions) was significant. None of the effects of Blocks was significant. This, again, showed that before subjects were transferred, the quality of paths chosen in the training blocks was roughly the same. The difference in RPT in Lo–Hi therefore cannot be simply explained by inertia from previous training trials. Fig. 13 and Table 4 also show a good fit of the model to the data in both the training and transfer trials ($R^2 = 0.92$, $RMSE = 0.015$).

9.2.3. Relative trial time

Fig. 13 shows the RTT (the same definition as that in E1) in the transfer trials of the High Incentive group. Results from the ANOVA on the RTT in the transfer trials showed that the main effect of Blocks was significant, $F(1, 75) = 7.084$, $MSE = 0.245$, $p < .01$. Subjects finished the map faster in Block 2 compared to Block 1. The main effect of kinds-of-

transfer was not significant $F(1, 75) = 0.252$, $MSE = 0.07$, $p > .6$. The kinds-of-transfer by Blocks interaction was significant, $F(1, 75) = 5.69$, $MSE = 0.209$, $p < .05$.

The difference between decreased-cost and increased-utility transfers was significant in Block 1, $F(1, 15) = 4.16$, $MSE = 0.147$, $p < .05$, but was not significant in Block 2, $F(1, 15) = 2.33$, $MSE = 0.07$, $p > .1$. Recalling previous results, subjects in the decreased-cost transfer had more information-seeking actions in Block 1 and found a better path. However, they were *slower* to finish the trials in the transfer blocks, compared to the increased-utility transfer. This suggests that when the cost was decreased, subjects spent *too much time* seeking information to find a good path, therefore ended up with a slower relative trial time. However, in Block 2, the effect was not significant.

For the training trials, the ANOVA showed a significant main effect of kinds-of-transfer, $F(1, 75) = 102.363$, $MSE = 31.34$, $p < 0.01$, and a main effect of Blocks, $F(1, 75) = 20.81$, $MSE = 1.46$, $p < .01$ (see Table 4). The difference between the Lo–Lo and Hi–Hi condition in Block 1 was significant, $F(1, 15) = 78.46$, $MSE = 15.97$, $p < 0.01$, and was also significant in Block 2, $F(1, 15) = 83.56$, $MSE = 15.37$, $p < .01$. The interaction of kinds-of-transfer and Blocks was not significant. Subjects were faster in both conditions with practice. The significant differences in RTT between the Hi–Hi and the Lo–Lo conditions were not surprising, given the fact that only the different utility conditions were normalized, not the different levels of cost. The relative trial times in the high cost conditions were significantly higher than low cost conditions mostly because of the extra time (the lockout time) spent on checking in the high cost conditions.

Fig. 13 and Table 4 also shows the model fit to the data in both the training and transfer trials. The quantitative fit was reasonably good, $R^2 = 0.66$, $RMSE = 0.16$. In general, the model captured the overall patterns of the empirical data and showed the same consequences of differential sensitivity to costs and benefits to performance. Specifically, in Block 1, when costs decreased the model planned more and found faster paths, but the overall performance (i.e., relative trial time) was worse.

9.3. Summary and conclusions of E2: Suboptimal tradeoffs between costs and utility

Based on results from E2, we conclude that subjects reacted to changes in costs faster than changes in utility. When the costs were reduced, subjects increased the number of information-seeking actions and found better paths than when utility was increased. We found the complementary pattern in the Low Incentive group. When the costs were increased, subjects made fewer information-seeking actions and found slower paths compared to the decreased-utility transfer. However, the overall performance (in both Incentive groups) in the switch-cost transfers was worse (slower) compared to the overall performance in the switch-utility transfers. This intriguing set of results suggests that subjects' decisions on when to stop seeking information were more sensitive to costs than utility. In fact, the results suggest that subjects "over-reacted" to changes in costs, thus leading to suboptimal tradeoffs between costs and utility of information. When costs were decreased subjects spent too much time gathering information. Unfortunately, this time spent was not a good investment, as the cost of information seeking was not recovered by the faster routes.

The overall fit of the model was promising; the model responded to changes in costs and utility the same way as the subjects. Since changes in cost updated the expected utilities of productions after the model clicked on a transfer, the responses were relatively fast. In

contrast, changes in utility of information updated the expected utilities of productions only after the model moved to and used the transfer. The changes in expected utilities of the productions affected how often the hill-climbing or the information-seeking strategy was used. The good fit of the model to the data shows that the differential rates of responses to changes in costs and utility matched well to those in the human data.

The time-based decay learning mechanism allows the model to discount remote past experiences so that it responds to recent past experiences faster. This mechanism is used because when the cost and information structures of the environment are changed, new observations need to have a larger impact on behavior than previous experiences in the old environment. The differential weighting of past experiences provides good explanations of the behavior observed in E2. It is possible that the mechanism captured the behavior of the subjects when they realized that the environment had significantly changed and began to put more weight on new observations.

The noisy conflict resolution controls the probability that a previously favorable (i.e., one that has a higher utility value) strategy would be used again. This was particularly important because in E2, subjects were given a series of different within-subject conditions, and a once favorable strategy may not continue to be favorable when the condition is changed. The model therefore had to “sample” the results of different strategies (even if they were less favorable based on past experiences) to know which strategy would be best in a particular condition. In E2, the model fit was better with a higher expected gain noise, suggesting that subjects could be exhibiting a higher probability of this kind of “sampling” behavior. The tradeoff between this kind of sampling behavior and adopting consistently the strategy that works the best is often cast as a tradeoff between exploration and exploitation. In a changing environment, the best strategy may not always be the best all the time, thus it may be beneficial to shift the balance to more exploration than exploitation. The increase in expected gain noise basically achieve this in the model: the model will be more likely to select different strategies (exploration) than to use the strategy that has the highest utility (exploitation).

10. Experiment 3

E3 tested the third prediction of the BSM—high cost may lead to poor exploration of the task space in a local-minimum environment. In E2, subjects were found to be more sensitive to costs than utility. High costs may therefore lower people’s willingness to explore the task environment because exploration requires higher costs that may or may not be compensated by better solutions. The use of the local decision rule in the BSM implies that when costs are high, information seeking may stop prematurely, thus leaving a major part of the task space unexplored.

The RAF maps in E3 were designed as shown in the notational Fig. 14. In contrast to the RAF maps used in E1 and E2, these are local-minimum maps in that a fast path to a location close to the end station could be found by a small number of information-seeking actions. However, this path would eventually lead to a slow transfer in the map (right before they reach the end station), thus making the overall path to the end station slow. The fastest path from the start station to the end station required more information-seeking actions and uses a path that initially went away from the end station, but did not require the use of any slow transfer to reach the end station. To increase the utility of information, the slow transfers in E3 were twice as slow as those in E1 and E2. The costs

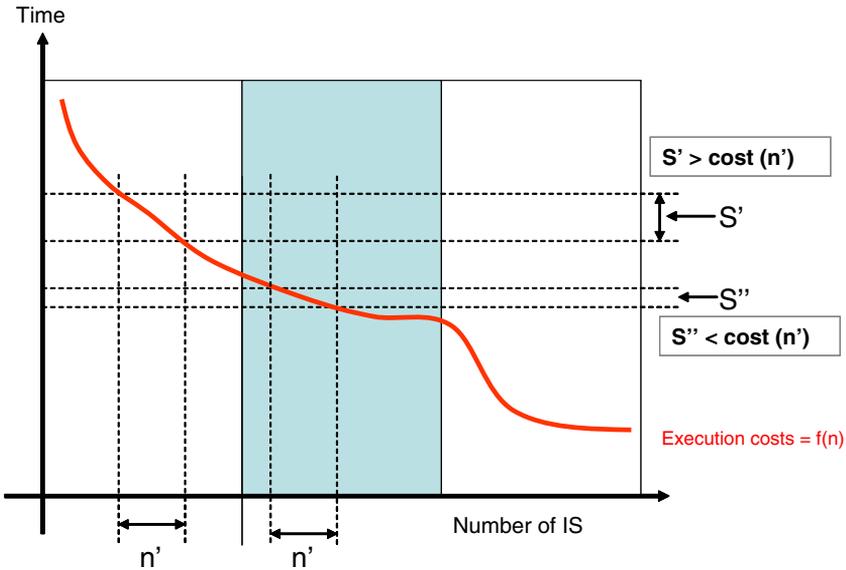


Fig. 14. Information seeking stops prematurely in a local-minimum environment. For the same additional number of information-seeking actions n' , the time savings in region A is S' . However, in region B, the same number of additional information-seeking actions n' will have a much smaller time saving of S'' . It is likely that S' will be larger than the cost of n' information-seeking actions, thus information-seeking will continue. However, S'' will be likely to be smaller than the cost, thus information-seeking is likely to stop in region B, especially when the cost is high. Even though much better solutions exist in region C, the premature stop of information seeking will lead to poor experience in region C, leaving a large portion of the task space unexplored.

were the same as in E1 and E2. E3 is a between-group study in which half of the subjects were assigned to a low cost condition, the other half to the high cost condition.

10.1. Design and procedures

Each of the 20 subjects was given 40 trials, half of the maps were direct-fastest maps and half were local-minimum maps. The order of the maps was randomized with the constraint that 4 direct-fastest and 4 local-minimum maps were given every 8 trials. This is to make sure that subjects would have roughly equal experience with each type of map during the experiment. Half of the subjects were randomly assigned to the high and low cost conditions.

10.2. Results

Predictions from the model were obtained by running the model 10 times. The major change specific to E3 was one that reflected the difference between the E3 task environment and the task environment used in E1 and E2. Specifically, the effort parameter of *use-slow-transfer* in E3 was set at 32, (i.e., twice as large as that in the High Utility condition of E1, see Appendix C). The model used here was the model used in E1 and E2. As in E1, the expected gain noise was set to 0.3. Since the task environment was constant throughout the study, the time-based decay mechanism was not used as in E1.

10.2.1. Number of information-seeking actions

Fig. 15 shows the number of information-seeking actions in the two between-subject conditions. A 2 (high/low cost, between-subject) \times 2 (map types, within-subject) ANOVA shows that the main effect of cost is significant, $F(1, 22) = 23.7$, $MSE = 6342.13$, $p < .01$. The main effect of map types is also significant, $F(1, 22) = 32.89$, $MSE = 377.82$, $p < .01$. The cost \times map types interaction is significant, $F(1, 22) = 6.30$, $MSE = 72.42$, $p < .05$. Subjects had significantly more information-seeking actions in the low cost condition, and in the local-minimum maps. The greater amount of information-seeking actions in the low cost condition was consistent with the results from E1 and E2. However, subjects had significantly more information-seeking actions in the local-minimum maps than in the

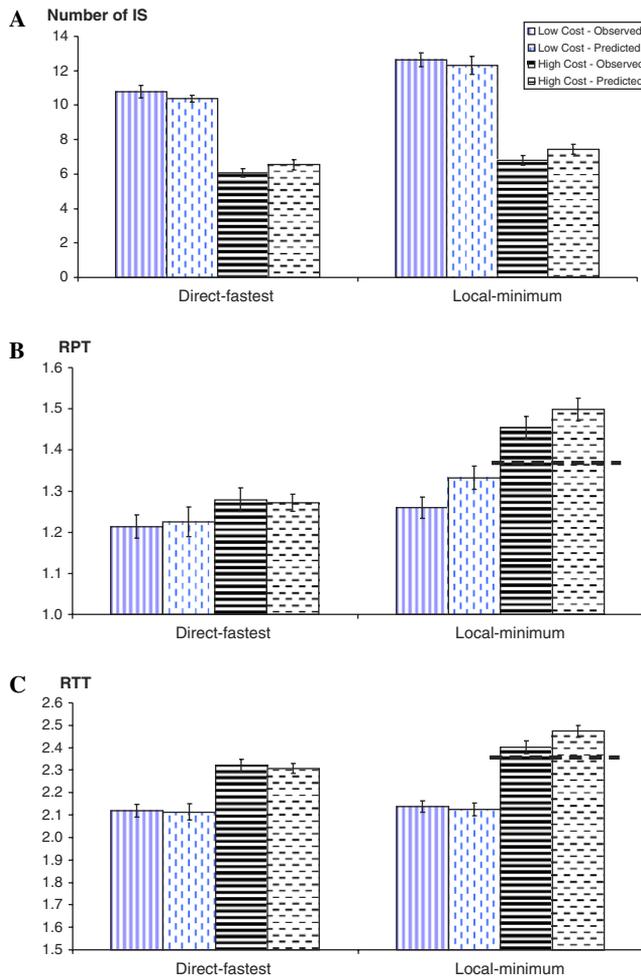


Fig. 15. (A) The number of information-seeking actions (IS), (B) relative path time (RPT), and (C) relative trial time (RTT) for each of the two map types in Experiment 3. Error bars represent standard errors of means. The error bars for the model are standard errors calculated from 10 iterations of model runs. The dashed lines represent the RPT and RTT when the number of IS was increased by 2 in the local-minimum maps in the high cost condition.

direct-fastest maps only in the low cost condition. The number of information-seeking actions between map types in the high cost condition was not significant. Subjects in the high cost condition seemed not able to distinguish between the two types of maps, and used similar numbers of information-seeking actions.

Fig. 15 also shows the model fit to the empirical data for the number of information-seeking actions in the two different map types. The overall fit was good, $R^2 = 0.90$, $RMSE = 0.07$. The model slightly under-predicted in the low cost condition and over-predicted in the high cost condition. The model, however, captured the overall pattern of the data. In the low cost condition, the number of information-seeking actions in the local-minimum maps was much higher than that in the direct-fastest maps. However, in the high cost condition, the number of information-seeking actions was roughly the same in the two kinds of maps.

10.2.2. Relative path time (RPT)

Fig. 15 shows the RPT in E3. Since for the direct-fastest maps, the hill-climbing path time equals the fastest path time, the original definition of RPT in E1 and E2 cannot be used (since the denominator will be zero). Instead, RPT is defined as: $RPT = \text{chosen path time}/\text{fastest path time}$. A 2×2 ANOVA shows that the main effect of cost is not significant, $F(1, 22) = 2.66$, $MSE = 3.85$, $p > .1$. The main effect of map types is significant, $F(1, 22) = 21.20$, $MSE = 2.79$, $p < .01$, and the cost \times map types interaction is also significant, $F(1, 22) = 7.26$, $MSE = 0.96$, $p < .01$. The difference in map types in the low cost condition is not significant, $F(1, 11) = 2.80$, $MSE = 0.47$, $p > .1$, but is significant in the high cost condition, $F(1, 11) = 39.45$, $MSE = 5.34$, $p < .01$. The difference between the high and low cost in the direct-fastest maps is not significant, $F(1, 22) = 0.74$, $MSE = 0.59$, $p > 0.3$, but is significant in the local-minimum maps, $F(1, 22) = 5.57$, $MSE = 5.73$, $p < .05$. Subjects used equally efficient paths in the direct-fastest maps, regardless of the cost. However, when the cost was high, subjects found significantly slower paths in the local-minimum maps than when the cost was low.

Fig. 15 shows the model fit to the data. The model fit the data well, $R^2 = 0.75$, $RMSE = 0.13$. The model found paths with similar quality in the direct-fastest maps, but over-predicted (i.e., found slower paths) in the local-minimum maps. The model also exhibited slightly larger variances than the human subjects. The model apparently did not consistently find the fastest paths in the low cost condition as the subjects did. On the other hand, the model apparently found fewer fast paths in the high cost condition.

10.2.3. Relative trial time (RTT)

Fig. 15 also shows the RTT (same definition as in E1 and E2) in E3. A 2×2 ANOVA yielded a significant main effect of cost, $F(1, 22) = 7.97$, $MSE = 13.38$, $p < .01$. The main effect of map types was not significant, $F(1, 22) = 2.38$, $MSE = 0.36$, $p > .1$, and the cost \times map types interaction was not significant, $F(1, 22) = 0.66$, $MSE = 0.11$, $p > .4$. The difference in map types in the low cost condition was not significant, $F(1, 11) = 0.42$, $MSE = 0.022$, $p > .4$, but was significant in the high cost condition, $F(1, 11) = 5.73$, $MSE = 0.39$, $p < .05$. The lack of effect of map types in the low cost condition suggests that, when using the local-minimum maps, subjects were able to find a path almost as fast as when using the direct-fastest maps. However, in the high cost condition, the overall performance in the local-minimum maps was significantly worse than that in

the direct-fastest maps. This was possibly because subjects spent less time seeking information, and obtained a worse path.

Fig. 15 also shows the model fit of the relative trial time. The model shows a good overall fit, $R^2 = 0.97$, $RMSE = 0.03$. Although the model fit the human subjects well in the low cost condition, it under-predicted in the direct-fastest maps, and over-predicted in the local-minimum maps in the high cost condition. The lower RTT in the direct-fastest maps in the high cost condition was probably because of the lower number of information-seeking actions of the model. The higher RTT in the local-minimum maps was probably because of the slower path the model used. Again, given the high constraints of the model, the fit was reasonably good.

One may wonder if subjects would have performed better if more information seeking were done. Given the reliability of the model across three studies and many measures, we turned to it for an answer to this question. To do so, we increased the expected utility of the *information-seeking-strategy* production in the high cost condition so that the mean difference of the number of information-seeking actions between the two types of maps was approximately 2 (note that mean difference in the low cost condition was approximately 1.85). In other words, we forced the model to perform more information-seeking and observed how well it performed. We found that with more information seeking, both the RPT and RTT were lower (see the dashed lines in Fig. 15). This suggests that the model found a faster path, and the time saved from the faster path was larger than the time spent on information-seeking. In other words, the subjects could have potentially attained much better overall performance if they had spent more time on information seeking. Our results showed that subjects had apparently stabilized at a suboptimal level of information-seeking.

10.2.4. Summary and conclusions of E3

Consistent with results from E1 and E2, when costs were high, subjects had fewer information-seeking actions. The key finding in E3 was the interaction between cost and map types. When the cost was low, subjects had significantly more information-seeking actions for the local-minimum than for the direct-fastest maps. As a result, the paths found when costs were low were of similar quality (i.e., similar relative path time) for both the local-minimum and direct-fastest maps. However, when the cost was high, subjects did not have more information-seeking actions for the local-minimum maps; as a consequence, the paths found were significantly slower paths for local-minimum than for the direct-fastest maps. Apparently when cost was high, subjects failed to adapt to the changing map types, and adopted a similar number of information-seeking actions in both types of maps.

The results of E3 highlight the importance of studying the course of adaptation. Many traditional cost-benefit analyses tend to extract an episode of behavior, calculate the costs and benefits of different alternatives, and infer the choices of the subjects based on the optimal cost-benefit tradeoffs of the existing alternatives (e.g., Beach & Mitchell, 1978; Christensen-Szalanski, 1978; O'Hara & Payne, 1998; Pirolli & Card, 1999). If this approach were adopted, performance of subjects in E3 would be considered irrational—because apparently non-perfect cost-benefit tradeoffs (that subjects failed to perform better in the local-minimum maps in the high cost condition) were observed.

It is possible that with longer practice (more than 40 trials), subjects might be able to learn the utility of information in the local-minimum maps in the high cost condition. However, three points lead us to believe that E3 performance is stable. First, the

interaction between costs and map types shows that subjects in the low cost condition did adapt to the task space in the trials provided. Second, in E2 our results showed that, with the right set of conditions, Subjects could and did change their information seeking behavior every 16 trials. Third, we found that the model had stabilized after 40 trials. In fact, we ran the model for an additional 40 trials and still did not find adaptation to local-minimum maps in the high cost condition. All in all, these points lead us to the conclusion that for the high cost condition, the cognitive system has stabilized on a suboptimal tradeoff.

The overall fit of the model was promising, with the model capturing the overall pattern of the empirical data for the three performance measures. Perhaps the most important aspect was that the model did exhibit different information-seeking behavior in different map types when the cost was low, but not when the cost was high. The behavior of the model provides a good account of this effect. In the local-minimum maps, the cost (the C parameter) of using the hill-climbing strategy will be much higher than in the direct-fastest maps. However, when the cost of information-seeking was high, information seeking stopped before the fastest path (the roundabout path) could be found because through interacting with the environment, the utility value of the production that initiated information-seeking actions became lower (because the C parameter was increasing) than the production that stopped the information-seeking strategy. The procedural learning mechanism therefore updated the utility values of the set of productions in such a way that competition between productions led to a low number of information-seeking actions. The low number of information-seeking actions was not sufficient to find the fastest path in the local-minimum maps, but was sufficient to find the fastest path in the direct-fastest maps. On the other hand, when the cost was low, more information-seeking actions were executed and the fastest paths were more likely to be found in both types of maps. Suboptimal tradeoffs were therefore realized by the conflict resolution and learning mechanisms of ACT-R, which updated the utility values of the information-seeking, hill-climbing, and the “give-up” productions according to the cost and information structures in different task environments. The set of productions, with different sets of utility values in different map types and costs, reflected the local decision rule as described in the BSM. The suboptimal tradeoffs found from the data therefore provide support for the BSM’s local decision rule and Bayesian learning mechanism.

11. General discussion

In the beginning of the article, we cast the problem of deciding when to stop seeking information as a process of adaptation to the environments. The Bayesian satisficing model (BSM) is able to obtain good level of performance in various environments through the global Bayesian learning process mediated by the local decision rule. The BSM, however, does not guarantee that the optimal tradeoff point can be found. The BSM predicts three related characteristics of information-seeking behavior that were supported by empirical data collected from human subjects. First, in a diminishing-return environment, the local decision rule was sufficient to lead to optimal performance. Second, in an environment where either the cost or information structure was changed, subjects responded to changes in costs faster than changes in utility of information. Third, when the cost was high in a local-minimum environment, subjects prematurely stopped seeking information and stabilized at suboptimal performance.

When faced with a new environment, a person may actively seek information to gain knowledge about the distal properties of the environment. Both our empirical and simulation results show that subjects used a local decision rule to decide when to stop seeking information. Although the local decision rule was effective in finding the right level of information seeking in most situations, the nature of local processing inherently limits the exploration of the environment. As we showed in Experiment 3, the local decision rule often implies “insufficient” information-seeking as high information-seeking costs discourage exploration of the environment. As a result, suboptimal performance emerges as a natural consequence of the dynamic interactions between cognition and the environment.

The combination of the conflict resolution, the procedural learning, and the credit assignment mechanisms of ACT-R were able to exhibit the same patterns of behavior across all three experiments in all dependent measures. The fact that the same model (only one parameter was different across the experiments) matched subjects’ behavior across the three experiments shows that these mechanisms were flexible enough to adapt the model’s behavior to large differences in its task environments. The result is that, across three studies, the model exhibited the same patterns of information-seeking behavior as the humans. Having verified the basic reliability and validity of the model, we then interrogated it to yield precise predictions of performance in the local-minimum environment if more information-seeking were performed and if more practice were given to the subjects. The ability of the models to generate these testable predictions is encouraging for the goals of cognitive psychology.

11.1. How persuasive are our good fits?

Roberts and Pashler (2000) argued against the overreliance on goodness of fit (GOF) as evidence to support theories, and suggested several other features of theory evaluation that are more important than GOF. We agree wholeheartedly. Below we examine our analyses in light of their three main principles of theory evaluation: (1) Flexibility—can the theory predict alternative results? (2) Data variability—do the data rule out what the theory rule out? (3) A priori likelihood of a good fit—can the theory fit any plausible result? We will examine our analyses based on these principles below.

11.1.1. Flexibility

The first principle focuses on how much a theory constrains the possible behavior of the model. For example, if a model has adjustable parameters, a particular good fit is only one example of what the model can produce. In general, ACT-R is complex and involves many parameters, and thus may seem able to fit a wide range of data. In fact, because of the interactions of the symbolic and subsymbolic representations, it is difficult to specify exactly how many parameters are free to vary in an ACT-R model (e.g., there are many possible symbolic representations for a given task). However, in our model, the symbolic representation is based on the BSM; and most of the subsymbolic parameters are set at their default values (as used in most of the models posted in the ACT-R website <http://act-r.psy.cmu.edu>). We did set 4 out of 35 parameters to obtain the fit in E1, but the parameter values were then fixed to fit data in both E2 and E3 (the only parameter we changed was the noise parameter to obtain the fit in E2). The parameter values were thus not “completely free” to vary to fit all the data across the three experiments.

A similar argument on the importance of the flexibility of a model to its validation is given by Pitt, Myung, and Zhang, 2002 (see also Pitt & Myung, 2002). They argued that

besides the number of parameters, the functional form is an important consideration when evaluating a model. A complex functional form will lead to high flexibility of the model to fit a wide range of possible patterns of data. Our model presented above rests on our theory that the adaptation process can be explained by the interaction of the Bayesian learning and the local decision components. The simulations of the BSM model clearly show that the simple functional form of the model is able to produce the different patterns of results obtained from the experiments. The ACT-R model is implemented based on the BSM, and thus has the same functional form. In fact, given the symbolic representation of the model as shown in Fig. 10 and Appendix C, none of the ACT-R parameter values affected the general functional form of the model. The good fits to the data are therefore not a result of an overly complex model.

Comparing results from E1 and E3, we found significantly different patterns of results: in E1, subjects adapted well to the different environments with different costs and utilities; however, in E3, subjects failed to adapt to the local-minimum environment when the cost was high. We showed that a single model, interacting with different environments, was able to produce the two different patterns of data by adjusting only the noise parameter of the model. The same model also produced the pattern of results in E2, in which the rates of approach to stable performance were faster when costs were changed than when utilities were changed. The good fits to the three sets of data show significant constraints to what the model *cannot* produce—for example, it cannot adapt to the local-minimum environment in E3 as well as it does to the diminish-return environments in E1.

11.1.2. *Data variability*

The second principle concerns the variability of the data on the constrained dimension that the theory predicts. The idea is to show how firmly the data agree with the predictions of the model. The major predictions of our model are focused on differences of the various dependent measures between the environments with different costs and utilities, and as we showed, all differences are statistically significant. Besides, we showed that the variability of the model closely matches that of the data in all sets of data. This shows that the data do seem to rule out what the theory rules out.

11.1.3. *A priori likelihood of a good fit*

The third principle rests on the well-known belief that a theory is useful only if it predicts unlikely, novel events. For example, a learning function, although correct, is not surprising if it merely predicts that performance improves according to a smooth function, reaches asymptotes, and stays there. Indeed, the likelihood of having a function that exhibits such characteristics is fairly high, and a good fit to the data does not gain much theoretical support for the function. However, if a theory predicts that performance stops improving and is able to specify the conditions under which this will happen, the likelihood of a good fit by chance will be much lower, and a good fit will be more meaningful. We argue that our theory does have a low a priori likelihood of good fits by chance to the three sets of data. Although it is not hard to imagine a simple learning function that improves and stabilizes at the level of information seeking as observed in E1, the finding that the improvement stops in a local-minimum environment when the cost is high cannot be easily predicted by a simple learning function, such as the power law of learning. In fact, we take the specific prediction and confirmation from E3 as strong evidence suggesting that a local decision process is used when subjects adapt to the cost and information structures of the task.

11.2. Relations to other suboptimal behavior

11.2.1. Suboptimal rules

The research presented in this paper falls into the general category of research that investigates how people learn from experience. It is obvious that most real-world decision-making behavior is action oriented; people have to choose what action to take to satisfy their goals, and they need to learn the degree to which their actions will lead to desirable or undesirable outcomes. Undoubtedly, much initial learning occurs by trial and error—i.e., one randomly chooses an option and observes the outcome. In our analyses, we assume that reinforcement from trial-and-error learning is the principal means to evaluate what actions to take in a particular environment. Indeed, we showed that that BSM does a good job characterizing the adaptation process of information-seeking behavior in different costs and utilities structures. However, one may speculate that in more complex situations people do inductively generalize action-outcome linkages and develop general rules to deal with specific instances. Learned values (i.e., costs and utility) of information seeking in one task environment may be transferred to a new task environment in which they may lead to the use of suboptimal rules (Einhorn, 1982). Indeed, we have found that suboptimal rules may persist even after years of experience with various software tools (Fu & Gray, 2004).

11.2.2. The sunk cost effect

The sunk cost effect refers to the higher tendency that one will continue an endeavor once an investment in money, effort, or time has been made. This tendency is often regarded as irrational in the normative standard, as decisions to continue or not should be based only on the marginal costs and benefits of future attempts. However, many studies find that the amount of resources spent on a task can have substantial impact on future decisions. For example, in a field study, Arkes and Blumer (1985) demonstrated that people who paid more for a 6-show theatre subscription package had a higher tendency to attend the shows. According to the normative standard of rationality, the likelihood to attend the shows should be the same once they had the tickets in hand, regardless of the price they paid for them. Apparently, the likelihood was higher for the group who had the higher sunk costs for the tickets.

In our task, subjects incurred continuous costs as they were seeking information until they reached the end station. This is analogous to a sunk cost situation: time already spent seeking information is the sunk cost; reluctance to use a direct path to the end station represents the desire to invest further in information-seeking. However, we argue that in our task the sunk cost effect is not as irrational as other examples of the sunk cost effect. The major difference of our task and other tasks is that continued information-seeking will increase the probability that a faster path will be found (see Fig. 2). There is therefore a prominent rational basis for continued information-seeking. In fact, many have argued that there exists a similar relationship between prior efforts and future likelihood of success in large information structures (Pirolli & Card, 1999) and problem-solving environments (Anderson, 1990). However, we do not intend to argue that the positive relationship between sunk costs and future likelihood of success is sufficient to explain the ubiquitous sunk cost effects. For example, in the theatre subscription package example discussed earlier, it is unlikely that people who paid more would anticipate that the shows would be more enjoyable. However, these situations may involve complex value functions of money and time that are obviously beyond the scope of this paper.

11.3. Implications

11.3.1. *Global and local control processes in choice behavior*

The good match of the BSM to the data suggests a mixture of global and local control of performance by trial-to-trial feedback. The model slowly adapts to the global characteristics (i.e. the relationship between number of information-seeking actions and overall performance time) environment by the Bayesian learning process, and decides to stop seeking information by a local decision rule. The BSM is in agreement with many choice models that treat adjustment of choice behavior in response to local samples of recent reward contingencies (Busemeyer & Myung, 1992; Davis, Staddon, Machado, & Palmer, 1993; Gonzalez-Vallejo, 2002; Herrnstein & Vaughan, 1980; Vaughan, 1981, 1985). However, our model assumes further that local samples of the environment are used to update the global knowledge of the environment through the Bayesian learning process. We also showed that compared to the local decision rule, the Bayesian learning process is slow and highly dependent on experienced samples of the environment. Although the model stabilizes at a suboptimal level of performance in a local-minimum environment, the process clearly has a prominent rational basis. In fact, as argued by others (Stephens & Krebs, 1987), for a forager in the natural environment, it makes more sense to pay attention to local changes of food patches than to the global trends of the environments.

11.3.2. *Implications for bounded rationality*

Although Bayesian learning is often associated with optimization, we showed that the BSM may not necessarily lead to the optimal level of performance. Rather, because of the use of a local decision rule, optimality is an exception, not a rule for the model. In fact, the use of a local decision rule leads to a stable level of performance, which is a necessary but not sufficient condition for optimality. Indeed, we share our views with others (e.g., Herrnstein, 1997) that stability provides a more general class of explanations than optimality. Since optimality implies stability, but not vice versa, we maintain that for models of bounded rationality, the level of explanation should not be more specific than the assumption of stability. In fact, as a mechanism of adaptation, the performance of our subjects and our models showed that the assumption of global optimization may be too constrained as a general guideline to explain human behavior.

Under the common assumption that cognition is well adapted to the environment to which it has evolved (e.g., see Anderson, 1990), the above arguments may lead to the question of why the optimal level of performance is not achieved, and what defines adaptivity. Our model takes the stance that adaptivity arises from the ability to take limited local information as controlling variables to myopically select the potentially “most profitable” actions. This myopic nature of adaptation, however, does not necessarily lead to globally optimal performance. On the other hand, we show that in environments with the general characteristic of diminishing returns, people are able to perform good tradeoffs between costs and utility, as we found in the simulation results of the BSM and from E1. We would like to think that cognition is evolved in natural environments where myopic processes are “good enough” to achieve reasonably high levels of performance such as our subjects did in E1. Unfortunately, we do not expect to obtain the data relevant to tracing the evolution of cognition that is needed to test this hypothesis.

The standard assumption that cognition is adapted to the most frequently encountered characteristics of natural environments, suggests that environments that result in suboptimal performance have characteristics that are not frequently encountered in nature. For better or worse today, more than ever, much of our daily environment is an artificial environment whose information-seeking characteristics may well be outside the normative range of most natural environments. Indeed, across the three studies, the empirical results and model predictions showed that the use of a simple local decision rule was able to achieve a reasonably good level of performance, except in environments with specific peculiarities, like the ones in E2 and E3. It is not too far-fetched to argue that these peculiarities stem from features of the task environment that are unrepresentative of the natural environment to which humans have evolved. In fact, it has been our goal to investigate how behavior deviates from optimality when people interact with environments with specific peculiarities, so that the underlying mechanisms can be revealed. We believe that the model provides a useful explanation of the patterns of behavior we observed—without the model we would have been forced to conclude that people are suboptimal in their information-seeking behavior. Instead, we conclude that cognition is adaptive—the simple, effective mechanisms are able to perform good tradeoffs in our “natural” environment and that complex computations are not necessary.

The models we constructed were based on our hypothesis that local information was utilized as a controlling variable in deciding the tradeoff point. This hypothesis was confirmed by the empirical data. When information-seeking costs were high, subjects in E3 stabilized at a suboptimal level of performance as predicted by the model. If global information were utilized as much as local information to decide the tradeoff point, subjects would have reached optimal performance even when information-seeking costs were high. Similar to the arguments made by many researchers studying human and animal choices, we conclude that bounded rationality is more likely to lead to a stable rather than an optimal tradeoff point.

11.3.3. Implications for interactive behavior

The results also contribute to recent research on interactive behavior, which stresses the importance of the study of the interaction of cognition, the environment, and the task (Byrne, 2001; Gray, 2000, *in press*; Kirsch & Maglio, 1994). Research on interactive behavior focuses on how internal cognitive processes are coupled with external physical activity in an interactive process. For example, Kirsch and Maglio (1994) showed that in a real-time interactive video game called Tetris, many manipulations are best understood as physical actions that complement complex computations in cognition. They argued that many process models of cognition often assume that information is already in the head, and have neglected the fact that people often use external physical actions to manipulate external objects to provide information to internal cognitive processes through our perceptual systems. Similarly, Gray and Fu (2004) showed that a small increase in information-seeking costs induced strategies that rely on imperfect knowledge in-the-head rather than accessing perfect knowledge in-the-world. The use of the suboptimal strategies was apparently a result of a local process that traded off information-seeking costs with memory retrieval costs. Similar to the studies in Gray and Fu, our manipulation of information-seeking costs (with or without the 1-s lockout time) was representative of that encountered in many human-system interactions. For example, people

often need to search for a particular command in the pull-down menu in a computer application, or they may need to use the “help” function to search for how a particular task can be done. In fact, information-seeking actions are ubiquitous in everyday tasks, in which we often find ourselves diving into the increasingly complex networked world of information and computer-mediated interaction. Our results show the importance of understanding the complex interaction between the task, cognition, and different information structures to facilitate the use of information. Specifically, in our experiments, we showed that even when useful information is lurking somewhere in the environment (e.g., an efficient shortcut in a computer application or a useful piece of information for decision-makers), people are not necessarily able to fully utilize the information to efficiently finish their tasks. Our results suggest the importance of the interaction between cognitive and perceptual-motor processes in understanding how information will be accessed and utilized to service the user’s goal.

Another aspect of interactive behavior is that it often involves a series of distributed choices on what actions to take. One of our recent studies (Fu & Gray, 2004) found that even after years of experience, users of a wide range of computer applications stabilized at a suboptimal level of performance. Carroll and Rosson (1987) call this phenomenon the “paradox of the active user.” Results from the current article suggest that the stable suboptimal level of performance could be a result of the localized process for choice of actions. As we observed in E3, when information-seeking costs are high, people tend to under-explore the task space. In other words, information-seeking costs may be affecting how likely people will be to use more efficient procedures in computer applications or other interactive environments. Our results therefore suggest an important warning for designers of interactive applications—even small differences in information-seeking costs may affect long-term learning and performance.

Acknowledgments

Support for this work was provided in part by grants to Wayne Gray from the Air Force Office of Scientific Research F49620-97-1-0353 and F49620-03-1-0143 as well as by the Office of Naval Research ONR #N000140310046. The current work is based on the dissertation of the first author. We want to thank Mike Oaksford, Raymond Nickerson, and John Wixted for their thoughtful comments on the manuscript.

Appendix A. Details of the Bayesian satisficing model (BSM)

A.1. Assumptions in the BSM

We make four general assumptions about the environment and the model to limit the number of variables without reducing its predictive power. By making these assumptions, we do not imply that all environments have the same kind of characteristics. Instead, we aim at deriving predictions of human behavior based on some general characteristics of environments. First, for the model, it is assumed that learning is a Bayesian combination of new and old information through each cycle of interaction with the environment. Second, we assume that the model uses a simple decision rule to stop seeking information based on existing knowledge of the environment. Third, for the environment, we assume that the more information the model has, the less time will be required to accomplish the task. Fourth, we assume that the task can always be finished by some heuristics (even with no

information-seeking action). However, depending on the number of information-seeking actions, the time required to finish the task varies. These assumptions allow us to focus on the study of how people behave in different environments with different relationships between costs and the utility of information.

The following variables are defined for the BSM:

- B —the mean execution costs required to finish the task
- n —number of information-seeking actions
- n' —number of information-seeking actions chosen by the model
- C —actual execution costs required to finish the task associated with a particular N

Learning in the BSM is concerned with estimating B from information observed from the environment. The model assumes that the execution costs required to finish the task uses the exponential function to represent the decreasing, diminishing-return relationship with the number of information-seeking actions (i.e., n). Research has shown that the exponential relationship fits the relationship well (e.g., see Pirolli & Card, 1999). The uncertainties of B are represented by a gamma distribution. The exponential distribution assumes minimal knowledge about the environment. The gamma distribution is a two-parameter (a and b in the following equation) general distribution that describes the uncertainties of B in a general environment:

$$p(B|a, b) = \frac{1}{\Gamma(a)b^a} B^{a-1} e^{-B/b}.$$

The likelihood function is assumed to be an exponential distribution

$$f(n|B) = \frac{A}{B} e^{-\frac{n}{B}},$$

where A is a proportionality constant and is assumed to be 1.

The gamma and exponential distribution are standard non-informative distributions in Bayesian analysis (Berger, 1985), which make minimal assumptions on the structure of the environment. Although the use of the exponential distribution assumes unrealistically that execution costs will approach zero as the amount of information increases, the decision criterion described below will almost guarantee that information-seeking will stop before the execution costs is too low. We do not claim that these distributions are able to describe characteristics of all possible task environments, but they are simple distributions that describe the general characteristic of $f(n)$ (i.e., decreasing, diminishing-return characteristic).

Mathematically, the distributions for B are assumed to be:

$$p(B|a, b) = \frac{1}{\Gamma(a)b^a} B^{a-1} e^{-B/b} \quad \text{s(Gamma distribution of the mean execution costs to solve the problem (B); } a, b \text{ are parameters of the distribution),}$$

$$f(n|B) = \frac{A}{B} e^{-\frac{n}{B}} \quad \text{(Exponential function of the execution costs to solve the problem, with the mean equal to } B, \text{ after } n \text{ information-seeking actions. } A \text{ is a proportionality constant.)}$$

From Bayes' theorem, the posterior probability density function of B given some observations of the outcomes of actions (i.e., the execution costs associated with n' steps of information-seeking) is:

$$p(B|n') = \frac{f(n'|B)p(B)}{f(n')} = \frac{f(n'|B)p(B)}{\int f(n'|B)p(B)dB} \quad \text{(Equation 1 : Bayesian learning equation)}$$

$p(B|n')$ will then be used as the updated "prior" estimation of the environment (i.e., $p(B)$) the next time the problem is solved. This allows the distribution $p(B)$ to be continually updated with new information.

With the definition of B , the function, $f(n)$, representing the amount of effort after n units of information-seeking can be calculated by:

$$f(n) = \int p(n|B')p(B')dB'$$

If w_n is the cost of a single information-seeking action, the optimal decision to stop seeking information is when the cost of an additional step of information-seeking exceeds its marginal utility of information. Mathematically, information-seeking will stop as soon as n satisfies the following equation:

$$f(n) - f(n + 1) < w_n \quad (\text{Equation 2 : local decision rule}).$$

To summarize, the prior distribution in the model represents the prior knowledge that people have of the relationship between the number of information-seeking actions and the execution costs required to finish the task in a general environment. With experience, the model gains information about the specific environment and updates its prior beliefs of $f(n)$. Based on this updated information, the number of information-seeking actions is decided. New observations from the environment can be obtained that allow a better estimation of when information-seeking will stop, and so on. This process allows the model to adapt progressively to the environment as more and more information is obtained from the environment. The BSM therefore provides a framework as to how people are able to seek information adaptively, based on limited knowledge about the environment.

Appendix B. ACT-R Architecture

ACT-R is a symbolic cognitive architecture with memory elements (i.e., chunks) in the declarative memory module and condition-action pairs (i.e., production rules) in the procedural memory module. ACT-R also has a subsymbolic level in which continuously varying quantities are processed, often in parallel, to participate in neural-like activation processes that determine the speed and success of access to chunks and production rules. These mechanisms are based on the “rational analysis” by Anderson (1990), in which Bayesian learning is one of the major assumptions of the analyses. Understanding of the models that we present requires knowledge of three of ACT-R’s subsymbolic components: the noisy conflict resolution mechanism, the procedural learning mechanism, and the credit assignment mechanism. These three mechanisms together determine the choice of different productions in different situations that are critical to the implementation of the BSM.

There are two major symbolic components in the ACT-R architecture (Anderson & Liebere, 1998)—a declarative knowledge component and a procedural knowledge component. Declarative knowledge corresponds to things that we know and can easily describe to others, and is represented as “chunks” in ACT-R. Each chunk has an activation value that determines the availability of the chunk. Goals are also represented as chunks that encode the system’s intentions. An example of a goal chunk, in which two has to be added to three and the answer is still unknown is shown below.

```
Goal-addition
ISA ADDITION
ADDEND1 TWO
ADDEND2 THREE
ANSWER NIL
```

Procedural knowledge is knowledge that we display in our behavior but of which we are not conscious. In ACT-R, procedural knowledge is represented as condition-action pairs, or production rules (or simply productions). Productions are in the form of IF ⟨condition – 1⟩ ⟨condition – n ⟩ THEN ⟨action – 1⟩ ⟨action – m ⟩. Procedural knowledge controls the operation of ACT-R’s buffers; for example, initiating retrievals from memory or shifts of attention on the action side and harvesting the results of these actions on the condition side. An example production rule that tries to solve an addition problem by retrieving an addition chunk is shown below. IF the goal is to add *num1* to *num2* and there is no answer THEN retrieve an addition chunk with *num1* and *num2* as addends

At any point in time, only a single production will fire. When there is more than one match, a mechanism called conflict resolution is used to decide which production to execute. The conflict resolution mechanism is based on a utility function. The expected utility of each matching production is calculated based on this utility function, and the one with the highest expected utility will be picked according to a mechanism referred to as conflict resolution, as described below.

B.1. Noisy conflict resolution

When there is more than one production rule that matches the goal specification, the noisy conflict resolution mechanism will decide which production to fire. Each production rule in ACT-R has a utility value E (referred to as “expected utility”) that determines the probability that the production will be selected. Since the conflict resolution mechanism is noisy, even the production with the highest utility is chosen only a certain proportion of time. Specifically, the probability that a production x will be selected is captured in this closed-form equation

$$\text{Probability of selecting } x = \frac{e^{E_x/t}}{\sum_j e^{E_j/t}},$$

where t (the expected gain noise) controls the noise added to the expected utilities. Noise is sampled from a logistic distribution with mean equals one and variance (σ^2) equals $\sigma^2 = \pi^2 s^2/3$, and $t^2 = 2s$. This equation is also known as the “softmax” or the Boltzmann’s machine equation in connectionist models (e.g., see Hinton & Sejnowsky, 1986). In general, the higher the value of t (or s , both are settable parameters in ACT-R), the higher the chance that productions with lower expected utility (E) will fire. This is an important characteristic because it balances exploration and exploitation behavior. In terms of the problem of choosing the number of information-seeking actions, the softmax equation allows sampling of the amount of effort required to finish the task for different numbers of information-seeking actions.

E is calculated as $E = PG - C$, where P is the expected probability that the goal will be achieved if that production rule is chosen, G is the value of the goal, and C is the expected cost of achieving the goal if the production is chosen. Similar to the BSM, the scale or currency for measuring utility in ACT-R is time. As described in the next section, the values of P and C will change with experience, allowing ACT-R to learn.

B.2. Procedural learning—a Bayesian learning mechanism

There are learning processes in ACT-R that produce statistical estimates of the appropriate quantities of two parameters—the probability of achieving the goal, P , and the cost of executing the production, C . P and C are calculated as

$$P = \frac{\text{successes}}{\text{successes} + \text{failures}} \quad C = \frac{\text{efforts}}{\text{successes} + \text{failures}}$$

The parameters successes and failures in both equations refer to the number of times the production has succeeded or failed to accomplish the goal. The parameter efforts is the total amount of time taken over all past uses of the production, successful or failed. In general, the chance of choosing a production increases when the number of times the production fires successfully is high, and the time taken over all past uses of the production is low. The formulae presented above were derived from a Bayesian estimation of the log odds of the probability that a particular action will successfully lead to its intended outcome (see Anderson, 1990 and Anderson & Lebiere, 1998).

The expected utility ($E = PG - C$) of a production is updated when the production fires and the outcomes of the production (i.e., success or failure and efforts) observed. This learning process in ACT-R is therefore similar to the Bayesian learning mechanism in the BSM. Both mechanisms learn by combining existing knowledge with new observations from the environment based on Bayes’ theorem. The expected utility of the productions that initiate information-seeking actions therefore reflects the function $f(n)$ in the BSM, which captures knowledge about the relation between the number of information-seeking actions and execution costs to finish the task. However, unlike the BSM, learning in ACT-R does not directly determine the number of information-seeking actions. Instead, the change in expected utilities of productions influences the probabilities that the productions will be chosen in the future. This introduces more variability in the learning process.

The above procedural learning mechanism, however, does not take into account the fact that people tend to weigh recent experiences more than distant experiences. ACT-R provides a time-based decay mechanism to the learning of expected value. This mechanism allows the influence of past experiences to decay with time. Thus, recent experiences will have a relatively larger influence on the production selection process. Consequently, an ACT-R model with a long history of past experiences can respond to recent experiences faster. The decay mechanism is important because the basic learning mechanism will respond to recent experiences more and more sluggishly when the number of older, past experiences increases (i.e., as the denominator gets larger and larger with experience).

The time-based decay mechanism has the same equations as the ones shown above, but successes and failures are calculated by

$$\text{successes}(t) = \sum_{j=1}^m t_j^{-d} \quad \text{failures}(t) = \sum_{j=1}^n t_j^{-d},$$

where t_j is defined as how long ago each past success or failure was. From the equation, the larger t is, the smaller the value will be weighted—i.e., the more remote a past experience is, the smaller its influence (i.e., discounting) will be to the current choice of actions. In other words, past experiences decay with time according to a power function with d as the index. The higher the value of d , the faster will be the decay of past experiences (when $d=0$, all past experiences will be equally weighted). This time-based decay mechanism allows faster response to new information, because recent information has a larger weight compared to old experiences.

B.3. Credit assignment mechanism

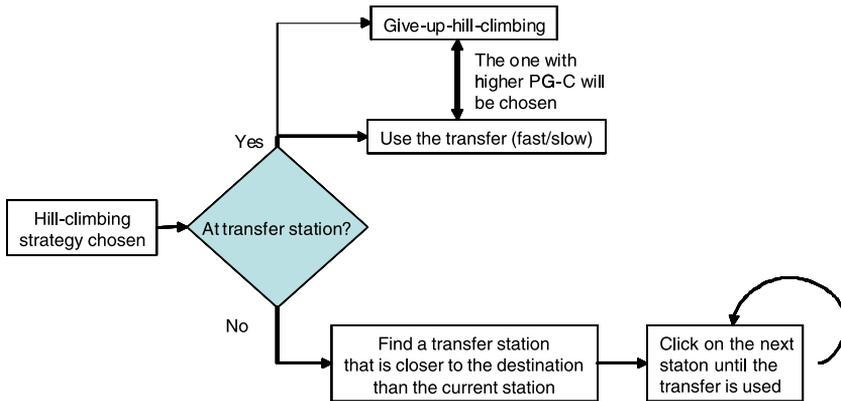
Another factor affecting the rate of learning is where the success and failure flag is set. The flags are set to tell ACT-R when the current goal is successfully accomplished and when the model fails to accomplish the goal. It can be seen that the later the flag is set, the larger will be the impact to the earlier productions (i.e., faster learning, see Appendix C). The credit assignment mechanism in ACT-R is therefore affected by where the flags are set. It determines how consequences of actions are credited to each of the production rules. In other words, the placement of the flags allows us to control when the consequences of different costs and utility of information will be fed back to the production rules. To preview our results, this mechanism is critical for the model to respond to costs faster than utility of information in the experiments.

Consider the example shown in the table below, where five productions, P_1 , P_2 , P_3 , P_4 , and P_5 were fired in succession. Assume that the initial number of successes and failures for P_1 is 1 and 0 respectively; the initial g parameter is set to 20; and the initial efforts parameter is set to 0.05 (all are default values). The change of the efforts parameter is shown in the second row. The initial utility value for P_1 will be 19.95 ($=1 \times 20 - 0.05$). If the success flag is set at P_3 , then after P_3 was fired, the procedural learning mechanism (without decay for simplicity) will update efforts of P_1 to $(0.15 + 0.05) = 0.2$. Since the number of successes is now 2, P remains 1 but C becomes 0.1 ($0.2/2$). Therefore the utility value will be updated to 19.9 ($1 \times 20 - 0.1$). However, if the success flag is set at P_5 , then the *efforts* parameter of P_1 will be updated to $(0.25 + 0.05) = 0.3$. The number of successes is still 2, P remains 1 but now C becomes 0.15 ($0.3/2$). The new utility value of P_1 will then be 19.85 ($1 \times 20 - 0.15$). It can be seen that the later the flag is set, the impact to the “upstream” productions (in this case, P_1 and P_2) will be higher (i.e., faster learning for the upstream productions).

Productions	P_1	P_2	P_3	P_4	P_5
Cumulative efforts	0.05	0.1	0.15	0.2	0.25

Appendix C. Details of the ACT-R model

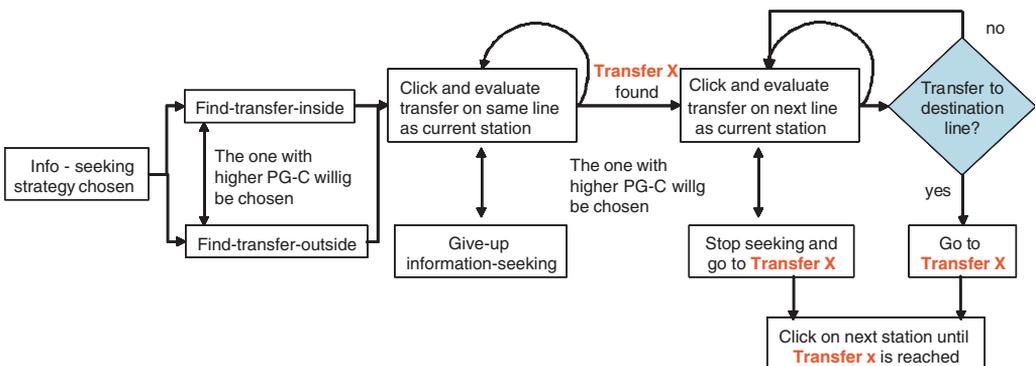
For the hill-climbing strategy, each box in the following figure represents a set of productions that implements the function as described. When the hill-climbing strategy is chosen, the model will look for and encode the transfer stations around (i.e., left, right, up, and down) the current station. Each transfer station is evaluated according to how much closer it is to the end station. If the model is at a transfer station, it will also consider using the transfer to change direction. If the model is not at a transfer station, it will go to the nearest transfer station between it and its end station. If it is at a transfer station, the model may decide to use the transfer (which could be either a fast or slow transfer) or to give up the hill-climbing strategy and use the information-seeking strategy. This decision is also a result of the conflict resolution mechanism, which is based on the utility values of the competing productions. The P and C parameters of these productions were updated through experience by the procedural learning mechanism. For example, the model will have learned that using a slow transfer has a high cost (i.e., an increasing C value, thus making the utility value lower). When the utility value of the production of using a slow transfer is lower than the production that gives up the hill-climbing strategy, the model may switch to the information-seeking strategy.



The credit assignment mechanism is controlled by a success flag after a transfer is used and a failure flag after the model gives up the strategy. The total cost (i.e., the C parameter) of executing the hill-climbing strategy is then credited to the initial production that chooses the strategy, and updates the utility value according to the procedural learning equations. The combination of the credit assignment and procedural learning mechanisms therefore influences the chance of selecting the hill-climbing strategy in the next cycle, in the same way as the Bayesian learning mechanism influences how frequently information-seeking actions will be chosen in the BSM.

C.1. The information-seeking strategy

For the information-seeking strategy, each box in the following figure represents a set of productions that implements the function as described. When the information-seeking strategy is chosen, the model will first look for transfer stations on the same line as the current station. There are two sets of productions that look for transfers to check. The first set of productions, *find-transfer-inside*, finds transfer stations that lie within the area enclosed by the current station and the end station (i.e., transfer stations that would be used if hill-climbing is used). The second set of productions, *find-transfer-outside*, finds transfer stations that lie outside the area enclosed by the current station and the end station. *Find-transfer-inside* has a higher initial utility value than *find-transfer-outside* (see table below). By doing this, we assume that initially subjects will tend to use a short path (which can be judged perceptually without deliberate information-seeking actions). As we will discuss later, this assumption is very useful when we model data from E3.



The procedural parameters of the critical productions in each of the two strategies

Productions/parameters	Successes	Failures	Effort	Utility
Hill-climbing-strategy	10	0*	0.05*	19.995
Give-up-hill-climbing	90	10	0.05*	17.9995
Use-train	500	0*	1	19.998
Use-fast-transfer	50	0*	0.5	19.99
Use-slow-transfer (high utility condition)	50	0*	16	19.68
Use-slow-transfer (med utility condition)	50	0*	8	19.84
Use-slow-transfer (low utility condition)	50	0*	2	19.96
Information-seeking-strategy	10	0*	0.05*	19.995
Find-transfer-inside (high cost condition)	50	0*	3	19.94
Find-transfer-inside (low cost condition)	50	0*	0.5	19.99
Find-transfer-outside (high cost condition)	40	10	3	15.94
Find-transfer-outside (low cost condition)	40	10	0.5	15.99
Stop-information-seeking	40	10	0.05*	15.995
Give-up-information-seeking	30	70	0.05*	5.9995

Critical productions with different parameter values in different experimental conditions are also shown. Values with “*” are the default values, so are the rest of the procedural parameters not shown in this table. Initial utility values are calculated based on the values shown in the table. With experience, these values change, and the change may affect utility.

Clicking on a transfer reveals its color (pink or orange) and the color codes the speed of transfer (pink fast, orange slow). This will continue until (1) a fast transfer is found, (2) another set of productions, *give-up-information-seeking*, is chosen, or (3) the model finds out that there is no fast transfer on the same line as the current station. Initially, *give-up-information-seeking* has a low utility value, but its chance of being chosen may increase through experience especially if the cost of information-seeking is high. When the information-seeking cost is high (i.e., the value of C increases through experience), the utility value of the set of productions that finds, clicks on, and evaluates a transfer will decrease, thus increasing the chance of giving up the information-seeking strategy.

When a fast transfer is found, the model will encode its location (labeled Transfer X) in the figure above and continue to evaluate transfers on the line to which the fast transfer leads. This will continue until (1) a fast transfer is found that leads to the line on which the end station lies, (2) the set of productions that stops information-seeking is chosen, or (3) the model cannot find any fast transfer. In cases (1) and (2), the model will move to the fast transfer on the same line as the current station (i.e., Transfer X) found earlier. In case (3), the model will give up the information-seeking strategy and use the hill-climbing strategy.

The following table shows where the flags were set. The flags specified where P and C will be calculated and all productions that were selected and led to the flagged production would be updated according to the procedural learning mechanism. Another cycle of credit assignment would then start after the flagged production. In the model, *success* flags were set after the train or the transfer was used. A *failure* flag was set after each “give-up” production. A failure flag has the effect of decreasing the P value of all productions selected that led to the flagged production, but the effect on the C parameter is the same as the effect of a success flag.

C.2. Relation of the ACT-R model to the BSM

The BSM combines the local decision rule and the Bayesian learning mechanism to decide on the number of information-seeking actions. In ACT-R, the local decision rule is implemented through the conflict resolution mechanism based on the expected utility of productions. The expected utility of a production is calculated as the difference between the expected gain and costs (i.e., $PG - C$), thus taking into account both the likelihood of reaching the end station and the costs associated with it. Similar to the BSM, the value of the C parameter is learned through each interaction with the task according to a Bayesian process. The credit assignment mechanism allows the actual cost of a series of actions to be credited to the productions that are responsible for initiating these actions. The experienced cost can therefore influence the chance of initiating the same series of actions in the future. For example, when the cost of information-seeking is high (i.e., C is high), the utility value of the productions that initiate the information-seeking action will

decrease. When the utility value of the production that initiates the information-seeking action becomes lower than the utility value of the productions that stop further information-seeking actions, the model will stop seeking information.

The tradeoff between costs and utility of information is reflected by the utility values of the productions that initiate and stop information-seeking actions, and those that initiate and stop the hill-climbing strategy (i.e., no information-seeking action). For example, when costs of information-seeking actions are high, the utility values of productions that initiate information-seeking actions will decrease, thus increasing the chance of selecting productions that stop information-seeking actions and productions that use the hill-climbing strategy. The number of information-seeking actions is then reduced. Similarly, when the utility of information is high (i.e., when the difference between the speeds of the slow and fast transfers is high), using the hill-climbing strategy to finish the task has a high cost, the utility values of the productions that initiate the hill-climbing strategy will decrease, thus increasing the chance of selecting productions that stop the hill-climbing strategy and productions that initiate information-seeking actions. The number of information-seeking actions is then increased. The interaction of the conflict resolution, procedural learning, and credit assignment mechanisms in ACT-R are able to model the tradeoff and implement the local decision rule and the Bayesian learning process in the BSM. Since the model can interact with the experimental software directly, the results from the model also allow quantitative measures that can be directly matched to the empirical data collected from human subjects.

The success and failure flag set in the model

Productions	Flag
Use-Train	Success
Use-Fast-Transfer	Success
Use-Slow-Transfer	Success
Give-Up-Hill-Climbing	Failure
Found-fast-transfer-to-end station-line	Success
Use-fast-transfer-after-information-seeking	Success
Give-up-information-seeking	Failure

The flags were the same the models for all three experiments.

References

- Anderson, J. R. (1990). *The adaptive character of thought*. Hillsdale, NJ: Erlbaum.
- Anderson, J. R., & Lebiere, C. (1998). *The atomic components of thought*. Hillsdale, NJ: Erlbaum.
- Arkes, H. R., & Blumer, C. (1985). The psychology of sunk cost. *Organizational Behavior and Human Decision Processes*, 35, 124–140.
- Ballard, D. H., Hayhoe, M. M., & Pelz, J. B. (1995). Memory representations in natural tasks. *Journal of Cognitive Neuroscience*, 7(1), 66–80.
- Beach, L. R., & Mitchell, T. R. (1978). A contingency model for the selection of decision strategies. *Academy of Management Review*, 3, 439–449.
- Berger, J. O. (1985). *Bayesian statistical decision theory* (2nd ed.). New York: Springer-Verlag.
- Brunswik, E. (1952). *The Conceptual framework of Psychology*. Chicago: University of Chicago Press.
- Busemeyer, 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(2), 177–194.
- Byrne, M. D. (2001). ACT-R/PM and menu selection: Applying a cognitive architecture to HCI. *International Journal of Human-Computer Studies*, 55, 41–84.
- Carroll, J. M., & Rosson, M. B. (1987). Paradox of the active user. In J. M. Carroll (Ed.), *Interfacing thought: Cognitive aspects of human-computer interaction* (pp. 80–111). Cambridge, US: The MIT Press.
- Christensen-Szalanski, J. J. J. (1978). Problem-solving strategies: A selection mechanism, some implications, and some data. *Organizational Behavior and Human Performance*, 22, 307–323.
- Davis, D. G. S., Staddon, J. E. R., Machado, A., & Palmer, R. G. (1993). The process of recurrent choice. *Psychological Review*, 100, 320–341.
- Einhorn, H. J. (1982). Learning from experience and suboptimal rules in decision making. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases* (pp. 268–286). New York, NY: Cambridge University Press.

- Fiedler, K. (2000). Beware of samples! A cognitive-ecological sampling approach to judgment biases. *Psychological Review*, *107*, 659–676.
- Fu, W.-T., & Gray, W. D. (2000). Memory versus perceptual-motor tradeoffs in a BlocksWorld task. In *Proceedings of the Twenty-second Annual Conference of the Cognitive Science Society* (pp. 154–159). Hillsdale, NJ: Erlbaum.
- Fu, W.-T., & Gray, W. D. (2004). Resolving the paradox of the active user: Stable suboptimal performance in interactive tasks. *Cognitive Science*, *28*(6), 901–935.
- Gigerenzer, G., & Todd, P. M. (1999). *Simple heuristics that make us smart*. New York, NY, US: Oxford University Press.
- Gonzalez-Vallejo, C. (2002). Making Trade-Offs: A probabilistic and context-sensitive model of choice behavior. *Psychological Review*, *109*, 137–155.
- Gray, W. D. (2000). The nature and processing of errors in interactive behavior. *Cognitive Science*, *24*(2), 205–248.
- Gray, W. D. (in press). The emerging rapprochement between cognitive and ecological analyses. In A. Kirlik (Ed.), *Adaptation in human-technology interaction*. New York: Oxford University Press.
- Gray, W. D., & Fu, W.-T. (2004). Soft constraints in interactive behavior: The case of ignoring perfect knowledge in-the-world for imperfect knowledge in-the-head. *Cognitive Science*, *28*(3), 359–382.
- Herrnstein, R. J. (1991). Experiments on stable suboptimality in individual behavior. *The American Economic Review*, *81*(2), 360–364.
- Herrnstein, R. J. (1997). *The matching law: Papers in psychology and economics*. Cambridge, MA: Harvard University Press.
- Herrnstein, R. J., & Prelec, D. (1991). Melioration: A theory of distributed choice. *The Journal of Economic Perspectives*, *5*(3), 137–156.
- Herrnstein, R. J., & Vaughan, W. J. (1980). Melioration and behavioral allocation. In J. E. R. Staddon (Ed.), *Limits to action: The allocation of individual behavior* (pp. 143–176). New York: Academic Press.
- Hinton, G. E., Sejnowsky, & T. J. (1986). Learning and relearning in Boltzmann machines. In D. E. Rumelhart, J. L. McClelland, & t. P. group (Eds.), *Parallel distributed processing: Explorations in the microstructure of cognition*. (Vol. Volume 1: Foundations, pp. 282–317). Cambridge, MA: MIT Press.
- Kirsch, D., & Maglio, P. (1994). On distinguishing epistemic from pragmatic action. *Cognitive Science*, *18*, 513–549.
- Mazur, J. E. (1981). Optimization theory fails to predict performance of pigeons in a two-response situation. *Science*, *214*, 823–835.
- O'Hara, K. P., & Payne, S. J. (1998). The effects of operator implementation cost on planfulness of problem solving and learning. *Cognitive Psychology*, *35*, 34–70.
- Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, *101*(4), 608–631.
- Oaksford, M., & Chater, N. (1996). Rational explanation of the selection task. *Psychological Review*, *103*(2), 381–391.
- Payne, J. W., Bettman, J. R., & Johnson, E. J. (1993). *The adaptive decision maker*. New York: Cambridge University Press.
- Pirolli, P., & Card, S. (1999). Information foraging. *Psychological Review*, *106*(4), 643–675.
- Pitt, M. A., & Myung, I. J. (2002). When a good fit can be bad. *Trends in Cognitive Sciences*, *6*(10), 421–425.
- Pitt, M. A., Myung, I. J., & Zhang, S. (2002). Toward a method of selecting among computational models of cognition. *Psychological Review*, *109*(3), 472–491.
- Roberts, S., & Pashler, H. (2000). How persuasive is a good fit? A comment on theory testing. *Psychological Review*, *107*, 358–367.
- Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological Review*, *63*, 129–138.
- Simon, H. A. (1996). *The sciences of the artificial* (3rd ed.). Cambridge, MA: MIT Press.
- Stephens, D. W., & Krebs, J. R. (1987). *Foraging theory*. Princeton, NJ: Princeton University Press.
- Stigler, G. J. (1961). The economics of information. *The Journal of Political Economy*, *69*, 213–225.
- Vaughan, W. Jr., (1981). Melioration, matching, and maximization. *Journal of the Experimental Analysis of Behavior*, *36*, 141–149.
- Vaughan, W. Jr., (1985). Choice: A local analysis. *Journal of the Experimental Analysis of Behavior*, *43*, 383–405.