Information search with situation-specific reward functions

Björn Meder*   Jonathan D. Nelson*

Abstract

The goal of obtaining information to improve classification accuracy can strongly conflict with the goal of obtaining information for improving payoffs. Two environments with such a conflict were identified through computer optimization. Three subsequent experiments investigated people's search behavior in these environments. Experiments 1 and 2 used a multiple-cue probabilistic category-learning task to convey environmental probabilities. In a subsequent search task subjects could query only a single feature before making a classification decision. The crucial manipulation concerned the search-task reward structure. The payoffs corresponded either to accuracy, with equal rewards associated with the two categories, or to an asymmetric payoff function, with different rewards associated with each category. In Experiment 1, in which learning-task feedback corresponded to the true category, people later preferentially searched the accuracy-maximizing feature, whether or not this would improve monetary rewards. In Experiment 2, an asymmetric reward structure was used during learning. Subjects searched the reward-maximizing feature when asymmetric payoffs were preserved in the search task. However, if search-task payoffs corresponded to accuracy, subjects preferentially searched a feature that was suboptimal for reward and accuracy alike. Importantly, this feature would have been most useful, under the learning-task payoff structure. Experiment 3 found that, if words and numbers are used to convey environmental probabilities, neither reward nor accuracy consistently predicts search. These findings emphasize the necessity of taking into account people's goals and search-and-decision processes during learning, thereby challenging current models of information search.

Keywords: information search, classification, optimal experimental design, payoffs, decisions from experience.

1 Introduction

When diagnosing and treating a patient, when choosing a job candidate or a mate, and in many other situations, one must make decisions without having all the relevant information. Are there widely applicable strategies for identifying useful queries? What governs people's information search? In information-acquisition situations where no particular benefits and costs apply, statistical *optimal experimental design* (OED) models provide one framework for evaluating the value of alternative queries (Fedorov, 1972; Good, 1950; Lindley, 1956; Myung & Pitt, 2009). A variety of experiments suggest that such models can also provide a reasonable description of human information-search behavior (reviewed by Nelson 2005, 2008), in situations with no explicit external payoffs.

But in many situations—for instance, when deciding whether something is safe to eat, whether a suspicious suitcase contains a bomb, or whether a patient should be sent to the cardiac care unit—strong asymmetries in consequences for particular correct or incorrect decisions apply. These asymmetries have implications for classification decisions. For instance, a potential cardiac patient should be sent to the cardiac care unit if the risk of heart attack is greater than some threshold. The threshold should clearly be less than a 50% chance of heart attack, which would be the threshold for maximizing overall classification accuracy.

Are there implications, for information search, of varying benefits and costs for different kinds of correct and incorrect decisions? Intuitively, it may seem that the best strategy is to conduct queries that allow determination of the true state of nature as accurately as possible, and to take payoffs into account only in the actual classification decision. That intuition is distinctly wrong. As our theoretical analyses and simulations illustrate, situation-

specific costs and benefits must be considered when determining which information to acquire (which test to conduct, which question to ask, which query to make), and not only when making classification decisions. In other words, many highly informative tests are useless, given the applicable situation-specific reward structures.

To what extent does people's information-search behavior appropriately reflect the costs or benefits associated with different correct or mistaken decisions? We address this in three experiments, using a probabilistic multiple-cue category learning and information-search paradigm. Importantly, we identify and investigate situations in which the goal of obtaining information that helps to maximize the number of correct classification decisions should in principle lead to different search behavior than is appropriate to maximize reward, given the particular environment's reward structure. (The environmental reward structure consists of the payoffs for each kind of correct classification decision, and the costs of each kind of incorrect classification decision.) We examine whether people's information-search behavior appropriately reflects situation-specific reward structures, or whether people may use *probability gain* (accuracy maximization, a psychologically plausible goal for information search in classification tasks, Baron, 1981, as cited in Baron, 1985; Nelson, McKenzie, Cottrell, & Sejnowski, 2010) to guide their search decisions, even when it is not adaptive to do so.

1.1 Overview of paper

We first review research on general-purpose methods for identifying useful informational queries, and empirical studies on human information selection in which no external payoffs apply. We then discuss related existing empirical research in psychology, investigating behavior in situations with different reward structures. This research shows some of the capabilities and limitations of human behavior in maximizing payoffs following experience-based learning, on non-information-search tasks. Some prior research also deals with information search given asymmetric rewards, but not in the context of experience-based learning, and not in contexts in which reward and accuracy make contradictory predictions.

We then introduce the mathematics that should in principle govern classification behavior in environments with external payoffs, and show how situation-specific payoffs can be incorporated into models of information search. Building on these equations, we use computer simulations to identify environments in which it is highly problematic to use probability gain (i.e., selecting information so as to maximize the probability of making a correct classification decision) to identify queries, if the goal is to maximize expected reward. We then report three experiments on human information-search behavior, in those environments. These experiments were designed to identify whether people use probability gain more broadly than it is adaptive to do so, and what conditions can facilitate identification of the most useful queries, given situation-specific reward functions. Finally, we discuss the implications of our findings in relation to Bayesian decision-theoretic models of human cognition, existing models of the value of information, and real-world information-search tasks.

2 Models of information search

How can one anticipate the usefulness of possible informational queries (questions, tests or experiments), before the answer (query result, experiment result, or test outcome) is known? In information-acquisition situations on classification tasks where no particular benefits and costs apply, optimal experimental design (OED) models[1] provide one framework for evaluating the value of alternative queries (Nelson, 2005, 2008; Appendix, Table A1). Mathematically, these models fall within a framework of expected utility maximization (Savage, 1954), where utility is defined according to a particular quantification of the value of information. Some OED models include maximizing improvement in probability of identifying the correct hypothesis, or category (*probability gain*, Baron, 1981/1985), maximizing change in beliefs (e.g., *Kullback-Leibler divergence*, Kullback & Leibler, 1951; or *impact*, Wells & Lindsay, 1980), and minimizing uncertainty (as measured with *Shannon entropy* or a related measure, Shannon, 1948; Lindley, 1956). OED models can in some cases themselves be exactly implemented by heuristic processes (Navarro & Perfors, 2011; Nelson, 2005, 2008, 2009). A number of heuristic algorithms, outside the mathematical framework of utility maximization, have also been proposed (Gigerenzer, Todd, & ABC Research Group, 1999; Green & Mehr, 1997; Luan, Schooler, & Gigerenzer, 2011; Martignon, Katsikopoulos, & Woike, 2008).

Some of these models have been proposed as normative and/or descriptive models in psychology. For example, Baron, Beattie, and Hershey (1988) used probability gain as a normative model in a medical test scenario.

---

[1]Related terms include *epistemic utility* and *quasi utility* (Good, 1950). Note that the idea with "optimal" experimental design models is not that they are globally optimal for sequential search. Indeed, computational constraints usually require planning only a limited number of steps into the future. Nor are OED models intended to optimize known external constraints. Rather, OED models are a statistical attempt to elucidate reasonable bases for selecting experiments, queries, or tests when external utilities do not apply or are not specifically known. This might be the case on perceptual or other categorization tasks, and scientific inference tasks. Various names of particular models have been used. We follow Nelson's (2005, 2008) nomenclature.

Crupi, Tentori, and Lombardi (2009) used probability gain in an analysis of the pseudo-diagnosticity paradigm. Oaksford and Chater (1994, 1996) used information gain to analyze Wason's (1966, 1968) selection task. Klayman and Ha (1987) used impact in their research on hypothesis testing (Klayman, 1987, also used information gain). Most of these studies focused on tasks where no particular benefits and costs apply. (Baron & Hershey's, 1988, medical diagnosis task, and Oaksford & Chater's, 1994, model of deontic versions of Wason's selection task, are notable exceptions.) This research has shown that information search that may look irrational from the viewpoint of classical deductive logic may make sense from the perspective of adaptively seeking information to facilitate probabilistic inductive inference (Chater & Oaksford, 2008; Hahn & Oaksford, 2007; Oaksford & Chater, 2007). Other work (Myung & Pitt, 2009; Cavagnaro, Myung, Pitt, & Kujala, 2010) shows how OED principles can be used to automatically design experiments to investigate different aspects of human cognition and discriminate between competing cognitive models (e.g., the shape of memory decay curves).

Which model best describes human intuition about the usefulness of possible informational queries, on classification tasks where no particular external payoffs apply? Nelson et al. (2010) used an experience-based category learning paradigm to pit alternate models against each other, in several experiments. Probability gain best described human information search behavior when environmental probabilities were learned through experience and no particular costs or benefits for different types of decisions applied. Subjects preferentially viewed the higher-probability gain feature in all the environments studied by Nelson et al., even in cases when all the other OED models (Appendix, Table A1) preferred a different query.

3　Prior research

Previous research does not address information search in the context of asymmetric reward functions, where the two goals of obtaining information for improving classification accuracy, and maximizing reward, contradict each other. Previous research addresses related issues, however. These issues include:

1. the circumstances under which classifications of stimuli can adapt to asymmetric payoff structures, in non-search tasks;

2. people's ability in rapid motor movement tasks, which do not involve information acquisition, to spontaneously adapt to asymmetric payoff structures; and

3. the factors driving people's choices among queries, in environments with asymmetric payoffs, but in which the alternate queries are objectively equally useful.

Our experiments bring together these research areas in novel ways.

3.1　Perceptual categorization, asymmetric rewards, and signal detection theory

Several studies on perceptual categorization have used a signal-detection-theory framework to examine how people's decision criteria vary as a function of the costs and benefits associated with different categorization decisions (Maddox, 2002; Maddox & Bohil, 1998, 2003; Maddox & Dodd, 2001; von Winterfeldt & Edwards, 1982). Typically, these categorization experiments present subjects with stimuli (e.g., lines of varying length), randomly sampled from two overlapping category distributions. The task of the subject is to categorize a given stimulus (e.g., as "short" vs. "long") in a way that maximizes accuracy or expected reward.

Maddox and Bohil (1998, 2001; see also Maddox, 2002; Maddox & Dodd, 2001) examined how people's categorization decisions vary as function of asymmetric payoffs. For example, if correct Category x responses are rewarded twice as highly as correct Category y responses, the expected value of a Category x response is higher than the expected value of a Category y response whenever P(Category x | stimulus) > 1/3. (The section on "Decision bounds in binary classification tasks," below, provides a formal treatment.) By contrast, to maximize accuracy, one must always choose the more likely category, meaning to predict x whenever P(Category x | stimulus) > 1/2. To maximize long-run expected reward, one must sometimes choose the less likely category, which necessarily leads to a higher number of incorrect classification decisions than if always choosing the more likely category. Thus, asymmetrically rewarded classification decisions induce a conflict between accuracy and reward.

Do people adopt decision criteria so as to maximize reward, in classification tasks? Typically, people shift their decision criterion away from 50% in the appropriate direction, but not as much as would be optimal from the perspective of expected reward. To account for these and related findings, Maddox and Bohil (1998) introduced the *COmpetition Between Reward and Accuracy* (COBRA) hypothesis. Since under asymmetric payoffs the reward- and accuracy-maximizing decision bounds conflict, the resulting criterion placement is suboptimal (for an overview, see Maddox, 2002). Thus, even given experience-based learning of the overlapping category distributions, when reward and accuracy conflict, peo-

ple's categorization decisions can fail to follow a reward-maximizing strategy. Note, however, that this line of research (and the COBRA model) is not concerned with information search, but rather with classification learning based on full stimulus information.

3.2 Rapid motor movements under uncertainty and spontaneous reward maximization

Another research program that uses experience-based learning to convey probabilistic information, and studies people's ability to spontaneously behave appropriately in the context of situation-specific reward structures, involves rapid motor movement tasks. Trommershäuser, Maloney, and Landy (2003a,b) introduced tasks that involved rapid pointing to a touch screen, with a payoff for hitting a reward region (e.g., a green circle), and a penalty for hitting a penalty region (e.g., a red circle). Because of the small size and close spatial (or overlapping) location of the reward and penalty regions, and the motor uncertainty in the rapid movement, it is generally not optimal to aim at the center of the reward region. Rather, the intended reach location should be appropriately shifted (according to the motor uncertainty in the rapid reaching movement), to maximize the expected payoff, aggregating the expected reward (amount times probability) from the hit region, minus the expected penalty (amount times probability) from the penalty region(s).

These motor tasks are mathematically equivalent to more traditional decision making under risk tasks studied in psychology (e.g., choices between gambles). Each pointing trial is the choice of a gamble, defined by the intended pointing location. The possible outcomes of a motor action (gamble) are determined by the probability, given the motor uncertainty, of actually touching each particular region (each reward, penalty, or overlap region) on the screen, and that region's associated payoffs. Trommershäuser et al. (2003a,b) found that people were close to optimal (typically earning 95+% of the theoretical maximum returns) in movement tasks, with monetary reward. In view of other psychological research on decision making under uncertainty, and the fact that the relevant probabilities were never explicitly conveyed to subjects, this is remarkable (Trommershäuser, Maloney, & Landy, 2008).

Note also that subjects' capacity to apply a reward-maximizing strategy did not require new learning experiences, or gradually shifting behavior, after the payoff scheme was introduced. Prior to the actual decision-making phase, in which the payoff scheme was imposed, subjects underwent a motor-task training phase, in which they internalized motor uncertainty without explicit payoffs for correct or incorrect pointing movements. In

the motor training phase, subjects had to reach quickly, and could learn how accurate they are in hitting a particular target location in the available time. After this experience-based learning of motor uncertainty, the reward function was introduced. At this point, subjects immediately adapted their movements to approximately maximize reward. These findings suggest that when probabilistic information (the motor uncertainty of the reach destination under time pressure) is properly internalized, people are able to adapt their behavior to maximize reward.

What are the limits to this kind of reward-maximizing behavior? Wu, Trommershäuser, Maloney, and Landy (2005) noted that in Trommershäuser et al.'s (2003a, 2003b) tasks, the optimal reach destination was always somewhere along a single imaginary line, about which the reward and penalty regions were symmetric. If a subject noticed and correctly intuited this symmetry, they could reduce the decision space from two dimensions to a single continuous dimension. (Wu et al. called this the *symmetry-axis heuristic*). Wu et al. (2005) considered a slightly more complex scenario, in which there were two penalty regions, which differed in their severity. In this scenario, the optimal movement endpoint goal was not on the imaginary symmetry line. Rather, the optimal movement goal was slightly within the lesser-penalty region, in a location that overlaps with the reward region. Although performance was high, the distributions of most subjects' reaches were shifted significantly away from the optimal locations, suggesting that this task is more difficult than the earlier maximum expected gain tasks.

To summarize, Trommershäuser and colleagues' research suggests that, following experience-based learning to internalize uncertainty in motor movement tasks, people have a remarkable, but not perfect, ability to take arbitrary payoff functions into account, without requiring new learning experiences. The conditions for optimality in the motor movement plans—namely, why Trommershäuser et al. (2003a,b) saw it, yet Wu et al (2005) did not—are not fully clear.

3.3 Information search under asymmetric reward functions

Whereas standard classification tasks require people to categorize items based on the full stimulus information, search tasks require people to consider which query (test, question, experiment) is most useful to achieve a certain goal (such as maximizing reward or accuracy) *before* making a classification decision. For each possible test outcome (e.g., for a positive or negative medical test result), such a task requires estimating its marginal probability, its implications (i.e., posterior probability that a person does or does not have the disease, given a par-

ticular test result), and its usefulness with respect to the agent's goals. For instance, it may be more important to correctly identify people who have a disease than those who do not have the disease, or vice versa, or to maximize overall classification accuracy.

Baron and Hershey's (1988) study of information acquisition included some problems in which the choice was whether or not to conduct a medical test, and other problems in which the choice was which of two tests to conduct. In each particular scenario, subjects were given information on the probability of a disease, and one or two tests. The information on each test was presented as the test's true positive rate (the probability of a positive test result in a patient who has the disease), and the false positive rate (the probability of a positive test result in a patient who does not have the disease). There were no explicit payoffs for correct diagnoses. The reward structure was described in terms of the cost (harm) of treating a person who does not have the disease, and the cost (neglect) of failing to treat a person who does have the disease. Both symmetric and asymmetric cost structures were used. In the symmetric cost structure, failing to treat a person with the disease was equally problematic as unnecessarily treating a healthy person. In the asymmetric cost structure, one of the kinds of errors had a greater cost, which was specified, than the other kind of error.

In some scenarios (Experiment 1, Cases 5–11), the task was to choose between two tests. The idea was to try to identify the cues that people use to select useful tests. Baron and Hershey found that subjects were often sensitive to normatively relevant variables, including prior probabilities of the disease, the difference between true and false positive rate (which corresponds to a test's impact; Nelson, 2005, 2009), and the applicable cost structure. Baron and Hershey also found that subjects use heuristic strategies, for instance by choosing a test that, relative to a given cost structure, minimizes the most harmful kind of errors.

However, because the tests were objectively equally useful (i.e., had the same utility given the applicable cost structure), subjects' preferences do not directly show how sensitive people are to the relative objective usefulness of different tests. In about half of the cases, subjects did not have a statistically reliable preference for one test over the other. In other cases, subjects did have a preference between the tests. Why? Some subjects' written justifications indicated that they had at least approximately calculated the relevant probabilities and utilities, and realized that the tests were roughly equally useful. (This does not explain why subjects preferred particular tests, in other scenarios.) Another possibility is that people used informational strategies, consistent with OED models, to pick queries. To address this, we re-analyzed these scenarios. However, no OED models consistently predicted behav-

ior (see Appendix, Table A2).

Still another possibility is that subjects have difficulties understanding and utilizing probabilistic information when it is presented with words and numbers, in this type of task. Traditional words-and-numbers formats are not very meaningful for inductive inferences (Cosmides & Tooby, 1996; Gigerenzer & Hoffrage, 1995; Krauss, Martignon, & Hoffrage, 1999), and lead to less-consistent search behavior than experience-based learning (Nelson et al., 2010). Baron and Hershey found evidence that subjects may have used particular heuristic strategies that could be applied given the words-and-numbers format of the probability and cost (reward structure) information.
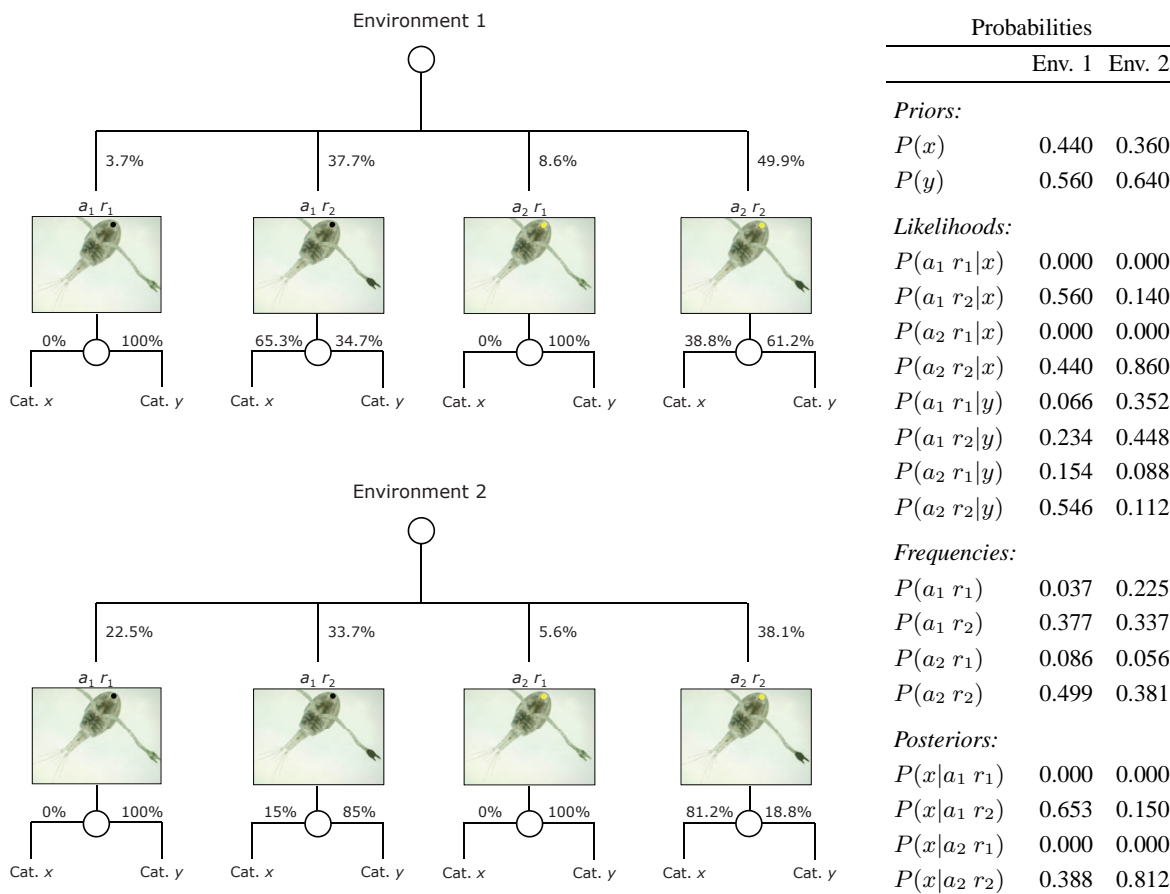
4   Aims of this paper

Previous research on classification decisions and on rapid movement tasks under asymmetric rewards gives a complicated picture regarding the circumstances under which people's behavior, following experience-based learning, can respond appropriately to externally imposed payoff structures. Moreover, this research does not examine information search. Although Baron and Hershey (1988) did study information acquisition under asymmetric payoff structures, they did not use experience-based learning of environmental probabilities, and they did not study circumstances under which the goal of gathering information to improve accuracy contradicts the goal of maximizing some external payoff function. Thus, it is hard to predict from the literature whether people will be able to identify the most useful tests, after learning environmental probabilities through experience, when situation-specific payoffs apply. The present paper has twin goals:

1. theoretically identify circumstances under which it is important to take situation-specific payoff schemes into account when searching for information, and
2. empirically address the circumstances under which people can take situation-specific payoffs into account when searching for information.

Our theoretical analyses integrate ideas from optimal experimental design, statistical decision theory, and signal detection theory. We use computational search techniques to identify environments in which searching with the goal to obtain the most reward is maximally incompatible with the goal of making accurate classification decisions. Empirically, we use experience-based learning and actual information-search tasks to address whether human subjects' information-search behavior can appropriately make use of situation-specific (symmetric and asymmetric) reward structures. We conduct these experiments in the environments identified through our com-

Figure 1: Statistical environments to differentiate usefulness of features A and R under symmetric vs. asymmetric reward functions. In each environment, there are four stimuli ('plankton'), constructed by combining two binary features, A ("eye") and R ("claw"). The numbers above the items indicate their frequencies; the numbers below indicate the probability of belonging to Category x or y, respectively. The table at right provides detailed information on the two environments.

Probabilities	Env. 1	Env. 2
Priors:		
$P(x)$	0.440	0.360
$P(y)$	0.560	0.640
Likelihoods:		
$P(a_1\,r_1|x)$	0.000	0.000
$P(a_1\,r_2|x)$	0.560	0.140
$P(a_2\,r_1|x)$	0.000	0.000
$P(a_2\,r_2|x)$	0.440	0.860
$P(a_1\,r_1|y)$	0.066	0.352
$P(a_1\,r_2|y)$	0.234	0.448
$P(a_2\,r_1|y)$	0.154	0.088
$P(a_2\,r_2|y)$	0.546	0.112
Frequencies:		
$P(a_1\,r_1)$	0.037	0.225
$P(a_1\,r_2)$	0.377	0.337
$P(a_2\,r_1)$	0.086	0.056
$P(a_2\,r_2)$	0.499	0.381
Posteriors:		
$P(x|a_1\,r_1)$	0.000	0.000
$P(x|a_1\,r_2)$	0.653	0.150
$P(x|a_2\,r_1)$	0.000	0.000
$P(x|a_2\,r_2)$	0.388	0.812

puter simulations, in which the goals of obtaining information to be accurate, and obtaining information for reward, maximally conflict.

Experiments 1 and 2 each consist of two tasks: a classification-learning task and a search task. In the learning task, both feature values of the stimuli were visible in every trial. The learning-task procedure was very similar to that of Nelson et al. (2010), and other multiple-cue probabilistic category learning tasks (e.g., Knowlton, Squire, & Gluck, 1994; Kruschke & Johansen, 1999). It was also similar to perceptual categorization classification tasks in the signal detection theory paradigm (e.g., Maddox, 2002). Figure 1 illustrates the category-learning task at a more conceptual level; Figure 4 illustrates a sample trial from the category-learning task. The goal of the learning task was to help subjects internalize environmental probabilities. Stimuli consisted of two dichotomous features: A (e.g., the eye feature), which can take values a_1 and a_2, and R (e.g., the claw feature), which can take values r_1 and r_2. Subjects' task was to classify stimuli as Category x or Category y, as a function of which stimulus was shown (a_1r_1, a_1r_2, a_2r_1, or a_2r_2).

Our research question concerned behavior in the subsequent search task. In this task, subjects could view only a single feature (Feature A or Feature R), before classifying stimuli. Figure 2 illustrates the decision problem that the search task presents, at a conceptual level. Once a feature is chosen to view, the specific value that it takes is revealed (Feature A can take values a_1 or a_2; Feature R can take values r_1 or r_2), according to the environmental probabilities. Based on this information, subjects had to classify the item as Category x or y.

The crucial manipulation concerned the monetary payoffs for correct classification decisions in the search task: in the *symmetric payoffs condition*, each type of correct classification paid the same amount of money (2€ for any

correct classification). In this case, maximizing accuracy will also maximize rewards, and Feature A is most useful. In the *asymmetric payoffs condition*, correct classifications of one category received a higher payoff than correct classifications of the other category (e.g., 2€ for correct Category x classifications vs. 0.2€ for correct Category y classifications). In this case, maximizing overall classification accuracy would not maximize rewards, and Feature R is more useful. This manipulation was designed to mimic the asymmetric rewards inherent in real-world scenarios, such as medical screening, in which it is more important to correctly identify patients who have a disease than those who do not have the disease.

Mathematically, the search task is very different from the learning task, as it requires people to determine which of the two features (A vs. R) would be most useful, relative to their goals (e.g., the applicable payoff function or an intrinsic goal, such as maximizing accuracy), *before* seeing the specific feature value or making a classification decision. Choosing a feature to view requires anticipating the usefulness of the possible outcomes of the search (e.g., usefulness of a_1 vs. a_2, in the case of Feature A), and aggregating the usefulness of each possible feature value according to its probability, in order to determine the aggregate usefulness of the feature. (Raiffa and Schlaifer, 1961, would call this a *preposterior analysis*.) Importantly, the usefulness of the individual feature states depends on the goal, such as maximizing accuracy or reward.

Note that the classification part of the search task is mathematically also very different from that of the learning task. For example, a subject may decide to search Feature A, and observe that it takes the value a_1. Given this single piece of information, the subject has to estimate the probability of the categories, e.g., P(Category $x \mid a_1$). However, this information may not have been learned, because in the learning task classifications were made based on full information about both feature states, e.g., P(Category $x \mid a_1 r_1$) and P(Category $x \mid a_1 r_2$). Thus, to estimate P(Category $x \mid a_1$), the subject has to remember P(Category $x \mid a_1 r_1$) and P(Category $x \mid a_1 r_2$), and then to average those two numbers according to the relative frequency of the $a_1 r_1$ and $a_1 r_2$ configurations.

4.1 Hypotheses

The environments used in our experiments entailed a conflict between the goals of improving accuracy and reward, when searching under asymmetric reward schemes: whereas Feature A improves overall classification accuracy, Feature R improves expected reward. The question was which feature people would prefer to view, according to the way they learned the environmental probabil-

ities and the actual (symmetric vs. asymmetric) search-task payoff structure.

Research on motor movement tasks suggests that if people meaningfully assimilate environmental probabilities, their performance might indeed approximate a reward-maximizing strategy, consistent with ideas from Bayesian decision theory. Thus, they should search Feature A under symmetric rewards (as in this case maximizing accuracy will also maximize rewards), but Feature R when searching under asymmetric rewards.

On the other hand, studies on experience-based information search (Nelson et al., 2010) show that maximizing accuracy (i.e., searching for information that helps to improve classification accuracy) best describes people's search behavior when no explicit external rewards are provided. Similarly, studies on perceptual categorization (which do not address information search) show that people have a general preference for accuracy, making it difficult to apply a reward-maximizing strategy in pure classification tasks. These findings lend support to the prediction that people may also have a preference for accuracy in information-search tasks, even when this may be suboptimal for maximizing rewards. On this view, people may preferentially search Feature A, even when searching under asymmetric payoffs.

Finally, it may be that people generally have difficulty identifying which informational query is most useful to achieve their goals when accuracy and reward conflict, even following experience-based learning of environmental probabilities. In this case, people might have no particular preference for Feature A vs. R, regardless of the applicable reward function.

In the following section, we briefly introduce the mathematics that should govern behavior (for an agent who wishes to maximize rewards, or utility) in two-way classification tasks with situation-specific rewards. Importantly, these equations also form the foundation of calculation of questions' usefulness in the context of information search, given asymmetric payoff structures. Subsequently, we describe our simulation experiments to identify environments to differentiate probability gain from situation-specific utilities, and behavioral experiments to identify what tests people select when reward and accuracy conflict.

5 Decision bounds in binary classification tasks

Consider a task in which stimuli must be designated as Category x or Category y. Given a particular reward function, including the rewards associated with correct x and y classifications, and the costs associated with incorrect x and y classifications, it is possible to determine for each

Figure 2: Information-search task illustrated. First one has to decide which feature to view (*A* or *R*, here "eye" and "claw" of plankton stimuli, respectively). The numbers show how likely one is to encounter a particular feature value, as well as the posterior probabilities of the two categories, given the feature value. Below the tree, the utility gain (Equation 5) of features (*A*, *R*) and feature values (a_1, a_2, r_1, r_2) is shown, for symmetric and asymmetric rewards. The height of the bars indicates the amount of utility gain: the width represents the frequency of occurrence. For example, in Environment 1, under the symmetric reward function, feature r_1 entails a high utility gain (0.440), but the probability of encountering this feature value is low (0.123). The tables provide detailed information on the two environments.
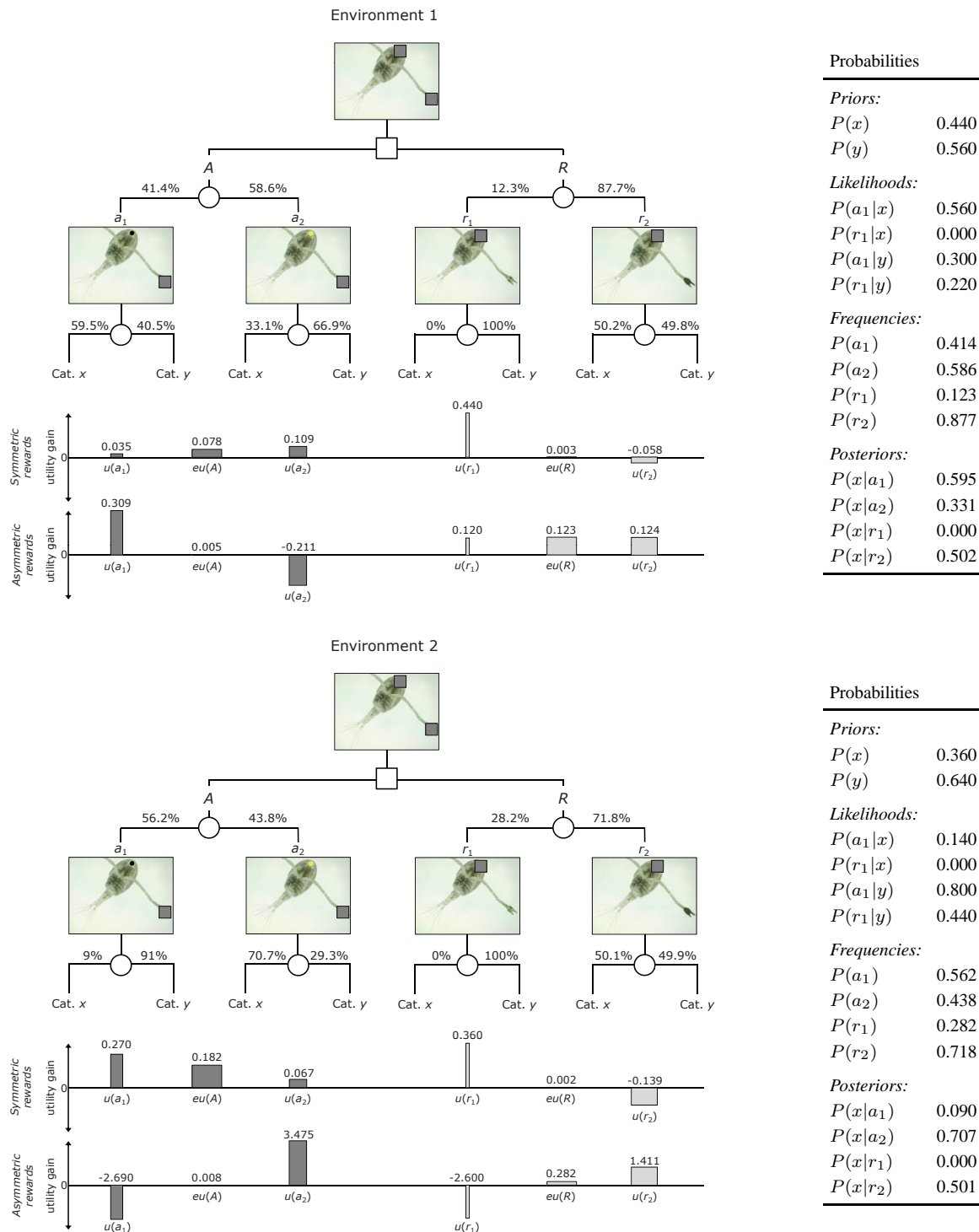
possible value of P(Category x | stimulus) whether designating an item as Category x or y has a higher expected reward (utility, or payoff). We refer to an environmental payoff structure in the form [k l m n] as follows (Figure 3a):

- k is the payoff for correct classification of a Category x stimulus ("predict x" when the true category is x);

- l is the payoff for correct classification of a Category y stimulus ("predict y" when the true category is y);

- m is the cost of erroneously categorizing a Category y stimulus as Category x ("predict x" when the true category is y);

- n is the cost of erroneously categorizing a Category x stimulus as Category y ("predict y" when the true category is x).

Given a particular reward structure[2], the reward-maximizing decision bound, which we call c_x, corresponds to the point of indifference, i.e., the probability of a stimulus belonging to Category x, $P(x)$, for which both possible categorization decisions ("predict Category x" and "predict Category y") have equal expected value (von Winterfeld & Edwards, 1982). The probability of Category x, $P(x)$, for which both categorization choices have the same expected reward is given by solving

$$kP(x) - m(1 - P(x)) = l(1 - P(x)) - n(P(x)) \quad (1)$$

for $P(x)$. The left side of the equation gives the expected value of classifying the stimulus as belonging to Category x, and the right side of the equation gives the expected value of classifying the stimulus as belonging to Category y. This implies that the indifference point, or decision criterion c_x, is given by

$$P(x) = c_x = \frac{l + m}{k + l + m + n} \quad (2)$$

---

[2]Our terms can be related to signal detection theory. If Category x is signal, and Category y is noise, then k is the payoff for a hit, l is the payoff for a correct rejection, m is the cost of a false positive, and n is the cost of a false negative (Figure 3). Our terms can also be related to the medical diagnosis scenarios used by Baron and Hershey (1988). If Category x is disease, and Category y is healthy, then k is the payoff for correctly treating a person with the disease, l is the payoff for correctly not treating a person without the disease, m is the cost (harm) from treating a person who does not have the disease, and n is the cost (neglect) from failing to treat a person who has the disease. The payoff structures we use (with positive payoffs for correct categorizations but no penalties for mistakes), and the payoff structures Baron and Hershey (1988) used (with penalties for mistakes but no rewards for correct diagnoses) can be easily equated. A [2 1 0 0] payoff structure, and a [0 0 1 2] payoff structure each lead to $c_x=1/3$; a [10 1 0 0] and a [0 0 1 10] payoff structure each lead to $c_x=1/11$, etc.

If $k - m = l - n$, then the decision bound $c_x=1/2$; we term this a *symmetric reward function* (or payoff structure). If $c_x=1/2$, selecting the more probable category maximizes both reward and accuracy (total number of correct classifications). Examples include [k l m n]=[1 1 0 0] (Figure 3b), a symmetric payoff structure under which correct classifications of either category receive one unit payoff, and there are no costs for making erroneous classification decisions, but also the payoff structure [5 10 3 8], which may not look symmetric on first glance.

An *asymmetric reward function* applies when $c_x \neq 1/2$. For example, [k l m n] = [10 1 0 0] refers to a reward scheme according to which correct Category x decisions are rewarded ten times as highly as correct Category y decisions, and there are no costs for erroneous decisions (Figure 3d). Plugging these values into Equation 2 yields $c_x=1/11$, meaning that one should predict Category x if $P(x) > 9\%$, and predict y otherwise.

Figure 3 (right hand side) exemplifies the situations graphically. For each value of P(Category x | stimulus) one can determine the action ("predict x" vs. "predict y") that has the higher expected value. The decision bound c_x is given by the intersection of the two reward functions for the "predict x" and "predict y" responses.

6 Simulation experiment: Utility gain vs. probability gain

Are there implications of symmetric vs. asymmetric reward structures for information search, when one or more properties in the environment can be queried, before making a classification decision? How can one quantify the usefulness of alternative queries, such as when deciding whether to look at Feature A ("eye") vs. Feature R ("claw") in Figure 2?

In this section we introduce probability gain, which is a psychologically plausible optimal experimental design (epistemic utility) method for selecting queries (Baron, 1985; Nelson et al., 2010). We also introduce *utility gain*, which uses the situation-specific utility structure to identify the most useful query. Finally, we use computer search techniques to identify environments in which the asymmetric reward structure strongly suggests that one query (e.g., looking at Feature R) is most useful, but probability gain strongly prefers a different query (e.g., looking at Feature A). We will subsequently use those environments in three experiments to explore whether human search can adapt to environment-specific reward structures.

Let F denote a feature (a random variable) before its specific state is known. The possible states of the feature F are $f_1, f_2, \ldots, f_m$. The expected usefulness (utility)

Figure 3: Payoff functions in a binary classification task (Category *x* vs. Category *y*)

a) *General case:* The reward-maximizing decision bound c_x is given by $(l + m) / (k + l + m + n)$. If $P(\text{Category } x \mid \text{stimulus}) > c_x$ one should predict Category *x*, if $P(\text{Category } x \mid \text{stimulus}) < c_x$ one should predict Category *y*.

	True category	
Decision	Cat. *x*	Cat. *y*
predict *x*	*k*	*m*
predict *y*	*n*	*l*
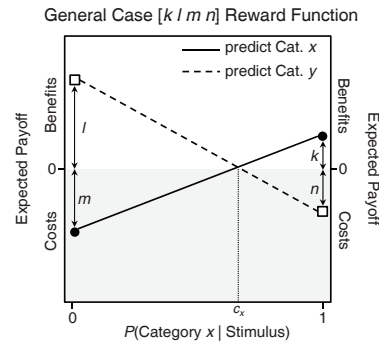
b) *Symmetric payoffs:* If $k + m = l + n$, then the decision bound $c_x = 1/2$, meaning that one should always select the more likely category, to maximize rewards. This is the implicit reward function when the goal is to maximize accuracy.

	True category	
Decision	Cat. *x*	Cat. *y*
predict *x*	1	0
predict *y*	0	1
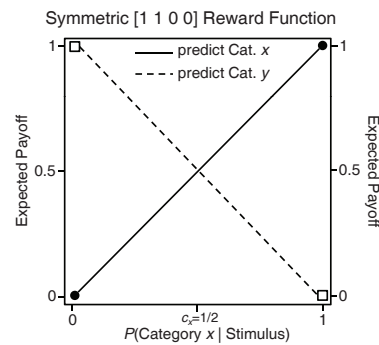
c) *Moderately asymmetric payoffs:* Correct Category *x* classifications are rewarded twice as much as correct Category *y* classifications; $c_x = 1/3$.

	True category	
Decision	Cat. *x*	Cat. *y*
predict *x*	2	0
predict *y*	0	1
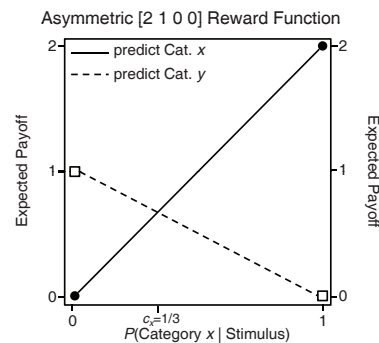
d) *Strongly asymmetric payoffs:* Correct Category *x* classifications are rewarded ten times as much as correct Category *y* classifications; $c_x = 1/11$.

	True category	
Decision	Cat. *x*	Cat. *y*
predict *x*	10	0
predict *y*	0	1

of Feature F can be defined as the average of the usefulness of the possible states of F, $u(f_j)$, weighted by their probability (Savage, 1954):

$$eu(F) = \sum_{j=1}^{m} P(f_j)\, u(f_j) \qquad (3)$$

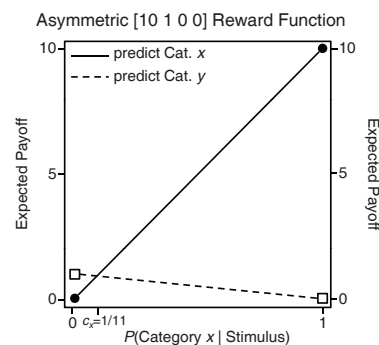Different models have been suggested to calculate the usefulness of choosing to view a feature, before the particular feature state is known. We focus on probability gain (PG; Baron, 1985), which appears to best capture people's intuitions about the usefulness of different features in classification tasks when environmental probabilities are learned through experience (Nelson et al., 2010). Probability gain quantifies the expected utility of a feature query (a test question) as the probability of correctly classifying a stimulus after the state f_j of Feature F is known, minus the probability of making a correct decision without asking the question:

$$eu_{pg}(F) = \sum_{j=1}^{m} P(f_j) \max_i (P(c_i|f_j)) - \max_i (P(c_i)) \quad (4)$$

where $P(c_i|f_j)$ denotes the posterior probability of category c_i given that feature f_j has been observed, and $P(c_i)$ is the prior probability of category c_i. As the model assumes selection of the most probable hypothesis, it is only concerned with the maximum of the prior and posterior distributions. Using this measure to quantify the usefulness of a datum and as basis for information search will maximize accuracy (i.e., number of correct classifications); the "currency" of this model is the expected improvement in correct classifications. Thus, a fixed decision bound of $c_x = 1/2$ is built in to this model.[3]

But what if different types of correct classification decisions are not equally rewarded? In this case we may

generalize the probability gain model to the *utility gain model*, which defines a datum's usefulness as the extent to which it increases the expected utility of a classification decision (Savage, 1954). The expected utility of a feature query is the utility associated with classifying a stimulus after the state of Feature F is known, minus the utility of making a decision without searching for information:

$$eu_{ug}(F) = \sum_{j=1}^{m} P(f_j) \max_i (u(\text{predict } c_i|f_j)) -$$

$$\max_i (u(\text{predict } c_i)) \qquad (5)$$

where in the two-category case "predict c_1" corresponds to "predict x" and "predict c_2" corresponds to "predict y", respectively (see Equation 1).

Conceptually, the utility gain model is similar to the probability gain model, except that utility gain is based on the maxima of the prior and posterior *utility* distribution, taking into account the costs and benefits of different types of classification decisions. The "currency" of this model is improvement in expected utility, and the implicitly entailed decision bound maximizes utility (reward). Under a symmetric reward function, the utility gain model reduces to the probability gain model (i.e., both models operate with the same decision criterion $c_x = 1/2$).[4]

---

[3]To illustrate, consider Environment 1 in Figure 2. Without querying Feature A or R, one should assign an item to Category y, which has the higher base rate (56%) (see table in Figure 2). On average one would achieve 56% classification accuracy. Imagine we can view just one feature, and we decide to look at Feature A. We will observe a_1 with 41.44% probability and a_2 with 58.56% probability. If we observe a_1, we should classify the items as Category y, achieving 59.45% accuracy. If we observe a_2, we should assign the stimulus to Category y, and will be correct on 66.94% of the cases. Weighing these outcomes by the frequencies of each respective feature value yields an expected accuracy of $0.4144 \times 0.5945 + 0.5856 \times 0.6694 = 0.6384$. Thus, looking at A will improve our classification accuracy from 56% to 63.84%, on average; the probability gain of this feature is 0.0784. Now consider Feature R. If we search this feature, we will observe r_1 with 12.32% probability and r_2 with 87.68% probability. If r_1 is observed, we know that the stimulus definitely belongs to Category y. If we encounter r_2, x is slightly more likely (50.18%). The expected accuracy is $0.1232 \times 1 + 0.8768 \times 0.5018 = 0.5632$. Thus the probability gain of Feature R is effectively zero, because querying Feature R does not meaningfully improve accuracy, versus the 56% accuracy that can be achieved without looking at either feature.

[4]To illustrate, consider again Environment 1 (Figure 2), but now assume that correct Category x classifications pay two units reward, whereas correct Category y decision pay one unit reward (i.e., a [2 1 0 0] reward function applies). Without querying a feature, the expected reward when classifying an item as Category x (which is the less likely, but higher rewarded category) is higher than Category y ($0.44 \times 2 = 0.88$ vs. $0.56 \times 1 = 0.56$). If looking at Feature A, one will observe a_1 with 41.44% probability and a_2 with 58.56% probability. If we observe a_1, then the probability of Category x is 59.45%, so we should classify the items as Category x, which yields $0.5945 \times 2 = 1.1891$ units of reward. If we observe a_2, the probability of Category y is 66.94% (which is greater than the threshold of 2/3), so we should assign the stimulus to Category y, yielding $0.6694 \times 1 = 0.6694$ units of reward. Thus, the expected value of Feature A is $0.4144 \times 1.1891 + 0.5856 \times 0.6694 = 0.8848$, effectively unchanged from the 0.88 units of utility that can be obtained without looking at either feature. Thus, utility gain of Feature A (0.0048) is effectively zero. Now consider Feature R. We will observe r_1 with 12.32% probability and r_2 with 87.68% probability. If r_1 is observed, we know that the stimulus definitely belongs to Category y, yielding one unit reward. If we encounter r_2, both categories are equally likely (50.18% vs. 49.82%), but since x is higher rewarded, the item should be assigned to this category, yielding $0.5018 \times 2 = 1.0037$ units reward. The overall expected reward of Feature R is $0.1232 \times 1 + 0.8768 \times 1.0037 = 1.0032$. Thus, the utility gain expected from querying R is $1.0032 - 0.88 = 0.1232$ units reward.

6.1 Environments for differentiating probability gain and utility gain

Would using the utility gain model lead to making different queries than using probability gain, when searching for information under asymmetric payoffs? We used computer simulations to search for environments in which the goals of accuracy and reward make maximally contradictory claims about which of two features would be most useful to view. We considered a moderately asymmetric [2 1 0 0] payoff structure, as well as a strongly asymmetric [10 1 0 0] payoff structure, in separate optimization procedures. Given these payoff functions we searched in environments with two mutually exclusive categories, Category x and Category y, and two features: A, which can take values a_1 and a_2, and R, which can take values r_1 and r_2. For instance, for the [2 1 0 0] reward structure, we searched for values of $P(x)$, $P(a_1|x)$, $P(a_1|y)$, $P(r_1|x)$, and $P(r_1|y)$, such that the probability gain of Feature A is much higher than that of Feature R, but the utility gain (relative to the [2 1 0 0] payoff structure) of Feature R is much higher than that of Feature A.[5] A further constraint was that both features would be needed in order to achieve the maximum possible accuracy in classifying stimuli as Category x vs. Category y. This constraint was included to insure that both features would be learned, in the learning task of our experiments. The features were class-conditionally independent (i.e., the state of Feature A was independent of the state of Feature R, and vice versa, conditional on the true category).

Figure 1 outlines the environments found by these optimizations. The two trees illustrate the frequency of the different stimulus configurations in the two environments, and the probability of Category x and y for each configuration, in each environment.[6] The different decision rules built into the probability gain and utility gain model directly influence the usefulness of features A and R (Figure 2). Consider Environment 1 (Figure 2, top). If all correct classifications are rewarded equally and there is no penalty for incorrect classifications, then Equations 4 and 5 may be used to verify that probability gain and utility gain both consider Feature A to be more useful

than Feature R (remember that probability gain and utility gain are identical under symmetric rewards). By contrast, if the moderately asymmetric [2 1 0 0] reward structure applies, according to which correct Category x classifications are rewarded twice as highly as correct Category y classifications, Feature R is more useful in terms of maximizing reward than Feature A is (Figure 2). Given this environment and reward structure, maximizing accuracy and reward at the same time is not possible when selecting a single feature to view. In Environment 1, in the experiments, when we refer to an asymmetric reward structure, it has a [2 1 0 0] payoff function.

We also identified a second environment (*Environment 2*) in which a strongly asymmetric [10 1 0 0] payoff structure makes Feature R more useful, but in which Feature A leads to higher classification accuracy (Figure 2). In Environment 2, in the experiments, when we refer to an asymmetric reward structure, we mean a [10 1 0 0] payoff function.

The results of these optimizations show that, if one's goal is to behave adaptively in an environment with asymmetric payoffs, it is not adequate to use informational utilities to select pieces of evidence to acquire. The choice of queries (experiments, or features to view), and not only the eventual classification decision, must reflect the environment-specific payoff structure.

7 Overview of Experiments

We studied information search under symmetric and asymmetric reward functions, in the two environments identified through our computer simulations. To give a richer perspective, we manipulated across experiments the way in which subjects learn about environmental probabilities and the relevant payoff structure.

In Experiment 1, a neutral experience-based multiple cue category-learning task was used to convey environmental probabilities to subjects. The categorization task was followed by a search task, in which subjects gathered information under symmetric or asymmetric real-money payoffs. When searching under symmetric rewards, correct classifications of either category were paid the same amount of money. Under asymmetric payoffs correct Category x classifications were paid more money than correct Category y classifications.

The search task in Experiment 2 was virtually identical to Experiment 1, involving real-money payoffs, with the reward structure (symmetric vs. asymmetric) manipulated between conditions. The difference was that in the initial experience-based learning task subjects learned to classify stimuli under an asymmetric reward structure. Subjects received points for correct classification deci-

---

[5]The optimization procedure itself was similar to that of Nelson et al. (2010). The main distinction is that Nelson and colleagues contrasted various OED models with each other, whereas we contrasted each asymmetric reward function with probability gain.

[6]For example, in Environment 1 configuration a_2r_2 occurs with a probability of 49.9% and belongs with a probability of 38.8% to Category x. Imagine your goal is to maximize total number of correct classifications. In this case, the stimulus should be classified as Category y, which is the more likely category ($P(y|a_2r_2) = 61.2\%$).

But what if correct Category x classifications are rewarded twice as highly as correct Category y predictions? In this case, the item should be assigned to x, as $0.388 \times 2 > 0.612 \times 1$.

Figure 4: Classification-learning task illustrated. A stimulus ("plankton") is shown and must be categorized as *x* or *y* ("Species A" or "Species B"). In Experiment 1, if the item is correctly categorized, feedback in form of a smiley appears; if incorrectly classified, a frowny face appears. The learning task in Experiment 2 was virtually the same, except that instead of a smiley or frowny face, points were associated with correct and incorrect classifications, with the amount of points earned depending on the reward function (small inset picture, at top right). Erroneous classifications were associated with zero points.
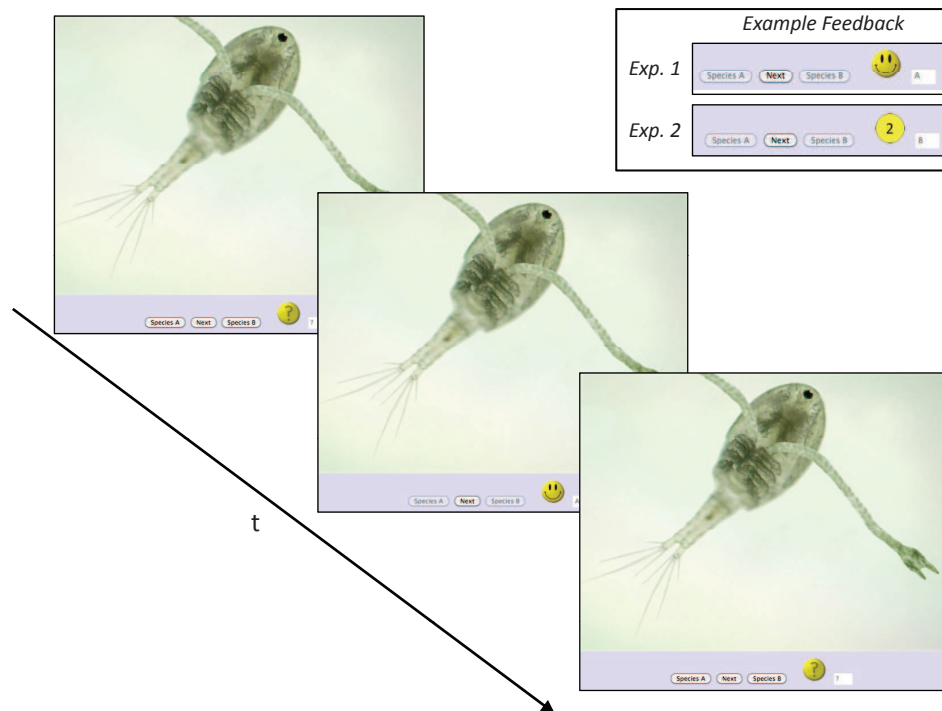
sions, according to an explicit asymmetric payoff function (i.e., correct Category *x* classifications received more points than correct Category *y* classifications). Their task was to learn to make categorization decisions that would maximize expected reward (points).

In Experiments 1 and 2 alike, after the experience-based classification task, the search-task payoff structure was announced with words and numbers; the real-money payoffs were provided immediately after the experiments. The objective was to study people's ability to respond to asymmetric payoff structures after learning about the statistical structure of the environment through experience. No feedback was provided during the search task itself, to prevent the possibility that search strategies could be adjusted according to search-task feedback.

Finally, in Experiment 3 we compared search behavior under symmetric vs. asymmetric payoffs in a completely description-based task, in which both environmental probabilities and payoff functions were presented with words and numbers.

8 Experiment 1

Experiment 1 examined whether information search can adapt to asymmetric reward structures, if people learn about environmental structure and probabilities through experience. An initial experience-based multiple cue probabilistic category-learning task was used to help subjects internalize the probabilistic structure of a particular environment (Environment 1 or 2, as described above and in Figure 1).

Our focus was on the subsequent information-search task, which utilized the same stimuli. In the search task, however, the features were obscured; people could view only a single feature (*R* or *A*), of their choice, before categorizing the stimulus, in each trial (Figure 2). Before the search task, subjects were informed that real-money payoffs would be provided after the experiment, according to the announced (asymmetric or symmetric) payoff structure.

What goals do people have in search? If people want to maximize external rewards, they should prefer to view

Feature *R* under asymmetric rewards. However, under symmetric rewards they should preferentially view Feature *A*, which maximizes accuracy *and* reward. By contrast, if people have a general preference to search for information that helps to improve overall accuracy of classification decisions, they should search Feature *A* regardless of the applicable external payoff function.

8.1 Subjects

Subjects were 91 volunteers—largely university students—from Berlin (48% female, mean age 26). They received a show-up fee of 10€ and could earn an additional bonus of up to 20€. Subjects were randomly assigned to one of four conditions: one of two environments {Environment 1, Environment 2} × one of two search-task reward structures {symmetric, asymmetric}.

8.2 Materials and Procedure

The experiment consisted of two tasks, an initial classification-learning task and a subsequent information-search task. The learning task involved categorizing simulated plankton specimens as Category *x* or Category *y*, based on the full feature configuration (Figure 1). (The categories were described as "Species A" or "Species B".) In each trial, a stimulus was chosen randomly according to the environmental probabilities and the subject categorized the item as belonging to Category *x* or *y* (Figure 4). Both features were visible in each stimulus presented, in each trial, throughout the categorization learning task. Subjects were familiarized with the two forms of each feature (e.g. a_1 vs. a_2, and r_1 vs. r_2) beforehand. After a categorization decision was made, in a trial, feedback on the true species was given, accompanied by a smiley or frowny face, depending on whether the categorization decision was correct or incorrect (Figure 4, upper-right inset). The assignment of physical features to probabilities, the polarity of each feature (e.g., which version of Feature *A* is considered a_1 and which version is considered a_2), and whether Category *x* or *y* was labeled as "Species A", were randomized. For each subject, two of three possible features ("tail", "eye", and "claw" of the plankton specimens) were chosen at random to be used. This led to 96 randomizations in each condition, one of which was chosen at random for each subject.

Subjects were instructed to learn to correctly classify the stimuli, based on the states of features *A* and *R*. No explicit rewards were provided in the learning task. Learning continued until criterion performance was reached, or the available time (around 2 hours) elapsed. For any stimulus, the optimal strategy was defined as predicting Category *x* when P(Category *x* | stimulus) >1/2 and predicting

Category *y* when P(Category *x* | stimulus) <1/2. Criterion performance was defined as (1) making at least 98% optimal (not necessarily correct) responses in the last 200 trials, irrespective of the specific stimuli in those trials; and (2) making an optimal response in the last five trials of every single stimulus type. The latter criterion ensures that even rare configurations are learned (see Figure 1 for frequencies of stimulus configurations). The purpose of the strict learning criterion was to ensure that subjects meaningfully assimilated the environmental probabilities, before the information-acquisition task. Subjects were also periodically informed at the accuracy they would achieve if they continued to respond as they did in the last 200 trials, and how well the optimal strategy would perform. This feedback was given on trial number 500, 750, 1000, 1250, etc.

In the subsequent information-acquisition task, which comprised ten trials without feedback, subjects continued to classify the plankton stimuli.[7] However, in this phase the stimuli's features were obscured, and subjects could reveal only one of the features on each trial (Figure 2). The crucial differences, versus the initial learning task, were that (i) prior to searching a feature it was uncertain which state it would take, so subjects had to take into account the individual feature states' usefulness and the probability of their occurrence, and (ii) after revealing a feature, classifications had to be made based on the state of a single feature alone, whereas both feature values were known in the learning task. The crucial question was which feature subjects would prefer to view. The same plankton stimuli, with the same environmental probabilities as in the learning task, were used; this was disclosed to subjects.

The payoff scheme was manipulated between subjects. In the *asymmetric reward conditions*, which were the main manipulation of interest, subjects' received 2€ for correct Category *x* classifications. The payment for correct Category *y* classifications was 1€ in Environment 1 ([2 1 0 0] asymmetric reward structure), and 0.2€ in Environment 2 ([10 1 0 0] asymmetric reward structure). In the *symmetric reward conditions* subjects received 2€ for each correct classification of either species. In the asymmetric payoff conditions, Feature *A* leads to highest classification accuracy, but Feature *R* leads to highest expected reward (Figure 2). In the symmetric rewards conditions, Feature *A* leads to highest classification accuracy as well as to highest reward.

---

[7]Why just 10 information-search trials? Nelson et al. (2010) used 101 information-search trials. However, within subjects, information-search behavior was very consistent from one trial to the next. In Nelson et al.'s Experiment 1, for instance, the median subject viewed the higher probability gain feature in 100 out of 101 search-task trials. Thus, having 100 or more search-task trials would not likely provide additional information about subjects' behavior.

The payoff scheme was presented to subjects before the information-search task. To prevent learning from feedback during the search task, no feedback (no smiley or frowny face, no count of accumulated earnings or number of correct decisions, or otherwise) was provided during the ten search trials themselves. After the end of the experiment, subjects were given feedback on their correct and incorrect Category *x* and *y* classifications, and paid in cash according to the previously announced asymmetric or symmetric payoffs.

To assess how well subjects were calibrated to environmental probabilities, subjects were subsequently presented with all stimuli (plankton specimens), and asked to estimate for each type the proportion that are each species, i.e., $P(\text{Category } x \mid \text{stimulus})$, for all four combinations of the features $a_1 r_1$, $a_1 r_2$, $a_2 r_1$, and $a_2 r_2$. Subjects also estimated the base rates of the two categories, $P(x)$ and $P(y)$.

8.3 Results

Most subjects (80/91=88%) achieved the learning criterion well within the available time ($\sim$2h). Two subjects (one in each environment) were replaced as the experienced probabilities did not match the true environmental probabilities closely enough to produce the intended ordering in features' relative usefulness.

Our main research questions concerned the information-search task (Figure 2), in which subjects could only select one feature to view, on each of the ten trials (Table 1). Most subjects with the symmetric search-task payoff structure viewed Feature *A* on more than half of the search-task trials (27/39 = 69% of subjects; two-tail binomial p=.02). Surprisingly, among subjects with the asymmetric payoff structure, a majority also preferentially viewed Feature *A*, which has higher probability gain, rather than Feature *R*, which has the higher expected reward under asymmetric payoffs (28/38=74% of subjects, two-tail binomial p<.01).

Analysis of the mean percentage of views to Feature *A*, aggregating across subjects, gives a similar picture. There was no effect of reward manipulation ($M_{symmetric}$=68%, $M_{asymmetric}$=68%). The results were similar in Environment 1 and Environment 2 (M_{env1}=63%, M_{env2}=73%). The possible slight difference between the environments could reflect some subjects' sensitivity to information gain, which prefers Feature *A* in Environment 2, but Feature *R* in Environment 1 (see General Discussion and Figure 5).

These results extend Nelson et al.'s (2010) finding that subjects prefer to view higher-probability gain features on classification tasks in which no particular costs or benefits apply. Importantly, *the preference to search higher-probability gain features may apply even when it is maladaptive* given the environmental probabilities and situation-specific payoffs.

Subjects' probability estimates for the different feature configurations were also analyzed. There was good correspondence between subjects' estimates and the true environmental probabilities, indicating that subjects acquired reasonable knowledge of environmental structure (Table 2).

8.4 Discussion

How can it be that subjects' probability estimates appeared well calibrated, yet subjects did not search to maximize payoffs under asymmetric reward structures? Did subjects fail to view the reward-maximizing feature because of a strong desire to maximize classification accuracy? Or were subjects unable to flexibly use their knowledge of environmental probabilities to identify the reward-maximizing feature, given an asymmetric reward function?

One explanation for the preference for Feature *A* is that maximizing accuracy serves as a kind of overriding goal in search behavior, which dominates the external reward function (i.e., monetary payoffs). For example, over the course of life, including school history, people may have learned that being accurate is an important goal in many situations, and have therefore developed a general preference for searching for information that allows them to improve accuracy.

An alternative explanation is that subjects in the asymmetric payoff conditions did want to earn the most money possible, but for some reason perceived Feature *A* to be most useful, relative to that goal. This idea would match the (incorrect) intuition that it is reasonable to conduct queries with the goal of learning the true state of nature as accurately as possible, and to take asymmetries in the reward structure into account only in the actual classification decision.

A third explanation is that, the theoretical complexities in calculating individual features' usefulness notwithstanding, features' usefulness was somehow encoded relative to an implicit accuracy-based reward structure in the initial classification-learning task. For instance, people might build up a decision tree that orders features relative to the accuracy-based learning task structure. People might not be able to use their knowledge of environmental probabilities in a flexible way to identify the most useful queries, in response to the novel asymmetric search-task payoff structures. Such a finding could be problematic for a Bayesian-decision theoretic account of cognition and behavior (see General Discussion).

Table 1: Search-task views to Feature *A*, which maximizes accuracy, across Experiments 1, 2, and 3

Rewards; Environment	Experiment 1		Experiment 2		Experiment 3
	% of SS	% of views (CI)	% of SS	% of views (CI)	% of SS
Symmetric; Env. 1&2	69% (27/39)[*]	68% (55%-78%)	9% (2/23)[****]	12% (6%-25%)	50% (11/22)
Asymmetric; Env. 1&2	74% (28/38)[**]	68% (55%-79%)	0% (0/20)[****]	2% (0%-6%)	63% (12/19)
Symmetric; Env. 1	60% (12/20)	59% (41%-75%)	8% (1/13)[**]	11% (3%-28%)	54% (7/13)
Symmetric; Env. 2	79% (15/19)[*]	77% (60%-89%)	10% (1/10)[*]	14% (4%-40%)	44% (4/9)
Asymmetric; Env. 1	70% (14/20)	67% (47%-82%)	0% (0/11)[***]	1% (0%-5%)	45% (5/11)
Asymmetric; Env. 2	78% (14/18)[*]	70% (52%-83%)	0% (0/9)[**]	2% (0%-9%)	88% (7/8)

Note. Under symmetric payoffs, Feature *A* also leads to highest reward; under asymmetric payoffs, Feature *R* leads to highest reward. In Experiment 1, following a neutral category-learning task, subjects preferentially viewed the accuracy-maximizing feature (*A*). In Experiment 2, following an asymmetric learning-task reward function, subjects preferentially viewed the other feature (*R*). Experiment 3 used a verbally described information-search scenario; no consistent preference for either feature was observed. 95% confidence intervals (CI) for mean proportion Feature *A* views were calculated using bootstrap sampling (bias-corrected and accelerated, in Matlab). In Experiment 1, three subjects viewed Feature *A* and Feature *R* equally often and were excluded (for computing percentage of views to feature *A*, these subjects were included). In Experiment 3 one subject ranked Feature *A* and *R* as equally useful, and one subject ranked the uninformative foil feature as more useful; both subjects were excluded. Two-tail, uncorrected binomial *p*-values are reported as follows: $p < .05$ = *, $p < .01$ = **, $p < .001$ = ***, $p < .0001$ = ****. In Experiment 3, and where not specifically noted, results were not significantly different from 50% of subjects preferring to view Feature *A*.

9 Experiment 2

Experiment 2 was designed to investigate the possible influence of the learning task on search-task behavior, by explicitly cueing asymmetric payoff structures during learning, but was otherwise virtually identical to Experiment 1. As before, a classification-learning task was followed by an information-search task. The exact same plankton stimuli, environmental probabilities, and asymmetric or symmetric search-task payoffs were used. The main difference was a slight modification to the learning-task feedback, to explicitly cue the asymmetric environmental payoff structure in the learning task.

After each learning-task trial of Experiment 2, the subject was shown a number of points (rather than a smiley or frowny face, as in Experiment 1) according to their classification decision, and whether it was correct. Correct Category *x* classifications were worth more points than correct Category *y* classifications, according to the payoff structure of the environment: twice as much in Environment 1, and ten times as much in Environment 2. Subjects were asked to make classification decisions that would maximize expected reward (points), even when that requires assigning an item to the less likely category.

The information-search task of Experiment 2 was virtually identical to Experiment 1, with the reward structure (symmetric vs. asymmetric) manipulated between subjects, and the same real-money payoffs. Our goal was to examine how learning to classify under asymmetric reward functions would influence subsequent information search under symmetric vs. asymmetric payoffs.[8] One possibility is that people generally prefer to search for information that improves accuracy. This would be especially consistent with Nelson et al.'s (2010) interpretation of their findings. It could also be seen as being consistent with Maddox and Bohil's (1998) competition between reward and accuracy (COBRA) hypothesis, although this model was designed to apply to classification decisions, not to information-search behavior. In this case, similar

---

[8] In each environment (Figure 1), the asymmetric learning-task reward structure ([2 1 0 0] in Environment 1, [10 1 0 0] in Environment 2) induces a conflict between accuracy and reward maximization, for a *conflict configuration* in which one decision has a higher expected reward, even though another decision is more likely to be accurate. In Environment 1, the probability that item $a_2 r_2$ belongs to Category *y* is 61.2% (Figure 1). However, under the asymmetric [2 1 0 0] payoff scheme (Figure 1c), maximizing points earned requires consistently categorizing this item as belonging to the less likely Category *x*. This will yield $2 \times 0.388 = 0.776$ points on average, as opposed to $1 \times 0.612 = 0.612$ points when categorizing it as belonging to the more probable Category *y*. Similarly, in Environment 2 under the [10 1 0 0] reward function, item $a_1 r_2$ has a probability of only 15% of belonging to Category *x*. However, categorizing this item as belonging to the less likely Category *x* is the reward-maximizing strategy, which will yield $10 \times 0.15 = 1.5$ points on average, as opposed to $0.85 \times 1 = 0.85$ points for categorizing it as Category *y*.

findings as in Experiment 1 should be obtained, with a preference to search Feature *A*, regardless of the search-task reward structure.

However, if the asymmetric learning-task payoffs help people identify the reward-maximizing strategy, a preference for Feature *A* should be observed when searching under symmetric payoffs, and a preference for Feature *R* when searching under asymmetric payoffs. Finally, if information search is driven by the usefulness of features relative to the learning-phase payoff structure, a general preference for Feature *R* could be observed, even for subjects with symmetric search-task payoffs, for whom Feature *A* leads to highest accuracy and payoffs.

9.1 Subjects

Subjects were 80 young adult (largely university students; 58% female, mean age 25 years) volunteers from Berlin. Subjects were randomly assigned to one of four conditions: one of two environments {Environment 1, Environment 2} × one of two search-task reward structures {symmetric, asymmetric}. They received a show-up fee of 5€ and could earn an additional bonus of up to 20€ in the information-acquisition task.

9.2 Materials and Procedure

Experiment 2 employed the same stimuli and probabilistic environments as Experiment 1. However, explicit asymmetric reward structures were presented during the classification learning task. Prior to learning subjects were informed about the reward structure. In Environment 1 (asymmetric [2 1 0 0] reward structure; Figure 3c), subjects were told that correct classifications of Category *x* were associated with 2 points, and correct classifications of Category *y* were associated with 1 point. In Environment 2 (asymmetric [10 1 0 0] reward structure; Figure 3d), correct Category *x* classifications were associated with 2 points and correct Category *y* classifications were associated with 0.2 points. This point manipulation was intended to be a heavy-handed way to introduce each asymmetric reward structure in the learning task. Subjects were explicitly instructed to classify stimuli in a way that maximizes points earned on average. No points were accumulated, however, as the goal was for subjects to learn the probabilities and the optimal response strategy, rather than to incentivize subjects to quickly amass a large number of learning trials. The assignment of environmental structure to specific labels and physical features in the stimuli was randomized across subjects, as in Experiment 1.

As in Experiment 1, in each trial a plankton specimen was chosen randomly according to the environmental probabilities (Figure 1). After making a classification de-

cision, a "?" symbol at the bottom of the screen changed to a 2, 1 or 0.2 (according to the reward function and the category) if the decision was correct, or to a 0 if the decision was incorrect (Figure 4, upper-right inset). Learning continued until the subject's performance approximated a reward-maximizing classifier that on each trial chooses the category with the higher expected reward. The precise learning criterion was analogous to Experiment 1, namely both (1) making at least 98% reward-maximizing responses in the last 200 trials, irrespective of the stimuli in those trials; and (2) making reward-maximizing responses in the last five trials of every single stimulus type. Feedback on the points they would earn on average, if they continued as in the last 200 trials, and the number of points that the optimal classifier would earn on average, was given after trial 200, 400, 600, 800, etc. (Feedback was slightly increased, vs. Experiment 1, because pilot work suggested that the learning task of Experiment 2 would be more difficult.) The instructions, and this periodic feedback, both emphasized that in some cases, it may be necessary to choose the less-probable category, to obtain the highest expected points.

The information-search task of Experiment 2 was virtually identical to that of Experiment 1. Half of the subjects were assigned to symmetric search-task payoffs, in which correct Category *x* and correct Category *y* classifications each paid 2€. Given this symmetric payoff structure, Feature *A*, which leads to higher classification accuracy, also leads to higher expected payoffs (Figure 2). In this case, note that neither the goal of maximizing accuracy nor the goal of maximizing reward would suggest viewing Feature *R*. Half of the subjects were assigned to an asymmetric search-task reward function. When searching under asymmetric rewards, Feature *A* maximizes classification accuracy (but does not improve expected reward), whereas Feature *R* maximizes expected reward (but does not improve overall classification accuracy). In Environment 1, correct Category *x* classifications paid 2€ and correct Category *y* classifications paid 1€. In Environment 2, correct Category *x* classifications paid 2€ and correct Category *y* classifications paid 0.2€.

In each environment, the asymmetric search-task reward structure exactly matched the point structure that had been experienced in the learning task. The payoff structure that would apply was presented to subjects before the information-search task. Subjects were explicitly instructed to choose a feature to view, so as to earn the most money. As in Experiment 1, there were ten information-search trials, with no feedback during those trials. Real-money payoffs were given after the experiment according to the decisions made and the applicable payoff function. After the search task, subjects were given a separate questionnaire with which they rated environmental probabilities, as in Experiment 1.

Table 2: Subjects' median probability estimates.

	Environment 1			Environment 2		
Item	True percent	Experiment 1	Experiment 2	True percent	Experiment 1	Experiment 2
$P(x)$	44%	36%	50%	36%	25%	35%
$P(x \mid a_1\, r_1)$	0%	0%	0%	0%	8%	0%
$P(x \mid a_1\, r_2)$	65%	73%	70%	15%	11%	25%
$P(x \mid a_2\, r_1)$	0%	0%	0%	0%	2%	0%
$P(x \mid a_2\, r_2)$	39%	35%	53%	81%	85%	90%

Note. The item being judged is in the left column; its true percent next; and the median of subjects' estimated percentages next. In this table, "*x*" denotes whichever category was less probable in a particular subject's randomization. Overall, subjects were well-calibrated, especially in Experiment 1. Most individual subjects appeared qualitatively well-calibrated, as well. In Experiment 2, there may be a tendency to overestimate the probability of Category *x*, perhaps due to the asymmetric learning-task reward structure in which correct Category *x* classifications have a higher payoff than correct Category *y* classifications.

Table 3: Learning difficulty across Experiments 1 and 2.

	Experiment 1 % learners	Experiment 2 % learners	Experiment 1 Median learning trials	Experiment 2 Median learning trials
Environment 1	82% (40/49)	60% (24/40)	469	733
Environment 2	95% (40/42)	48% (19/40)	280	794

Note. The number of learning trials is based on learners only. Experiment 2's asymmetric learning-task reward structure made learning much more difficult than in Experiment 1. This is reflected by the smaller proportion of learners, and the learners' higher number of learning-task trials. The especially strong difference in number of learning trials required in Environment 2, between the experiments, could be due to its especially high conflict (with the highly asymmetric [10 1 0 0] reward structure) between reward and accuracy maximization (configuration $a_1 r_2$, see Figure 1). See Appendix, Figure A1, for more detailed learning-task results.

9.3 Results and Discussion

Only about half of the subjects (43/80 = 54%) reached the learning criterion in the ~2h available for learning, in contrast to virtually all (80/91 = 88%) subjects in Experiment 1. Even restricting consideration to the learners, a greater number of trials were needed, in the same environments, than in Experiment 1 (Table 3). Analysis of learning data shows that the difficulty in learning in Experiment 2 stemmed from the need to choose the less-probable (but higher-rewarded) category, on the conflict configurations for which Category *y* was more probable but a Category *x* response had higher expected reward (Appendix, Figure A1). On these stimuli, the vast majority of subjects first responded in a way that would maximize accuracy, rather than reward, and only later shifted to the reward-maximizing response strategy. The learning task in Experiment 2 can itself be seen as an extension of the signal detection theory paradigm into a multiple-cue probabilistic category learning task. These learning-task results suggest that the competition between reward and accuracy (COBRA) hypothesis may extend to multiple-cue category learning tasks under asymmetric rewards.

Our research questions concerned information search under asymmetric-vs.-symmetric rewards, in the subsequent information-search task. In stark contrast to Experiment 1, almost all learners preferentially viewed Feature *R* (Table 1). The preference to view Feature *R* was seen when that feature improved expected reward, in the asymmetric reward conditions (20 of 20 subjects: two-tail binomial $p < .0001$). Remarkably, this preference was also seen when a symmetric reward function—under which Feature *A* maximizes both accuracy and reward—was used (21 of 23 subjects: two-tail binomial $p < .0001$). Analysis of the mean percent of views to Feature *A* confirms these findings ($M_{symmetric} = 12\%$, $M_{asymmetric} = 2\%$).

This is particularly remarkable, as under symmetric payoffs Feature *R* does not improve classification accuracy *or* reward.

This clarifies the meaning of results from Experiment 1, as well as Nelson et al. (2010). *People do not always prefer to view higher-probability gain features*. In Experiment 2, which used an asymmetric environmental reward structure during learning, the lower-probability gain feature (Feature *R*) was preferentially viewed in the search task, even when doing so was suboptimal with respect to both reward and accuracy, given a symmetric search-task payoff structure.

Together with Experiment 1, these results suggest that, rather than subjects having the capacity to flexibly adapt search behavior following experience-based learning and a goal of being accurate, their information search seems to be driven by the importance of features as identified during learning under an explicit asymmetric or (in Experiment 1) implicit symmetric reward function. In Experiment 2, the strong finding is that, irrespective of the search-task payoff structure (the manipulation of which had little-to-no effect on search behavior), subjects preferred to view Feature *R*. Across both experiments, subjects tended to view the feature that *would have been the most useful, if the learning-task reward structure were preserved in the search task, irrespective of the actual search-task reward structure*. Put more poignantly, it may be that for people to search appropriately in environments with asymmetric payoff structures, the payoff structures themselves must also be learned through experience.

These results strongly argue against the idea that people have a universal accuracy goal in search. They also argue against the idea that people can adaptively search according to novel announced payoff structures, following experience-based learning of environmental probabilities. Relative to the COBRA hypothesis, there is no evidence of a conflict between reward and accuracy in the search-task behavior per se. This conflict is seen in the learning task. In the search task, however, search behavior follows whichever way the reward-accuracy conflict was resolved in the learning task.

10   Experiment 3

Experiments 1 and 2 both used experience-based learning to convey environmental probabilities to subjects. Most research on information acquisition, however, has used words and numbers to convey environmental probabilities. Experiment 3 therefore examined information-search behavior under asymmetric payoffs when environmental probabilities are conveyed through summary statistics. Subjects were the 43 people who successfully completed Experiment 2. Upon completion of the experience-based learning task and information-search task in Experiment 2, subjects were presented with a verbally described information-search scenario in which the task was to categorize fictitious aliens into two species (adapted from the "Planet Vuma task", Skov & Sherman, 1986). In this scenario, subjects' task was to identify the species of an invisible alien ("Glom" vs. "Fizo"), by querying certain features (e.g., "wearing a hula hoop"). For each subject, the environmental probabilities (categories' base rates and feature likelihoods) and the symmetric or asymmetric reward function were identical to the experience-based plankton classification task that they had just completed in Experiment 2. (The homology between the tasks was not disclosed.) In Experiment 3, subjects ranked the features' usefulness, relative to the explicitly provided reward function, rather than actually viewing features and categorizing individual stimuli. Base rates were verbally described in terms of percentages (e.g., "Out of one million creatures on planet Vuma, 44% are Fizos and 56% are Gloms"). Feature likelihoods were presented in a table, denoting what percentage of each species possessed each feature (e.g., "56% of Fizos wear a hula hoop"). An uninformative third feature, present in 0% or 100% of both species, was also included, as a foil to ensure that subjects understood the information presented.

The same payoff structure as the subject had just experienced in Experiment 2 was used. Here, it was described in terms of points (e.g., "For each correct classification of a Glom you get 2 points. For each correct classification of a Fizo you get 1 point."). Subjects were asked to rank order the questions according to their usefulness: "Considering the information given, what questions would most help you to earn the most points possible?" A bonus of 5€ was given if the questions were correctly ranked in order of their usefulness (i.e., *A > R* under symmetric rewards and *R > A* under asymmetric rewards, with the useless feature not ranked higher than Feature *A* or *R*).

10.1   Results and Discussion

In the summary-statistics-based scenario there was no discernible preference between the features. There does not appear to be any effect of the payoff manipulation, either (Table 1). If anything, the trend is in the wrong direction: under asymmetric rewards, 63% of people ranked Feature *A* to be more useful than Feature *R* (two-tail binomial *p*=.17), whereas under symmetric payoffs 50% of the subjects ranked Feature *A* to be more useful.

These results add to Nelson et al.'s (2010) finding that there is little relationship between actual search behavior following experience-based learning and judgments of features' usefulness, based on summary statistics. The present data further show that there may be

no effect of an explicit payoff manipulation, in an abstract summary-statistics-based information search scenario. These data contribute to a body of research, which has focused on risky-choice gambling decisions, examining the circumstances under which there are differences in description- versus experience-based decisions (Hadar & Fox, 2009; Hertwig, Barron, Weber & Erev, 2004; Hertwig & Pleskac, 2010; Ungemach, Chater & Stewart, 2009).

11   General Discussion

Does experience-based classification learning provide a foundation for adaptive information search in environments with asymmetric reward structures? We addressed this in Experiment 1, using experience-based classification learning, with natural sampling and immediate feedback, to train people in the statistical structure of two probabilistic environments. Subsequently subjects searched for information under symmetric vs. asymmetric payoffs. Remarkably, even given experience-based learning and subjects' reasonable explicit knowledge of environmental probabilities, there was no perceptible effect of the payoff manipulation on the feature subjects preferred to view in the search task. Rather, most subjects preferentially searched Feature *A*, which had higher probability gain (i.e., led to higher accuracy), irrespective of whether it led to higher rewards or not, given the monetary payoff structure in the information-search task.

Is there any way for people to learn to choose information adaptively in environments with asymmetric reward structures? Experiment 2 addressed this by explicitly giving subjects asymmetric reward structures in the initial categorization-learning task. Subjects were instructed to classify so as to obtain the highest average reward, even if that required assigning an item to the less-likely (but higher-rewarded) category. In the subsequent information-search task, which itself was identical to Experiment 1, subjects could view only a single feature before making a categorization decision. In stark contrast to Experiment 1, the vast majority of subjects preferentially viewed Feature *R*, irrespective of whether doing so was adaptive given the search-task payoff structure. This refutes the idea that people have a general tendency to maximize accuracy in search. Experience-based learning does not necessarily lead to a preference to view the higher-probability gain feature. Rather, search behavior seems to have been driven by the (implicit or explicit) reward-scheme experienced during the previous classification learning task. The difficulty subjects had in learning in Experiment 2 suggests that accuracy is an especially intuitive payoff function; even with great encouragement, around half of the subjects failed to achieve

learning criterion in Experiment 2. The proximal driver of search-task information-acquisition behavior does not seem to be the actual search-task payoff structure, but rather something taken more directly from the learning task.

Experiment 3 found that, when probabilistic information and reward structures were both conveyed through summary statistics, people were indifferent regarding the usefulness of the alternative features. This was found for both symmetric and asymmetric reward schemes.

11.1   Did the search-task payoff manipulations have any effect?

Did the manipulations perhaps shift decision criteria in the desired direction, but just not far enough? Research on perceptual category learning—which is not concerned with information search, but rather with classification decisions when full stimulus information is available—has found that people are able to take asymmetric payoff functions into account, but only to a limited extent, when setting their decision criterion for classification decisions (e.g., Maddox & Bohil, 1998). We therefore analyzed whether the asymmetric search-task payoff manipulation could perhaps have shifted the decision criterion somewhat, in the correct direction, but not enough to cause subjects to preferentially view the feature with higher expected payoff. Note that this hypothesis is compatible with the ideas that people effectively learn environmental probabilities through their experience, and can search adaptively relative to asymmetric reward structures. It is just that under this hypothesis it is difficult to internalize an asymmetric reward structure.

Consider Environment 1, in which a [2 1 0 0] asymmetric reward structure applied in the search task. In this environment, one can calculate that if the reward function is a less-asymmetric [1.38 1 0 0], then Feature *A* and Feature *R* would have objectively equal usefulness (Equation 5). This hypothetical indifference reward function implies a decision criterion $c_x=42\%$, rather than the reward-maximizing decision criterion of 33% (Equation 2). The upshot is that subjects could have been slightly influenced by the search task payoff manipulation, with their internal decision criterion shifting in the appropriate direction (say, to 46%), and yet Feature *A* would have been objectively more useful than Feature *R*, relative to this slightly-asymmetric internalized reward function. In Environment 2, similarly, a person who internalized a reward function of the form [4.57 1 0 0], rather than the actual [10 1 0 0] payoff structure, would find Feature *R* and Feature *A* to be equally useful.

If making mistakes is intrinsically psychologically unpleasant, irrespective of the actual extrinsic payoff function (which in the present experiments included zero

Table 4: Informative search-task classification decisions.

Search-task rewards	Higher-rewarded response, given feature r_2	Exp. 1: Responses of Category x, given feature r_2	Exp. 2: Responses of Category x, given feature r_2
Symmetric	$x \approx y$	58% (62/106)	96% (152/158)
Asymmetric	x	72% (73/101)	100% (162/162)

Note. Subjects' responses, given that they had searched Feature R, and observed r_2, in the information-search task in Experiments 1 and 2 (aggregated over Environments 1 and 2). Given r_2, categories x and y are effectively equally likely. Under symmetric payoffs, both responses have the same expected value, but under asymmetric payoffs assigning the item to Category x has higher expected value. Results across the experiments suggest that the learning task is the primary determinant of the feature subjects view in the search task, and that there is little or no influence of the search-task payoff manipulation. The difference in proportion of Category x responses given feature r_2, under symmetric versus asymmetric rewards, is not statistically reliable (see text).

penalty for mistakes), this would also result in a less-extreme decision criterion c_x. Suppose a subject in the [2 1 0 0] payoff condition intuits one unit cost for each mistake, thereby internalizing an effective [2 1 1 1] reward function. This subject should adopt a decision criterion $c_x = 40\%$, closer to 50% than the optimal 33% criterion (Equation 2). Any intrinsic psychological cost of making mistakes, where that cost applies equally to each kind of mistake (i.e. where $m = n$), would have a similar effect in shifting the decision criterion towards 50%. Thus, Experiment 1's information-acquisition results do not preclude that there may be a small influence, in the appropriate direction, of the explicitly stated search-task reward function. Rather, the results show that such an influence, if it exists, is very small.

Search-task categorization responses, which required subjects to make decisions based on a single feature value, may provide some additional insights into the (in)efficacy of the search-task payoff manipulation. In cases where Feature r_2 was observed, the posterior probabilities of Category x and Category y were effectively equal (Table 4; Figure 2). Thus, if the reward manipulation had any effect whatsoever, subjects in the asymmetric reward conditions should strongly prefer to choose the more highly rewarded Category x, given Feature r_2, in the search task. Note further that, if either of the above hypotheses are true—namely that people applied a decelerating nonlinear function to the utilities (e.g. and perceived [2 1 0 0] payoffs as [1.38 1 0 0])—or if the intrinsic cost of error hypothesis is true (e.g., [2 1 0 0] payoffs are perceived as [2 1 1 1]), then subjects should still overwhelmingly respond Category x given Feature r_2, under asymmetric payoffs, in both environments. The data from Experiment 1 strongly contradict even these weaker hypotheses about the possible influence of the search-task payoff manipulation; only about 72% of subjects' responses were for Category x, given Feature r_2 (Table 4).

However, the raw data do suggest a trend in which subjects with asymmetric rewards had a higher propensity to choose Category x given r_2 than did subjects with symmetric rewards (Table 4). We used bootstrap sampling to estimate 95% confidence intervals for true proportion of Category x responses, given r_2, under symmetric vs. asymmetric rewards: $M_{\text{symmetric}}=58\%$, CI from 39% to 76% and $M_{\text{asymmetric}}=72\%$, CI from 54% to 87%. The highly overlapping confidence intervals show that there is no statistically reliable effect of the payoff manipulation.[9]

What about Experiment 2? In Experiment 2, almost all classification decisions, given feature r_2, were for Category x. Given this ceiling effect, it is not possible to address whether the search-phase reward manipulation had an effect in Experiment 2. What is overwhelmingly clear, in Experiment 2 and Experiment 1 alike, is that subjects meaningfully assimilated the *learning-task* reward structure.

11.2   Alternate explanations

People are often risk-averse, for instance by preferring a retirement investment with a smaller expected return but much-less-variable return to a highly variable investment with a higher expected return. When small amounts of money are involved, little risk aversion is usually observed. That makes risk-aversion an a priori improbable, though still conceivable, hypothesis in Experiments 1 and 2. Could risk-aversion explain these results, on the assumption that people understood the underlying probabilities and payoff structure? In the asymmetric payoff conditions, in each environment, although Feature A has

---

[9]A standard difference-of-proportions test would falsely assume that the underlying data (including successive responses from the same subject, which tend to be the same) are independent.

Table 5: Expected values (and standard deviations) of Features *A* and *R* (in €).

	Symmetric rewards		Asymmetric rewards	
	Feature *A*	Feature *R*	Feature *A*	Feature *R*
Env. 1	1.28 (0.96)	1.13 (0.99)	0.88 (0.77)	1.00 (0.94)
Env. 2	1.64 (0.77)	1.28 (0.96)	0.72 (0.86)	0.78 (0.92)

a lower expected payoff than Feature *R*, Feature *A* also has lower standard deviation in payoff (Table 5). Thus, it would be conceivable that risk aversion could explain subjects' preference to view Feature *A*, given asymmetric payoffs, in Experiment 1. In the symmetric payoff conditions, however, Feature *A* has higher expected payoff as well as lower standard deviation in expected payoff (Table 5). Thus, in Experiment 2, risk aversion cannot possibly explain why subjects overwhelmingly preferred to view Feature *R*, given symmetric payoffs. Risk aversion is not a possible explanation of results across the two experiments.

Can alternate optimal experimental design (OED) models (Nelson, 2005, 2008) or heuristic strategies (e.g., Martignon et al., 2008) explain people's choices of features to view? Different models, such as information gain and impact, make various claims about the usefulness of individual features, in each environment (Figure 5; Appendix, Table A1). For example, in Environment 1 the information gain (expected reduction in Shannon entropy) of Feature *R* is higher than that of Feature *A*, whereas in Environment 2, Feature *A* has a higher information gain than Feature *R*. Thus, information gain cannot explain why people preferred Feature *A* in both environments in Experiment 1, but Feature *R* in both environments in Experiment 2, nor why people had no preference in Experiment 3. The OED and heuristic models, as articulated to date, were designed to provide general-purpose strategies for information acquisition. They were not designed to apply to situation-specific payoff functions. Accordingly, these models do not predict changes as a function of learning- or of search-task reward manipulations, or information formats. Hence, these models cannot explain why search behavior differs between Experiments 1, 2, and 3.

Were the stakes not high enough? Nelson et al. (2010, Experiment 3, Condition 2), in a task with no external rewards or payoffs, found that a difference of as little as 4.5 percentage points in features' probability gain was enough to induce a strong preference to view the higher-probability-gain feature. The consistency across subjects in the present experiments, especially the preference to view Feature *R* in Experiment 2, suggests that subjects

did not pick features at random to view, but were in fact highly motivated.

Do people dispense with probabilities altogether, and learn according to experienced outcomes and expected rewards (i.e., learn only expected values of actions; e.g., Barron & Erev, 2003)? The learning data strongly contradict this idea. Experiment 1 demonstrated that both environments are learnable. However, in Experiment 2 subjects struggled a great deal with the learning task (Table 3; Appendix, Figure A1), due to the conflict configuration in each environment, in which Category *y* was more probable, but a Category *x* classification choice had higher expected reward. This should not be a problem for a purely expectation-based system, which could easily identify which categorization action has the highest expected reward, for each configuration, within a few hundred trials. Experiment 2 showed that it takes a great deal of training for human subjects to respond contrary to accuracy in the learning task. Anecdotal evidence for this also comes from one of our subjects who, after failing to learn to classify under asymmetric payoffs, stated "It feels weird to be wrong" in choosing the less-probable category in the conflict configuration. In any case, it is not trivial mathematically to go from the learning task to the search task, because the search-task requires marginalizing over the unobserved feature, and a preposterior analysis of the expected usefulness of each feature. Therefore, a simple reinforcement-based account of the learning task could not simultaneously explain search-task behavior.

11.3 Representations, decision strategies, and reward structures in learning and information search

For developing psychological theories, Anderson (1990) proposed beginning with minimal mechanistic assumptions, and making more specific processing assumptions only when necessary. Following the results of Nelson et al. (2010), we began with the idea that people can become familiar with environmental probabilities through experience-based learning, and that we would investigate people's goals for information search when asymmetric payoffs apply. In line with Bayesian decision theory (Savage, 1954) we assumed that people would have separate representation of beliefs (probabilities) and utilities (costs and benefits), which would allow them to determine possible questions' usefulness, relative to their goals.

Subjects in Experiments 1 and 2 appeared to have developed a reasonable understanding of environmental probabilities (Table 2). Nevertheless, they were unable to use that knowledge in a flexible way to identify the most useful query, given novel search-task payoff structures. These results point to the importance of more precisely

Figure 5: Information-search behavior: data and theoretical models. Dark grey represents Feature *A*, light grey Feature *R*. Empirical search-task results are displayed in the top row (% of subjects preferentially viewing Features *A* vs. *R*) and next-to-top row (mean views to Features *A* vs. *R*); subsequent rows show predictions of alternate informational OED models (Table A1). MaxVal and ZigVal (Martignon et al., 2008), two heuristic models, also prefer Feature *R*. None of these models captures the differences between Experiment 1 and Experiment 2, as none of these models makes different predictions according to the procedure during the categorization learning task. The final row, Learning-phase Reward, captures the idea that following experience-based learning people preferentially view whichever feature would have been most important, relative to the reward structure and goals in the learning task (see text and Figure 6).
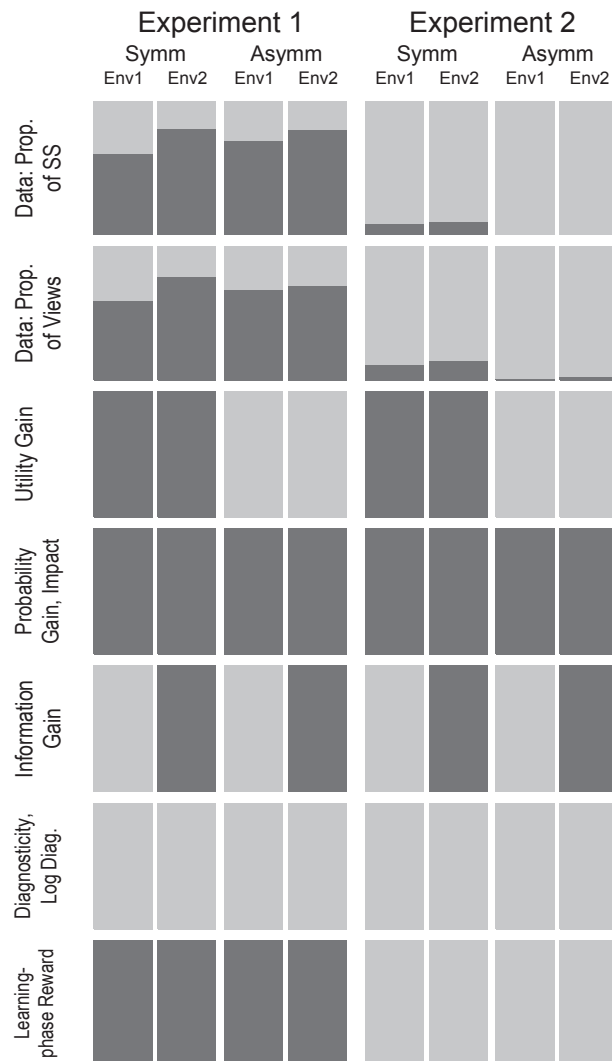
Figure 6: Decision trees that might be established during the learning task. Depending on the goal of the classification task (maximizing overall accuracy in Experiment 1 vs. maximizing rewards in Experiment 2), features' relative usefulness differs. In Experiment 1, subjects were trained to choose whichever category is most probable, given the presented stimulus. To most efficiently achieve this, with minimal feature views, Feature *A* should be the root node. By contrast, in Experiment 2 subjects learned to classify under asymmetric rewards, with the goal of categorizing stimuli in a way that maximizes expected reward. This goal is most efficiently achieved by first querying Feature *R*, which has higher usefulness than Feature *A* (i.e., higher utility gain). (In fact, categorizing stimuli based on the state of Feature *R* alone is sufficient to maximize expected rewards. Therefore, in the trees, both states of Feature *A* lead to the same decision.)
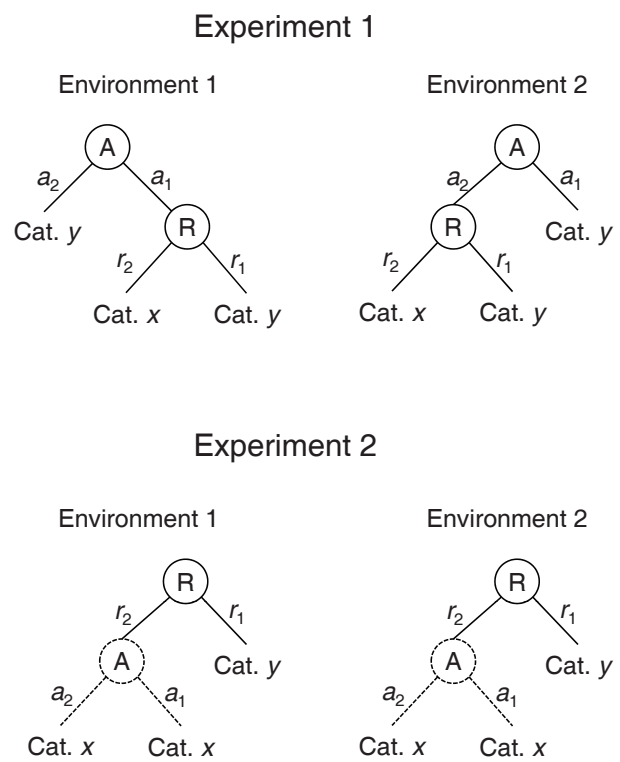
characterizing the algorithm-level (in Marr's, 1982, classification) cognitive processes and representations that develop during experience-based learning.

What learning processes might underlie the present experiments? Research on concept learning shows that, even in tasks in which people could in principle view every feature in each trial, they learn to allocate attention to efficiently view features sequentially (Blair, Watson, Walshe & Maj, 2009; Rehder & Hoffman, 2005). Computational models have been developed to explain the allocation of eye movements to specific stimulus features on Shepard, Hovland and Jenkins's (1961; Nosofsky, Gluck, Palmeri, McKinley & Glauthier, 1994) concept formation task (Love, 2010; Nelson & Cottrell, 2007).

With respect to the present findings, we propose that people do something similar, namely learn the decision tree that is most efficient—i.e., that requires the smallest number of feature views, on average—subject to the constraint of having optimal performance, relative to the (implicit or explicit) reward function during learning. When constructing these trees (Figure 6), we used the explicit asymmetric learning-task reward structure for Experiment 2, and a symmetric reward structure for Experiment 1. In this case, the result is the acquisition of a *fast-and-frugal decision tree* (Bergert & Olsson, in preparation; Green & Mehr, 1997; Luan et al., 2011; Martignon et al., 2008).

We hypothesize that this tree is what later drives people's search-task behavior: Choices of which feature to view reflect the learning-task search hierarchy, rather than any judgment of features' relative usefulness per se. This process works well if the search task retains the learning-task payoff structure, but can work poorly otherwise, even when the new payoff structure is as straightforward as monetary payoffs for accuracy.

To illustrate, consider the tree for Environment 1, in Experiment 1. When a stimulus is presented, one first looks at Feature A. If a_2 is observed, the stimulus can be assigned to Category y, as this is now the more probable category regardless of the state of Feature R (Figure 1). The state of R needs to be checked only if a_1 is observed, as the item is more likely to belong to Category y when $R=r_1$, but more likely to be x when $R=r_2$, given a_1. The tree structure is similar for Environment 2.

By contrast, in Experiment 2, where people had to assign stimuli to the higher rewarded category, Feature R is more useful, and is the root node of the search trees. In fact, people could in principle reach criterion performance by making decisions based on Feature R alone. Whenever $R=r_1$, stimuli should be assigned to Category y, and when $R=r_2$, to Category x (Figure 1). However, the different kinds of mistakes made for the various configurations during learning show that subjects considered both features (Appendix, Figure A1). Subjects' probability estimates also show that they were sensitive to the influence of the state of Feature A. For example, the median estimate for $P(x \mid a_1\ r_2)$ was 25%, but for $P(x \mid a_2\ r_2)$ was 90% (Table 2), close to the true values of 15% and 81%, respectively. Therefore, we kept both features in the tree for Experiment 2.

Introducing a new reward function in the search task requires rearranging the tree in some cases, to achieve the most efficient tree, and changing classification decisions associated with some exit nodes. Such a re-ordering may be difficult, and might even require new learning experiences and feedback. The announcement, via words and numbers, of a new search-task payoff structure was not enough, no matter how well people appeared to have assimilated the environmental probabilities in the learning task.

In sum, for development of a comprehensive theory of human information acquisition, the present results suggest (1) not taking a simplistic decision-theoretic view of probability learning as distinct from rewards that could drive search behavior, and (2) focusing on the nature of the learners' goals, decision strategies, and any specific habits in the learning process (e.g., eye movement search ordering among the features).

11.4 The value of information in real-world decisions

What about real-world search decisions? Physicians sometimes use fast-and-frugal trees to search for information and to make medical diagnoses (Fischer et al., 2002; Green & Mehr, 1997).[10] This fits with our ideas on the use of search-and-classification trees learned through experience, which can adapt to situation-specific costs and benefits through the arrangement of the exit nodes (Luan et al., 2011). Such simple decision trees are also used to train medical personnel for making classification decisions during mass casualty incidents, enabling first responders to classify and prioritize victims according to the severity of their injuries, as they are easy to apply, even under stressful conditions (Super, 1984, as cited in Luan et al., 2011).

In applied contexts within medical decision making and environmental toxicology, some studies have explicitly employed quantitative value of information analyses (Yokota & Thompson, 2004; see also Benish, 2002, 2009). However, the explicit use of this methodology has developed slowly, perhaps in part because of the complicated mathematics of the real world. For instance, it is not

---

[10]There may be a general bias in favor of medical testing, and conflicts of interest that can arise when the practitioner is specifically paid to conduct a test. We view these as exogenous issues that could affect medicine severely, but which are separate from people's underlying information-search capacities.

trivial to identify a reasonable utility function to incorporate both financial costs, and changes in life expectancy and quality of life. Whichever utility function is adopted, it is important to appropriately discount future returns according to when they will occur. However, Yokota and Thompson note that around half of studies did not report use of temporal discounting in their analyses.

It does seem that in some circumstances, for instance in deciding to test for a rare but serious disease, people may appropriately take asymmetric payoffs into account. Another example would be airport security personnel's thorough pre-flight screening of a person exhibiting mildly suspicious behavior. It need not be more likely than not, or even very likely at all, that the individual present a threat; the costs of failing to detect a bomb are high, justifying low thresholds for screening. In these examples the basic reward structures (e.g., the high cost of missing a serious-yet-treatable disease, or of missing a bomb on an airplane) are very intuitive. There is also ample opportunity to train practitioners on the payoff structure applicable in particular medical or security contexts. We take our present results to suggest, at a minimum, that without a situation with an intuitive and easily-internalized rationale for a particular asymmetric payoff structure, spontaneously adaptive search behavior could be difficult to achieve.

11.5 Future directions

One important area for future research will be to directly compare scenarios with intuitive and strongly asymmetric payoff structures with theoretically-identical abstract scenarios. Baron and Hershey (1988) described different diseases abstractly. One manipulation could compare abstract disease names with names that cue strongly asymmetric consequences, for instance where one disease is almost certainly deadly, but the other disease is akin to having the flu. Experience-based learning, perhaps in a way similar to our present experiments, could be used in both cases. It would also be interesting to investigate whether the type of learning-phase feedback matters, when the extrinsic search-task payoff structure is intuitive.

From the perspective of human *learners*, not all utilities are created equal. Learning is considerably easier under an implicit symmetric payoff structure (or at least, without an explicitly asymmetric payoff structure), as seen in Experiment 1. Learning is considerably more difficult under an asymmetric payoff structure, as seen in Experiment 2, where roughly half of subjects failed to achieve the learning criterion. Search behavior following learning, however, is an entirely different animal. Following experience-based learning in a particular environment, with a particular reward structure, people are able to spontaneously identify which features would be most

useful to query, even in abstract probabilistic simulated plankton categorization tasks. This is a positive finding that supports a theory of diverse cognitive abilities. It suggests that even in environments with arbitrary reward structures (e.g., in which a disease is very serious, and its treatment—if given unnecessarily—is fairly harmless), if people learn the environment, individual features, *and the reward structure* through their own experience, they will spontaneously have very good intuitions as to which queries are most useful, even in situations where the most useful queries do not improve classification accuracy.

References

Anderson, J. R. (1990). *The Adaptive Character of Thought*. Hillsdale, NJ: Lawrence Erlbaum.

Baron, J. (1985). *Rationality and Intelligence*. Cambridge, England: Cambridge University Press.

Baron, J., & Hershey, J. C. (1988). Heuristics and biases in diagnostic reasoning: I. Priors, error costs, and test accuracy. *Organizational Behavior and Human Decision Processes*, *41*, 259–279.

Baron, J., Beattie, J., & Hershey, J. C. (1988). Heuristics and biases in diagnostic reasoning: II. Congruence, information, and certainty. *Organizational Behavior and Human Decision Processes*, *42*, 88–110.

Barron, G., & Erev, I. (2003). Small feedback-based decisions and their limited correspondence to description-based decisions. *Journal of Behavioral Decision Making*, *16*, 215–233.

Benish, W. A. (2002). The use of information graphs to evaluate and compare diagnostic tests. *Methods of Information in Medicine*, *41*, 114–118.

Benish, W. A. (2009). Intuitive and axiomatic arguments for quantifying diagnostic test performance in units of information. *Methods of Information in Medicine*, *48*, 552–557.

Bergert, F. B., & Olsson, H. (in preparation). A new method for constructing fast and frugal trees that makes them faster, more frugal, and more accurate.

Blair, M. A., Watson, M. R., Walshe, C. R., & Maj, F. (2009). Extremely selective attention: eye-tracking studies of the dynamic allocation of attention to stimulus features in categorization. *Journal of Experimental Psychology: Learning, Memory and Cognition*, *35*, 1196–1206.

Cavagnaro, D. R., Myung, J. I., Pitt, M. A., & Kujala, J. V. (2010). Adaptive design optimization: a mutual information-based approach to model discrimination in cognitive science. *Neural Computation*, *22*, 887–905.

Chater, N., & Oaksford, M. (2008). *The Probabilistic Mind: Prospects for Rational Models of Cognition.*

Oxford: Oxford University Press.

Cosmides, L., & Tooby, J. (1996). Are humans good intuitive statisticians after all? Rethinking some conclusions from the literature on judgment under uncertainty. *Cognition*, *58*, 1–73.

Crupi, V., Tentori, K., & Lombardi, L. (2009). Pseudodiagnosticity revisited. *Psychological Review*, *116*, 971–985.

Fedorov, V. V. (1972). *Theory of Optimal Experiments*. New York: Academic Press.

Fischer, J. E., Steiner, F., Zucol, F., Berger, C., Martignon, L., Bossart, W., Altwegg, M., & Nadal, D. (2002). Using simple heuristics to target macrolide prescription in children with community-acquired pneumonia. *Archives of Pediatric and Adolescent Medicine*, *156*, 1005–1008.

Gigerenzer, G., & Hoffrage, U. (1995). How to improve Bayesian reasoning without instruction: Frequency formats. *Psychological Review, 102*, 684–704.

Gigerenzer, G., Todd, P., & the ABC research group (1999). *Simple Heuristics That Make Us Smart*. New York: Oxford University Press.

Good, I. J. (1950). *Probability and the Weighing of Evidence*. New York: Griffin.

Green, L., & Mehr, D. R. (1997). What alters physicians' decisions to admit to the coronary care unit? *Journal of Family Practice*, *45*, 219–226.

Hadar, L., & Fox, C. R. (2009). Information asymmetry in decision from description versus decision from experience. *Judgment and Decision Making*, *4*, 317–325.

Hahn, U., & Oaksford, M. (2007). The rationality of informal argumentation: A Bayesian approach to reasoning fallacies. *Psychological Review*, *114*, 704–732.

Hertwig, R., & Pleskac, T. J. (2010). Decisions from experience: Why small samples? *Cognition*, *115*, 225–237.

Hertwig, R., Barron, G., Weber, E. U., & Erev, I. (2004). Decisions from experience and the effect of rare events in risky choice. *Psychological Science*, *15*, 534–539.

Klayman, J. (1987). An information theory analysis of the value of information in hypothesis testing. Retrieved May 23, 2005, from http://www.chicagocdr.org/cdrpubs/

Klayman, J., & Ha, Y.-W. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Psychological Review*, *94*, 211–228.

Knowlton, B. J., Squire, L. R., & Gluck, M. A. (1994). Probabilistic classification learning in amnesia. *Learning and Memory*, *1*, 106–120.

Krauss, S., Martignon, L., & Hoffrage, U. (1999). Simplifying Bayesian inference: The general case. In L. Magnani, N. Nersessian and P. Thagard (Eds.), *Model-based reasoning in scientific discovery* (pp. 165–179). New York: Kluwer Academic/Plenum.

Kruschke, J. K., & Johansen, M. K. (1999). A model of probabilistic category learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *25*, 1083–1119.

Kullback, S., & Leibler, R. A. (1951). Information and sufficiency. *Annals of Mathematical Statistics*, *22*, 79–86.

Lindley, D. V. (1956). On a measure of the information provided by an experiment. *Annals of Mathematical Statistics*, *27*, 986–1005.

Love, B. C. (2010). Looking to learn, learning to look: attention emerges from cost-sensitive information sampling. Paper presented at the 2010 meeting of the Psychonomic Society, November 18–21, 2010, St Louis, Missouri, USA. Retrieved November 17, 2011, from http://www.psychonomic.org/pdfs/Abstracts10.pdf

Luan, S. Schooler, L., & Gigerenzer, G. (2011). A signal detection analysis of fast-and-frugal trees. *Psychological Review*, *118*, 316–338.

Maddox, W. T. (2002). Toward a unified theory of decision criterion learning in perceptual categorization. *Journal of the Experimental Analysis of Behavior*, *78*, 567–595.

Maddox, W. T., & Bohil, C. J. (1998). Base-rate and payoff effects in multidimensional perceptual categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *24*, 1459–1482.

Maddox, W. T., & Bohil, C. J. (2003). A theoretical framework for understanding the effects of simultaneous base-rate and payoff manipulation on decision criterion learning in perceptual categorization. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *29*, 307–320.

Maddox, W. T., & Dodd, J. (2001). On the relation between base-rate and cost-benefit learning in simulated medical diagnosis. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *27*, 1367–1384.

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. San Francisco: W. H. Freeman and Company.

Martignon, L., Katsikopoulos, K. V., & Woike, J. K. (2008). Categorization with limited resources: A family of simple heuristics. *Journal of Mathematical Psychology*, *52*, 352–361.

Myung, J. I., & Pitt, M. A. (2009). Optimal experimental design for model discrimination. *Psychological Review*, *116*, 499–518.

Navarro, D. J., & Perfors, A. F. (2011). Hypothesis generation, sparse categories, and the positive test strategy. *Psychological Review*, *118*, 120–134.

Nelson, J. D. (2005). Finding useful questions: On Bayesian diagnosticity, probability, impact, and infor-

mation gain. *Psychological Review*, *112*, 979–999.

Nelson, J. D. (2008). Towards a rational theory of human information acquisition. In N. Chater and M. Oaksford (Eds.), *The Probabilistic Mind: Prospects for Rational Models of Cognition* (pp. 143–163). Oxford: Oxford University Press.

Nelson, J. D. (2009). Naïve optimality: Subjects' heuristics can be better-motivated than experimenters' optimal models. *Behavioral and Brain Sciences*, *32*, 94–95.

Nelson J. D., & Cottrell, G. W. (2007). A probabilistic model of eye movements in concept formation. *Neurocomputing*, *70*, 2256–2272.

Nelson, J. D., McKenzie, C. R. M., Cottrell, G. W., & Sejnowski, T. J. (2010). Experience matters: Information acquisition optimizes probability gain. *Psychological Science*, *21*, 960–969.

Nosofsky, R. M., Gluck, M., Palmeri, T. J., McKinley, S. C., & Glauthier, P. (1994). Comparing models of rule-based classification learning: A replication and extension of Shepard, Hovland, and Jenkins (1961). *Memory & Cognition*, *22*, 352–369.

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, *101*, 608–631.

Oaksford, M., & Chater, N. (1996). Rational explanation of the selection task. *Psychological Review*, *103*, 381–391.

Oaksford, M., & Chater, N. (2007). *Bayesian Rationality*. Oxford: Oxford University Press.

Raiffa, H., & Schlaifer, R. O. (1961). *Applied Statistical Decision Theory*. Cambridge, MA: Division of Research, Graduate School of Business Administration, Harvard University.

Rehder B., & Hoffman, A. B. (2005). Eyetracking and selective attention in category learning. *Cognitive Psychology*, *51*, 1–41.

Savage, L. J. (1954). *The Foundations of Statistics*. New York: Wiley.

Shannon, C. E. (1948). A mathematical theory of communication. *The Bell System Technical Journal*, *27*, 379–423, 623–656.

Skov, R. B., & Sherman, S. J. (1986). Information-gathering processes: Diagnosticity, hypothesis-confirmatory strategies, and perceived hypothesis confirmation. *Journal of Experimental Social Psychology*, *22*, 93–121.

Slowiaczek, L. M., Klayman, J., Sherman, S. J., & Skov, R. B. (1992). Information selection and use in hypothesis testing: What is a good question, and what is a good answer? *Memory & Cognition*, *20*, 392–405.

Super, G. (1984). *START: A Triage Training Module*. Newport Beach, CA: Hoag Memorial Hospital Presbyterian.

Trommershäuser, J., Maloney, L. T., & Landy M. S. (2003a). Statistical decision theory and trade-offs in the control of motor response. *Spatial Vision*, *16*, 255–275.

Trommershäuser, J., Maloney, L. T., & Landy M. S. (2003b). Statistical decision theory and rapid, goal-directed movements. *Journal of the Optical Society of America*, *20*, 1419–1433.

Trommershäuser, J., Maloney, L. T., & Landy M. S. (2008). Decision making, movement planning and statistical decision theory. *Trends in Cognitive Science*, *12*, 291–297.

Trope, Y., & Bassok, M. (1982). Confirmatory and diagnosing strategies in social information gathering. *Journal of Personality and Social Psychology*, *43*, 22–34.

Ungemach, C., Chater, N., & Stewart, N. (2009). Are probabilities overweighted or underweighted, when rare outcomes are experienced (rarely)? *Psychological Science*, *20,* 473–479.

von Winterfeldt, D., & Edwards, W. (1982). Costs and payoffs in perceptual research. *Psychological Bulletin*, *91*, 609–622.

Wason, P. C. (1966). Reasoning. In B. M. Foss (Ed.), *New Horizons in Psychology* (pp. 135–151). Harmondsworth, England: Penguin.

Wason, P. C. (1968). Reasoning about a rule. *Quarterly Journal of Experimental Psychology*, *20*, 273–281.

Wells, G. L., & Lindsay, R. C. L. (1980). On estimating the diagnosticity of eyewitness nonidentifications. *Psychological Bulletin*, *88*, 776–784.

Wu, S.-W., Trommershäuser, J., Maloney, L. T., & Landy, M. S. (2006). Limits to human movement planning in tasks with asymmetric gain landscapes. *Journal of Vision*, *6*, 53–63.

Yokota, F., & Thompson, K. M. (2004). Value of information literature analysis: A review of applications in health risk management. *Medical Decision Making*, *24*, 287–298.

Appendix

Table A1: Alternative optimal experimental design (OED) models.

Probability Gain (*PG*)
(Baron, 1985)

$$u_{PG}(d) = \max_i \left(P(c_i \mid d) \right) - \max_i \left(P(c_i) \right)$$

Information gain (*IG*)
(Lindley, 1956)

$$u_{IG}(d) = \sum_{i=1}^{n} P(c_i) \, log_2 \frac{1}{P(c_i)} \;-\; \sum_{i=1}^{n} P(c_i \mid d) \, log_2 \frac{1}{P(c_i \mid d)}$$

Kullback-Leibler divergence (*KL*)
(Kullback & Leibler, 1951)

$$u_{KL}(d) = \sum_{i=1}^{n} P(c_i \mid d) \, log_2 \frac{P(c_i \mid d)}{P(c_i)}$$

Impact (*Imp*)
(Wells & Lindsay, 1980)

$$u_{Imp}(d) = \sum_{i=1}^{n} \left| P(c_i) - P(c_i \mid d) \right|$$

Bayesian Diagnosticity (*BD*)
(Good, 1950)

$$u_{BD}(d) = max \left(\frac{P(d \mid c_1)}{P(d \mid c_2)}, \frac{P(d \mid c_2)}{P(d \mid c_1)} \right)$$

log_{10} Diagnosticity ($log_{10}BD$)

$$u_{log_{10}BD}(d) = log_{10} \left(max \left(\frac{P(d \mid c_1)}{P(d \mid c_2)}, \frac{P(d \mid c_2)}{P(d \mid c_1)} \right) \right)$$

Note. Alternative optimal experimental design (OED) models proposed to quantify the usefulness of a datum *d* (a feature value, test result, answer) to identify an object's category $C = \{c_1, \ldots, c_n\}$ (for reviews, see Nelson 2005, 2008). The expected usefulness of a query (test, question, experiment) is calculated as the average usefulness of the data, where the usefulness of each datum *d* is weighted by its probability (Equation 3). See Figure 5 for model predictions for the current experiments. Bayesian Diagnosticity and Log_{10} Diagnosticity are only defined for binary categories.

Table A2: Analysis of Baron and Hershey's (1988) scenarios in which study subjects chose which of two medical tests (T1 or T2) was most useful (Experiment 1, Cases 5–11).

Case	5	6	7	8	9	10	11
P(disease)	0.75	0.50	0.25	0.50	0.50	0.25	0.75
T1 true positive	0.84	0.88	0.92	0.84	0.92	0.88	0.88
T1 false positive	0.32	0.28	0.24	0.32	0.24	0.28	0.28
T2 true positive	0.76	0.72	0.68	0.76	0.68	0.72	0.72
T2 false positive	0.08	0.12	0.16	0.08	0.16	0.12	0.12
harm	1	1	1	1	3	1	3
neglect	1	1	1	3	1	3	1
T1 expected utility	0.050	0.300	0.050	0.100	0.100	0.450	0.450
T2 expected utility	0.050	0.300	0.050	0.100	0.100	0.450	0.450
T1 probability gain	0.050	0.300	0.050	0.260	0.340	0.010	0.090
T2 probability gain	0.050	0.300	0.050	0.340	0.260	0.090	0.010
T1 information gain	0.167	0.289	0.280	0.212	0.383	0.212	0.231
T2 information gain	0.280	0.289	0.167	0.383	0.212	0.231	0.212
T1 impact	0.195	0.300	0.255	0.260	0.340	0.225	0.225
T2 impact	0.255	0.300	0.195	0.340	0.260	0.225	0.225
T1 diagnosticity	3.096	4.343	7.177	3.307	6.213	4.771	3.914
T2 diagnosticity	7.177	4.343	3.096	6.213	3.308	3.914	4.771
T1 $\log_{10}$ diagnosticity	0.480	0.615	0.816	0.507	0.749	0.657	0.573
T2 $\log_{10}$ diagnosticity	0.816	0.615	0.480	0.749	0.507	0.573	0.657
SS' prefs.,	T2	T1≈T2	T1	T1≈T2	T1 ≈T2	T1	T2
(percentage format)	$t=-2.29$		$t=2.94$			$t=3.62$	$t=-3.50$
SS' prefs,	T1≈T2	T1≈T2	T1≈T2	T1≈T2	T1≈T2	T1	T2
(odds format)						$t=3.18$	$t=-9.75$

Note. The scenarios were described in terms of the prior probability of the disease, the true and false positive rate of each test, the harm caused by treating a patient who does not have the disease, and the cost of neglecting to treat a patient who does have the disease. In each case, the two tests had equal utility. OED models (Table A1, see also Nelson, 2005, 2008) of the relative usefulness of each test were calculated. The two lowermost rows give subjects' preferences and the *t*-statistic reported by Baron and Hershey, for cases in which subjects significantly (uncorrected two-tail $p < .05$) preferred one of the tests. There are two numbers for each scenario, reflecting responses from different informational formats. The first version used a percentage (e.g., "75 percent") to denote the prior probability that the patient had the disease; the second version used odds ("three to one") to describe the probability that the patient had the disease. It does not appear that any of the OED models offer a plausible explanation of subjects' choices on this task. Probability gain and information gain wrongly predict that Test 2 will be preferred in Case 10, and that Test 1 will be preferred in Case 11. Impact predicts a preference in Cases 8 and 9, which was not observed, and is tied in Case 10 and Case 11, where subjects showed preferences. Bayesian diagnosticity and $\log_{10}$ diagnosticity show strong preferences in Case 8 and Case 9, whereas subjects were statistically indistinguishable from indifference. In Case 8 and Case 9, the trend (which Baron and Hershey reported as nonsignificant in each instance) was for subjects to prefer Test 1 in Case 8, and Test 2 in Case 9. All the OED models, however, have the opposite preference, namely for Test 2 in Case 8, and Test 1 in Case 9.

Figure A1: Learning data from Environment 2, Experiments 1 and 2. Each subject is one row; each column is one feature configuration, sorted according to their frequency from left to right. Trials are plotted from top to bottom, and from left to right, for a particular subject and a particular configuration. In each trial, a decision that is consistent with which category is most probable (Experiment 1) or is most rewarded (Experiment 2), is plotted with a white rectangular pixel. Suboptimal decisions are plotted with black rectangular pixels. The top two panels show learning data from Experiment 1, in which people's task was to classify stimuli according to which category is most probable (i.e., with no explicit reward function during learning). Most people (38/40) achieved the learning criterion. The bottom two panels show learning data from Experiment 2, in which an explicit asymmetric reward function applied in the learning phase. Only 19 out of 40 people achieved the learning criterion. The results show that subjects struggled a great deal with the conflict configuration (second column from left), for which accuracy and reward conflict (i.e., subjects had to choose the less likely category in order to maximize expected reward).

$P(a_2 r_2) = 0.38$ $P(a_1 r_2) = 0.34$ $P(a_1 r_1) = 0.23$ $P(a_2 r_1) = 0.06$
$P(x \mid a_2 r_2) = 0.81$ $P(x \mid a_1 r_2) = 0.15$ $P(x \mid a_1 r_1) = 0$ $P(x \mid a_2 r_1) = 0$