COGNITIVE SCIENCE
A Multidisciplinary Journal

Cognitive Science 40 (2016) 100–120
Copyright © 2015 Cognitive Science Society, Inc. All rights reserved.
ISSN: 0364-0213 print / 1551-6709 online
DOI: 10.1111/cogs.12220

# Self-Directed Learning Favors Local, Rather Than Global, Uncertainty

Douglas B. Markant,[a] Burr Settles,[b] Todd M. Gureckis[c]

[a]*Center for Adaptive Rationality, Max Planck Institute for Human Development*
[b]*Duolingo, Inc.*
[c]*Department of Psychology, New York University*

Received 3 September 2013; received in revised form 14 October 2014; accepted 24 October 2014

**Abstract**

Collecting (or "sampling") information that one expects to be useful is a powerful way to facilitate learning. However, relatively little is known about how people decide which information is worth sampling over the course of learning. We describe several alternative models of how people might decide to collect a piece of information inspired by "active learning" research in machine learning. We additionally provide a theoretical analysis demonstrating the situations under which these models are empirically distinguishable, and we report a novel empirical study that exploits these insights. Our model-based analysis of participants' information gathering decisions reveals that people prefer to select items which resolve uncertainty between two possibilities at a time rather than items that have high uncertainty across all relevant possibilities simultaneously. Rather than adhering to strictly normative or confirmatory conceptions of information search, people appear to prefer a "local" sampling strategy, which may reflect cognitive constraints on the process of information gathering.

*Keywords:* Information sampling; Self-directed learning; Active learning; Machine learning

## 1. Introduction

A cornerstone of many educational philosophies is that people learn more effectively when they direct their own learning experience (Boekaerts, 1997; Bruner, 1961). Although there are many ways that such control might influence learning, one important factor is the ability to choose among different sources of information, a decision-making process we refer to as self-directed information sampling (Gureckis & Markant, 2012). A

Correspondence should be addressed to Douglas Markant, Center for Adaptive Rationality, Max Planck Institute for Human Development, Lentzeallee 94, 14195 Berlin, Germany. E-mail: markant@mpib-berlin.mpg.de

canonical example of this type of decision making is a doctor deciding which diagnostic test to perform on a sick patient (e.g., either an MRI or a blood test). Given a set of illnesses that may be causing the patient's symptoms, the physician must decide whether a particular test, whose outcome is as yet uncertain, is most likely to reveal the correct diagnosis. The doctor plays a vital role in this assessment, relying on knowledge of the patient and potential illnesses to determine the value of a new source of information.

Relative to our understanding of how people evaluate alternatives and manage uncertainty in economic decision making (Glimcher & Rustichini, 2004; Kahneman & Tversky, 1979; von Neumann & Morgenstern, 1944), less is known about how people judge the usefulness of new sources of information during learning in order to make sampling decisions. The goal of the present paper is to advance our understanding of this aspect of human decision making and learning.

### *1.1. Two views on human information sampling*

Existing studies of information sampling have contributed to two, seemingly contradictory, theoretical positions. On one hand, a large body of work on hypothesis testing and reasoning suggests that people prefer to sample information that is consistent with their existing beliefs even though that information may be ineffective for learning (Klayman, 1995; Nickerson, 1998), a behavior we refer to as *confirmatory sampling*. One well-known example of this is the positive test strategy (PTS), whereby people focus on observations that are positive examples of their current hypothesis, without accounting for how those observations relate to alternative hypotheses (Klayman & Ha, 1987, 1989). In general, confirmatory sampling suggests a singular focus on an existing hypothesis and the data that are likely to be observed if it were true, while alternative hypotheses are neglected, often resulting in information that is less useful for learning.[1]

In contrast, a separate line of research has argued that people's information sampling decisions are consistent with normative principles, according to which alternative hypotheses are integrated to determine what information is most useful for adjudicating between them (see Nelson, 2005 for review). This normative framework, based on theories of "optimal experimental design" first developed in the statistics literature (Fedorov, 1972; Lindley, 1956), has accounted for sampling decisions in a range of information search tasks (Nelson, McKenzie, Cottrell, & Sejnowski, 2010; Oaksford & Chater, 1994), including relatively open-ended or complex domains such as visual search (Najemnik & Geisler, 2005), spatial search (Gureckis & Markant, 2009; Markant & Gureckis, 2012), causal structure learning (Steyvers, Tenenbaum, Wagenmakers, & Blum, 2003), and sequence learning (Austerweil & Griffiths, 2011). As demonstrated by Nelson (2005), different normative models within this framework (e.g., *probability gain* or *information gain*) may predict distinct sampling decisions depending on the task structure and the learner's goal, but as a group they share the principle of integrating across all possible hypotheses to determine the most useful source of information.

### 1.2. An intermediate position

Although these two perspectives have long been considered to be in opposition to each other, they can also be viewed as two endpoints along a continuum representing the degree to which information from multiple hypotheses contribute to behavior. Confirmatory sampling is consistent with a single hypothesis controlling behavior while normative information selection usually implies combining information from all viable hypotheses. Where people fall along this continuum may depend on their ability to consider alternative hypotheses and their relationships to possible outcomes, which may be limited in complex problems involving large numbers of hypotheses or other demands. For example, individual differences in working memory capacity predict how many alternatives influence judgments about the probability of a focal hypothesis, thereby determining how well people's judgments correspond to normative predictions (Dougherty & Hunter, 2003; Newstead, Thompson, & Handley, 2002; Sprenger et al., 2011). Predicting how people make sampling decisions may thus require an understanding of the cognitive constraints that limit performance in any particular task.

This perspective on information sampling, with confirmatory and normative sampling representing two extreme positions, raises a number of questions. First, there is the problem of formalizing intermediate models that make use of alternative hypotheses but to a lesser extent than predicted by normative theories. This idea is closely related to recent models of approximate Bayesian inference, which link sub-optimal learning or decision making to an impoverished representation of the hypothesis space (Griffiths, Vul, & Sanborn, 2012; Sanborn, Griffiths, & Navarro, 2010). There have been few attempts as yet, however, to define such a framework in the context of human information sampling (but see Steyvers et al., 2003, for a similar approach in a causal learning setting). In what follows, we describe one approach to formalizing intermediate models inspired by contemporary research in active machine learning (AML).

Second, it is as yet unclear what task or environmental circumstances might lead to sampling behavior that is better described by an intermediate model. For example, for certain kinds of hypothesis spaces, confirmatory sampling is consistent with normative goals, precluding the need to consider more than one alternative (Austerweil & Griffiths, 2011; Klayman & Ha, 1987; Navarro & Perfors, 2011; Nelson & Movellan, 2001). Learners who correctly account for these constraints might be expected to show more or less confirmatory sampling depending on the problem space. Similarly, previous work has shown that prior experience with a domain is associated with normative sampling, such as when dealing with familiar materials (Cox & Griggs, 1982; McKenzie, 2006) or in problems involving social information gathering (Trope & Mackie, 1987).

One relatively unexplored possibility that we consider in the present paper is that intermediate strategies are likely to manifest in sufficiently complex problems. One strategy for collecting information in complex or multivariate domains is to decompose a problem into simpler components and to reduce "local" sources of uncertainty. For example, when multiple features may be related to an outcome, a learner might hold one feature constant

while varying the other across multiple samples (Avrahami et al., 1997; Rottman & Keil, 2012), which is known as the "control of variables" strategy and is essential to scientific reasoning. Isolating variables helps people to search the space of potential hypotheses (Klahr & Dunbar, 1988) and is a key part of "learning to learn" about complex concepts (Kuhn & Dean, 2005). As a result of using this strategy during self-directed learning, sampling decisions may be better described by a model that focuses on local sources of uncertainty relevant to subsets of alternative hypotheses, rather than strictly confirmatory or normative accounts.

In this paper, we present a direct test of the idea that people prefer local sources of uncertainty when making sampling decisions, using a category-learning paradigm in which they control the selection of training items. We show that in this kind of problem, uncertainty about how to classify an item is directly related to how informative it is about the true category rule, raising the possibility that people rely on their uncertainty about how to predict the outcomes of potential queries in order to decide between them. This proposal is directly inspired by research on AML, in which such uncertainty sampling is a common, computationally efficient method for selecting training data for artificial classifiers. We describe a range of sampling models that vary in the degree to which they integrate information about alternative categories to predict what information is useful to sample. We then present an experiment designed to test between these alternative accounts of information sampling. After reporting the results of our experiment and subsequent modeling, we discuss the implications of this work for our understanding of how people direct their own learning.

## 2. Uncertainty sampling in AML

Effectively gathering information during learning is a problem that faces machine learners just as it affects human learners. A crucial factor in machine learning is the availability of labeled training data, which often requires costly human annotation. For example, credit card companies rely on automated systems for detecting fraudulent activity, but a human reviewer needs to label such instances before they can be used for supervised training of a classifier. Given the high costs of labeling, only instances that will improve the performance of the model should be selected for review. AML research has explored how selections should be made in order to maximize the accuracy of a machine learning model (Settles, 2012), with applications in a wide range of settings, including text classification (Olsson, 2009), natural language processing (Settles & Craven, 2008), and recommendation systems (Rubens, Kaplan, & Sugiyama, 2011).

Early work on AML (e.g., Cohn, Ghahramani, & Jordan, 1996; Mackay, 1992) drew upon the same framework of optimal experimental design that has guided recent psychological research. Accordingly, the same set of normative models explored in studies of human information sampling have been applied to problems involving machine classifiers. These models are often "prospective" in that they estimate the value of an

observation by simulating the effect of each of its possible outcomes on the current model. For example, given a potential training item, *information gain* formalizes the reduction in uncertainty that would be achieved for each possible labeling (e.g., a reviewer identifying a transaction as fraudulent or not), and the overall expected value of selecting that item is found by weighing each outcome's value by its predicted likelihood of occurring.

However, determining the best selection strategy in this way is computationally intractable in many machine learning applications. As a result, researchers have also developed methods for sampling an item based on a model's current uncertainty in how to classify it (a measure which does not require estimating the effect of observing that item). In comparison to prospectively evaluating the optimal decision, using the model's current classification uncertainty to decide what to learn about is less costly and in many cases achieves similar improvements in efficiency.

In the following sections, we describe a set of simple models, originally proposed in the AML literature (Settles, 2012), which predict the value of new information based on classification uncertainty. Importantly, these models vary in the way that uncertainty about alternative categories is integrated to make a sampling decision. For a potential training item $x$, there is a set of possible category labels $\{y_1, y_2, \ldots\}$ that could result from observing that item. The probability of each category label being assigned to $x$ is given by the distribution $p(y|x)$. The key way that the models differ is in how that probability distribution is used to predict whether $x$ should be sampled.

### 2.1. Information gain and label entropy

The first uncertainty sampling model we consider is *label entropy*, which is defined as the Shannon entropy over the label distribution $p(y|x)$ for an observation $x$ (see Settles, 2012):

$$LE(x) = -\sum_i p(y_i|x) ln\, p(y_i|x). \tag{1}$$

Shannon entropy measures the disagreement in predictions across possible labels for an item.[2] Entropy is highest when the model predicts each label equally, and it is lowest when a certain prediction is made for a single label. Thus, *label entropy* quantifies the amount of predictive uncertainty the observer has for item $x$. Those items that the observer has difficulty predicting the class membership for are assumed to be useful to observe, as they represent some aspect of the world where the learner is uncertain and would benefit from feedback.

While *label entropy* is commonly used in AML systems, it is also closely related to normative models that are used in cognitive psychology (Klayman & Ha, 1987; Nelson & Movellan, 2001; Oaksford & Chater, 1994; Steyvers et al., 2003). In fact, when hypotheses are deterministic (i.e., the likelihood of an observation is either 1 or 0 for all possible hypotheses), *label entropy* is formally equivalent to *information gain*, a prospective

normative model that has been used to account for sampling behavior in a number of tasks. Information gain is defined as the reduction in uncertainty resulting from a new observation $x$. Having observed a previous set of observations $D = \{\langle x, y\rangle_1, \langle x, y\rangle_2, \ldots\}$, uncertainty is given by the Shannon entropy of the posterior distribution (now defined over a hypothesis space $\mathcal{H}$ with uniform prior likelihood):

$$I(p(h|D)) = -\sum_{h\in\mathcal{H}} p(h|D) ln\, p(h|D) \tag{2}$$

$$= -ln\frac{1}{N} = ln\, N \tag{3}$$

where $N$ indicates the number of hypotheses in $\mathcal{H}$ that are consistent with the observations in $D$. Information gain is the reduction in uncertainty that would occur as a result of observing that item $x$ has label $y_i$:

$$IG(\langle x, y_i\rangle) = I(p(h|D)) - I(p(h|\langle x, y_i\rangle, D)) \tag{4}$$

$$= ln\, N - ln\, N_i = ln\frac{N}{N_i} \tag{5}$$

where $N_i$ is the number of remaining hypotheses that are consistent with $x$ having label $y_i$. Since the outcome of the observation is unknown, the expected information gain is weighted by the probability of each outcome occurring:

$$E[IG(\langle x, y\rangle)] = \sum_i p(y_i|x) IG(\langle x, y_i\rangle) \tag{6}$$

$$= \sum_i p(y_i|x) ln\frac{N}{N_i} \tag{7}$$

$$= -\sum_i p(y_i|x) ln\frac{N_i}{N} \tag{8}$$

$$= -\sum_i p(y_i|x)\ ln\, p(y_i|x) \tag{9}$$

Note that Eq. 9 is the same as the definition of *label entropy* in Eq. 1. Thus, for deterministic hypotheses, the expected information gain of a query (evaluated one step ahead) is equivalent to the entropy measured over its possible outcomes. This is maximized when an item is "globally" uncertain such that all outcomes have equal probability. If hypotheses have uniform prior probability (as is the case here), this occurs when each outcome is predicted by an equal number of plausible hypotheses.

This analysis provides a connection between one uncertainty sampling model (*label entropy*) and a prospective model (*information gain*) that has been used as a normative standard for understanding human behavior. When dealing with deterministic hypotheses the two approaches are identical, and sampling according to label entropy is consistent with a normative model that takes into account the full set of hypotheses.

### 2.2. Margin sampling

While focusing on items that are globally uncertain or unpredictable seems intuitively useful, there is reason to expect that it may not be the sampling strategy humans use, particularly when learning in complex, multivariate environments. One natural strategy, not captured by *label entropy*, might be to decompose a complex task into a set of simpler problems. We can formalize the strategy of focusing on separate components in a sampling model that values uncertainty about any boundary between only two categories. *Label margin* predicts that the learner will prefer instances for which the likelihood of any two categories is similar, independent of any other categories. When the label distribution $P(y|x)$ is ordered from highest to lowest probability $\{p_1, p_2, \ldots\}$, with $p_1$ indicating the highest label probability, *label margin* is based on the difference between the two most likely labels for $x$:

$$LM(x) = 1 - (p_1 - p_2) \tag{10}$$

Critically, *label margin* is not maximized for only those items about which the learner is globally uncertain. Instead, there is a preference for local sources of uncertainty between subsets of potential outcomes. Whereas *label entropy* integrates information about all possible labelings of an item, *label margin* relies on the two most likely outcomes, disregarding the rest of the label distribution. As discussed above, this model thus reflects an intermediate sampling model in that a subset of possible alternatives (e.g., different categories) is used to evaluate whether a potential training item is worth learning about.

### 2.3. Most certain

Finally, previous work on hypothesis testing suggests that people may prefer items that they can already classify with relative confidence. People have a well-documented bias toward seeking positive evidence of the hypothesis they are considering (Klayman & Ha, 1989; Wason, 1960). To quantify this strategy, we define the *most certain* measure as:

$$MC(x) = max(p(y|x)) \tag{11}$$

The predictions of this model directly contrast those of *label entropy*, with the highest value assigned to items that can already be classified with confidence. The *most certain* measure is one way of instantiating confirmatory sampling—it shows a preference for items for which the learner has a strong prediction about the category label.

## 3. Empirical studies of information sampling during category learning in humans

In a recent study, we examined the interaction of self-directed information selection and category learning (Markant & Gureckis, 2014). In this experiment, people learned about two categories of "antennas" that varied along two perceptual dimensions (circles that differed in size and the orientation of a central line segment, see Fig. 1) and received one of two television stations (CH1 or CH2). We compared a self-directed condition, in which participants designed stimuli to learn about, with a standard, passive condition in which instances were generated from predefined distributions. A main finding from this study was that for simple uni-dimensional rules, self-directed learners acquired the correct category rule faster than passive learners (see also Castro et al., 2008).

In light of evidence that self-directed sampling can speed learning, it is important to understand how people decide what data to collect. Given a potential observation, what information do people rely on to decide if it will be useful? As we have proposed above, one aspect that may explain a person's decision to sample an item is his or her uncertainty in how to classify it. Intuitively, a self-directed learner should direct his or her attention to items that are high in uncertainty while ignoring items that can already be confidently classified or predicted. Consistent with this strategy, the pattern of stimuli sampled by self-directed learners in our previous study revealed that participants systematically directed their samples toward the category boundary as the task progressed (see Fig. 1B), suggesting a preference for items that they were uncertain about how to classify.

However, from that study we were unable to identify which sampling model best accounted for people's decisions. One reason for this is that we did not directly measure

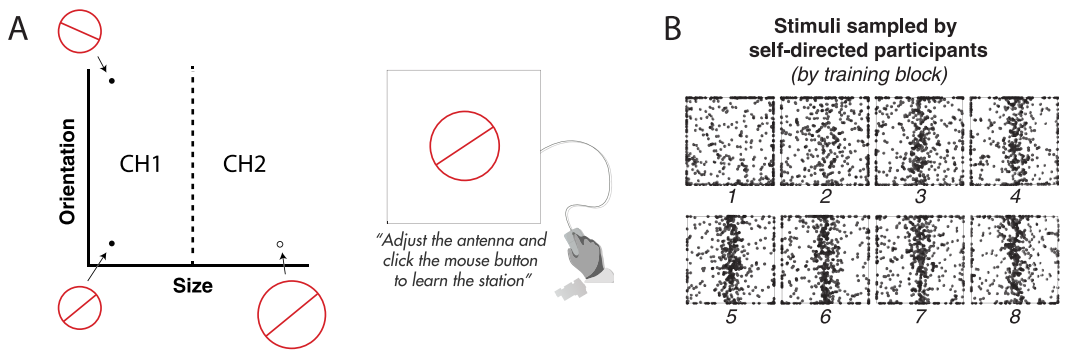


Fig. 1. (A) Abstract stimulus space used in Markant and Gureckis (2014), which is adapted for the current study. Stimuli were "loop antennas" that varied in diameter and orientation. Antennas were assigned to one of two categories based on the channel they received ("CH1" or "CH2"). Participants "designed" a stimulus they wanted to learn about using the mouse. Clicking the mouse button reveals the category membership of the item. (B) The pattern of sampling behavior observed by self-directed learners in Markant and Gureckis (2014) across eight training blocks. Each dot represents a single stimulus which was selected by a participant and is plotted in the stimulus space. In the first block, participants distributed samples widely over the entire stimulus space but then gradually focused their choices on the region surrounding the category boundary. (The category boundary has been rotated to always appear vertically in these plots.)

participants' uncertainty about the items they decided to learn about, and uncertainty can't be directly inferred from the items chosen. A given item might be associated with either high or low subjective uncertainty depending on how much a person has learned, regardless of where it falls in the stimulus space.

In addition, the binary classification task precludes a comparison of the *label entropy* and *label margin* models, both of which make highly similar predictions in that task, as seen in the top row of Fig. 2. Each heatmap describes the value assigned to a potential observation depending on the learner's uncertainty in how to classify it. For example, an item that can be confidently classified (e.g., $p(y|x) = (1,0)$, corresponding to the left edge of the heatmap) would be assigned a low value by *label entropy* and *label margin*, but a high value by *most certain*. For the binary classification problem, *label entropy* and *label margin* make highly similar predictions about how items will be valued. Items close to the center of the space have the highest value, and the ranking of items is identical between both models, making it difficult to distinguish between them. Settles (2012) observed that the predictions of these models diverge when considering more complex categorization tasks. For example, in a ternary classification task (see bottom row of Fig. 2), *label entropy* predicts a preference for items for which all three classes are likely (e.g., near the junction of the category boundaries). In contrast, *label margin* assigns the maximum value to items for which one category is highly unlikely but the learner is uncertain about the other two (shown in Fig. 2 by the high predicted value along the radial axes of the simplex, including the midpoints of each edge). In short, this model predicts that samples are likely to be allocated close to any boundary between two categories.

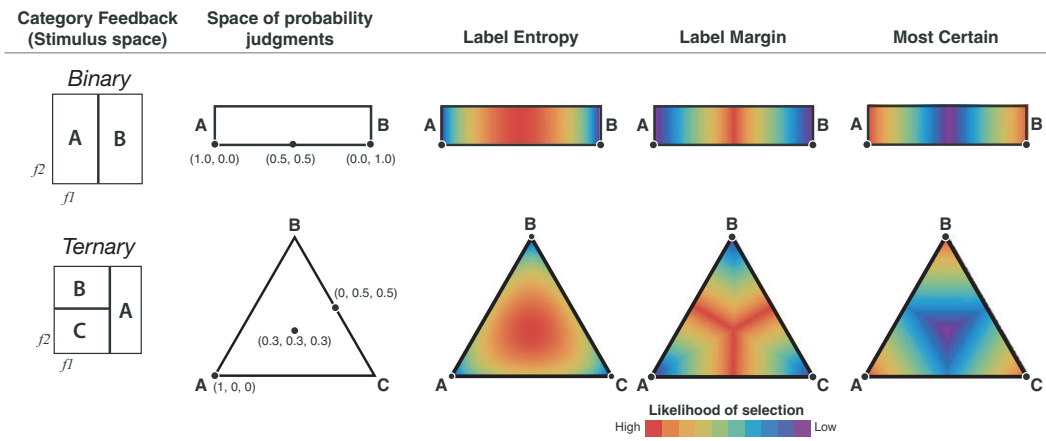


Fig. 2. Comparing predictions of sampling models (red = more highly valued choices, blue = less valued choices). Top: For a binary classification problem, a new observation in stimulus space will correspond to a location on the probability judgment scale, where the leftmost point reflects confidence that the observation will be classified "A" and the rightmost point reflects confidence it will be classified "B." For the binary problem, the predictions of *label entropy* and *label margin* are highly similar. Bottom: In a ternary classification problem, an item in stimulus space will correspond to a location in the three-category simplex depending on the learner's uncertainty. Here, the predictions of *label entropy* and *label margin* diverge, allowing us to test which of the two models better accounts for sampling behavior.

The design of the present experiment capitalizes on this distinction by extending the paradigm used in Markant and Gureckis (2014) to a ternary classification problem, thereby allowing us to separate the predictions of the sampling models presented above. Participants in the experiment collected information by sampling new instances and receiving feedback about their category membership. In order to obtain an estimate of a learner's uncertainty at any point in time, they also judged the likelihood that a sampled instance belonged to each of the three categories (before receiving feedback about its true label).

The goal of our design was to use these subjective judgments to test which model provides the best account of sampling decisions. Based on past work, we expected either the *most certain* or *label entropy* models would provide the best-fit to a majority of subjects. Furthermore, we hypothesized that participants best described by *label entropy* would be more successful at learning the classification rule because this sampling strategy makes more efficient use of the information in the task. To foreshadow, both of these predictions appear disconfirmed: Our results suggest that most participants were better accounted for by the *label margin* strategy and that these subjects were more likely to learn the target rule relative to other participants.

## 4. Experiment

### *4.1. Participants*

Sixty undergraduates at New York University participated in the study for psychology course credit. One participant was excluded for ending the task early. The experiment was run on standard desktop computers in a single 1-h session.

### *4.2. Stimuli*

The category label associated with each stimulus was deterministically defined by a ternary classification rule of the form shown in the bottom row of Fig. 2. In addition to the structure that is shown, three more rules were created through different rotations (90, 180, and 270 degrees) of the same boundaries in stimulus space. Each participant was randomly assigned to one of the four rules (leading to $N = \{16,14,15,14\}$ for the four rules) and a random mapping of labels ("CH1," "CH2," "CH3") to categories.

Training stimuli were chosen by participants according to the procedure below. Stimuli for each test block were generated by subdividing the stimulus space into a grid of 36 equally sized regions and generating a random stimulus from each region.

### *4.3. Procedure*

Participants were instructed that the stimuli in the experiment were television "loop antennas" and that each unique antenna received one of three channels (CH1, CH2,

or CH3). Their goal was to learn the difference between the three types of antennas so that they could correctly classify new antennas during the test blocks. The experiment alternated between training blocks (8 trials each) and test blocks (36 trials each). Participants were told that the experiment would end when they correctly classified 34 of 36 test items (94%) in a single test block. If a participant failed to reach that criterion, the experiment ended after 16 rounds or at the end of an hour (whichever occurred first).

#### *4.3.1. Training trials*

Participants sampled a new antenna by adjusting the size and orientation and receiving feedback about which channel was received. They were instructed that they should design antennas they thought were useful and that would help them to predict the TV channel for other designs they had not yet tested.

Each training trial began with the presentation of a randomly generated antenna. Participants then adjusted its size and orientation by moving the mouse from left to right while holding either the "Z" or "X" key, respectively. Only one dimension could be changed at a time, but participants could make any number of changes and were self-paced. When the stimulus was the desired size and orientation, they pressed the mouse button to obtain feedback.

Following their selection but before feedback was displayed, the participant judged the likelihood that the antenna would receive each of the channels using a series of three rating scales (shown in Fig. 3). The three scales were presented independently such that only one was visible at a time. When each scale appeared, the participants clicked on a location in the scale according to their belief that the antenna they had designed would receive that channel. Labels were provided to indicate how different locations corresponded to different degrees of subjective probability, but participants were instructed that they could respond anywhere within the rating scale. A response was required for each scale, and there was no time limit for entering the response. The initial position of the mouse cursor within each scale was randomized, allowing us to evaluate whether responses were influenced by the starting position.

Following the probability judgment, feedback was displayed above the selected antenna for 4s. The feedback display included the correct category label for that antenna, as well as the probability judgments that the participant had just entered (see Fig. 3, bottom). This helped the participant to evaluate the accuracy of his or her prediction, given the true category label.

#### *4.3.2. Test trials*

Each block of eight training trials was followed by 36 test trials. On each test trial, a single item was presented in the center of the display and the participant classified the item according to the channel he or she believed it was most likely to receive. A response was required to complete the trial, and participants responded at their own pace. No feedback was provided on individual test trials. At the end of each test block participants were told their overall accuracy from the test block they just completed.

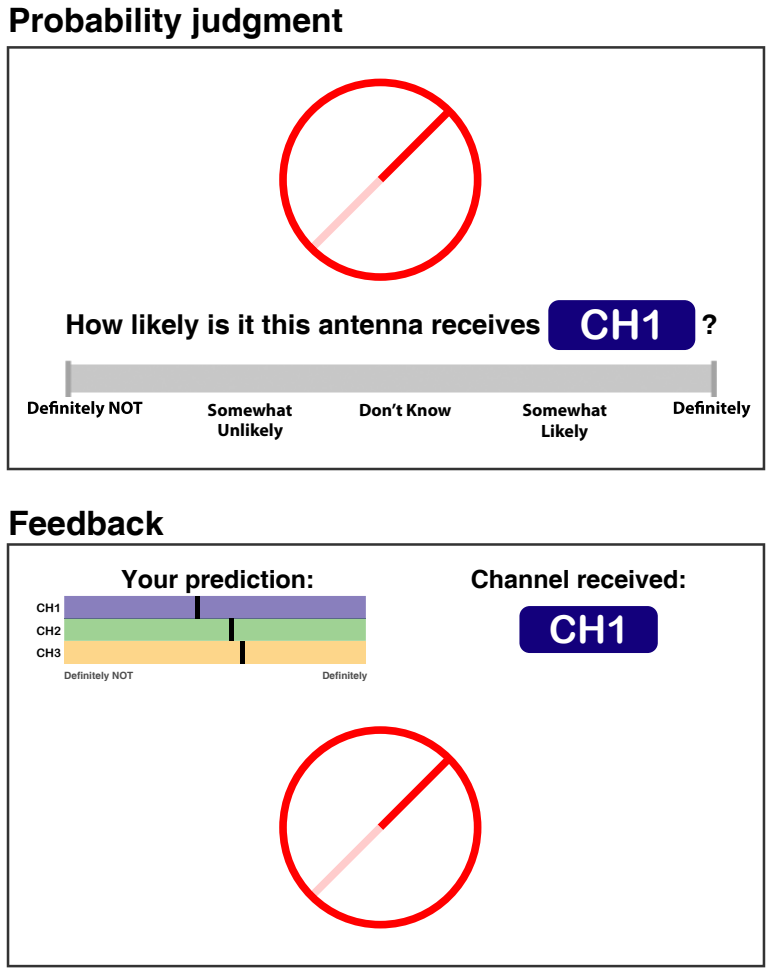


Fig. 3. Top: Probability judgments were entered by clicking on a scale for each of the three categories (CH1, CH2, and CH3). Bottom: Probability judgments were displayed alongside the category label during feedback.

## 4.4. Results

### 4.4.1. Classification performance

Thirty-six participants (62%) successfully reached the accuracy criterion of 94% correct within the available time.[3] Of those participants, the average number of blocks to criterion was 6 ($SD$ = 3.1). For the remaining participants, the average number of blocks completed was 9.8 ($SD$ = 3.6).

### 4.4.2. Probability judgments

On each training trial, the participant judged the likelihood that the stimulus they selected belonged to each of the three categories, resulting in three values between 0 and

1 based on where they clicked within the response scale. In order to verify that participants were not simply responding based on the position of the cursor, for each rating we measured the difference between the initial (random) cursor position and the participant's response. Any trial in which the response did not differ by more than 5% of the scale from the initial cursor position (for at least one of the three ratings) was classified as a non-response. This cutoff was chosen based on our intuition about how much the position would vary if the person simply clicked wherever the cursor appeared, and it was intended to be a conservative criterion to exclude any partial judgments that could bias the results. Using this exclusion criterion, the average proportion of non-responses was .33 (*SD* = .17). Two participants were excluded from further analysis because their proportion of non-responses was more than three standard deviations above the group mean (83% and 94%).

#### 4.4.3. Overall model fits

Our first goal was to assess the overall fit of the three sampling models to each participant's set of probability judgments. For each model, we used rejection sampling to create a relative frequency histogram that approximated a probability density function over the three-category simplex (shown in the bottom row of Fig. 2). We calculated the proportion of samples (out of 1 million) that fell within each of 400 equally sized triangular bins. Each triplet of ratings was then normalized to sum to one, corresponding to a location within the same space. The estimated probability of a set of ratings $r_t$ then corresponded to the proportion of samples falling within the same bin as that judgment, $F(r_t)$. For each participant's full set of judgments, we fit each of the sampling models using the softmax choice rule (Luce, 1959; Sutton & Barto, 1998):

$$p(r_t) = \frac{e^{F(r_t)/b}}{\sum_{z \in Z} e^{F(z)/b}} \tag{12}$$

where $b$ is the temperature parameter and $Z$ is the full set of possible response bins (see Table 1 for quartiles of best-fit values of $b$). A value of $b$ close to zero implies that the learner consistently chose samples with the maximum value according to the model, while higher values of $b$ indicate increasingly random sampling. We then calculated the log-likelihood of each judgment made by a participant and summed

Table 1
Quartiles for distribution of best-fit values of temperature parameter

| | Q1 | Q2 | Q3 |
|---|---|---|---|
| *LE* | .0003 | .0004 | .0010 |
| *LM* | .0006 | .0010 | .0031 |
| *MC* | .0016 | .0021 | .0053 |

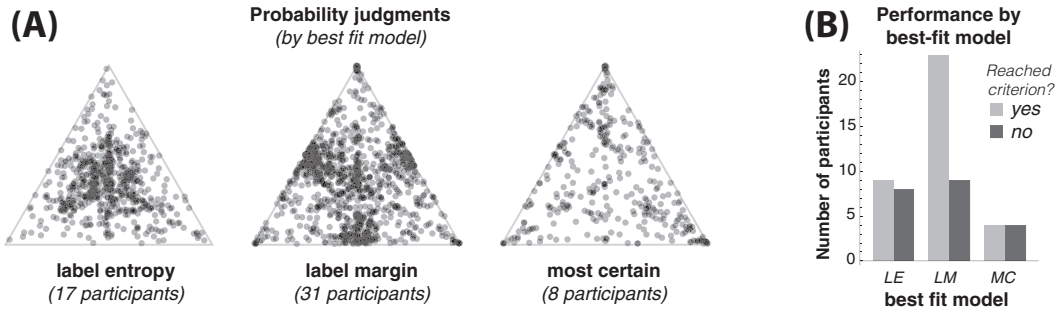


Fig. 4. (A) Normalized probability judgments plotted within the three-category simplex for participants best-fit by each of the three models (see Fig. 2 for reference). (B) Number of participants who reached the learning criterion, divided by which sampling model provided the best-fit.

across all trials to get an overall score for each model. These scores were calculated using only those trials that were not classified as non-responses as described above.

Classifying participants according to the model with the highest log-likelihood, we found that 17 people (30%) were best-fit by *label entropy*, 32 people (56%) were best-fit by *label margin*, and the remaining 8 people (14%) were best-fit by *most certain*. Judgments made by participants, separated by the best-fitting model, are plotted within the three-category simplex in Fig. 4A, with each point representing the probability judgment for a single sample chosen by a participant. A higher density of points reflects an increased tendency (as a group) to sample stimuli in a given region of the probability judgment space. Upon visual inspection, the overall pattern for each group corresponds to the predictions of the best-fitting model shown in Figure 2.

#### *4.4.4. Relating sampling decisions to learning*

We next examined whether success at learning the rule was related to the sampling strategy reflected in participants' probability judgments (see Fig. 4B). Of those participants who reached the learning criterion, 23 (64%) were best-fit by *label margin*, 9 (25%) were best-fit by *label entropy*, and the remaining 4 (11%) were best-fit by *most certain*. Among participants who failed to reach criterion, 8 people were best-fit by *label entropy*, 9 were best-fit by *label margin*, and the remaining 4 were best-fit by *most certain*.

Although a majority of participants were best-fit by *label margin* overall, in general participants had relatively heterogenous strategies, with some of their selections more likely according to one of the other models. Thus, rather than dividing people based on the best-fitting model, we next tested whether the frequency with which people made selections most consistent with each model was related to whether they learned the rule. We divided the probability space into three equally sized regions corresponding to the predictions of each model (see Fig. 5, left). Note that using this method, the *label margin*

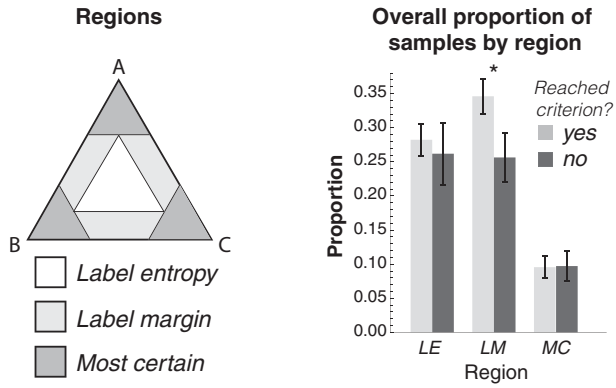


Fig. 5. Left: Each judgment was classified according to the model assigning it the highest likelihood, effectively dividing the probability space into three regions. Right: The frequency of sampling items within the *label margin* region (but not other regions) was related to successful learning.

region does not include samples in the center of the simplex that are also predicted by *label entropy*.

We then categorized each probability judgment according to the region it fell within and measured the proportion of samples that a participant made in each of the three regions. Participants who reached the learning criterion made more samples in the *label margin* region than those failed to reach the criterion ($t(55) = 2.04$, $p < .05$). However, there was no difference between learners and non-learners in the frequency of sampling in the *label entropy* ($t(55) = 0.44$, $p = .6$) or *most certain* ($t(55) = 0.06$, $p = .95$) regions. Thus, successful learning in the task was associated with increased sampling of items that are most consistent with the *label margin* model.[4]

## 5. Discussion

In many real-world contexts people can control what information forms the basis of their learning and decision making. As a result, their performance often hinges on how they make sampling decisions, and in particular, whether those decisions facilitate new learning or simply reinforce existing beliefs. Previous research has painted a mixed picture in this regard. A long history of work on hypothesis testing supports the view that people tend to be "confirmatory" information samplers, looking for data that is consistent or strongly predicted by a focal hypothesis (Beattie & Baron, 1988; Klayman, 1995; Wason, 1968). In contrast, theories of normative information acquisition propose that people search for information in a manner consistent with optimizing the amount of information conveyed by their actions (Klayman & Ha, 1987; Nelson, 2005; Oaksford & Chater, 1994).

Under many conditions, these two theories make qualitatively distinct predictions as to the kind of information people will collect as they learn. However, it is also useful to

consider that they are two extremes in terms of how information about alternatives contributes to sampling decisions. In many domains, the generation and/or use of alternatives has been linked to how people reason and the quality of their decisions (Dougherty, Gettys, & Thomas, 1997; Evans, 2007; Newstead et al., 2002; Sprenger et al., 2011; Thomas, Dougherty, Sprenger, & Harbison, 2008). The consideration of alternatives has often been proposed as a critical factor mediating whether people engage in confirmatory or normative sampling (Sanbonmatsu, Posavac, Kardes, & Mantel, 1998), but these behaviors have often been treated as a dichotomy rather than as endpoints on a continuum. In even moderately complex problems that require people to reason about alternative hypotheses, the best account for their sampling behavior may correspond to a more intermediate position on this dimension.

### *5.1. Margin sampling: A preference for "local" classification uncertainty*

In this study, we found clear evidence for an intermediate sampling behavior, with a majority of participants best described overall by the *label margin* model. According to this model, items are preferred when they are likely to belong to two categories (independent of the likelihood of any remaining categories). This formalizes the idea that people seek out information to reduce "local" sources of uncertainty related to a subset of alternatives. In contrast, normative models (including *information gain*, which is identical to *label entropy* in our task) predict that people should sample according to a "global" measure of uncertainty, with the strongest preference for items that are equally likely to belong to any category. For an ideal observer in this task, globally uncertain items convey the greatest amount of information about possible classification rules, and margin sampling should only reduce the efficiency of learning as it will tend to rule out fewer alternative rules per item sampled (a gap which increases with the number of possible categories).

In addition, our results showed that a tendency toward margin sampling was associated with more successful learning of the target rule, relative to the other models tested. People who were best-fit by *label margin* were more likely to reach the learning criterion overall. Moreover, across all participants, the frequency with which people sampled items that were strongly predicted by *label margin* (but not *label entropy*) was related to whether they reached the learning criterion.

Evidence of margin sampling suggests a general preference for a local form of exploration, but it is not diagnostic about the exact underlying sampling process. One possibility is that, when faced with a multidimensional task like the one used here, people decompose the problem into simpler components. This kind of piecemeal strategy may be more effective when it is difficult to simultaneously consider many alternatives or to process information about multiple feature dimensions. Thus, margin sampling may reflect an adaptation whereby people isolate individual components to learn in succession, akin to the "control of variables" strategy that is an important part of scientific thinking more generally (Klahr & Dunbar, 1988; Kuhn & Dean, 2005). Alternatively, even if people do not rely on this particular strategy, they may have learned over the course of prior

experience that margin sampling is more effective and, as a result, ascribe higher utility to items predicted by that model.

Another reason people might prefer the label margin strategy is that it is easier to process feedback from dichotomous tests. For example, if an individual is completely uncertain about an item, the feedback that it belongs to category "A" may be less easy to integrate with his or her understanding of the category (since he or she did not make a strong prediction to begin with). In contrast, for items where the participant is uncertain between only two categories, feedback can be decisive. This constraint may become increasingly important as the number of category labels or discrete outcomes that are possible for any given query increases.

One limitation of the current study is our dependence on participants' self-reported probability judgments, since there were no costs for failing to report their subjective belief accurately. Importantly, we were able to measure a failure to respond by randomly initializing the cursor position before each rating, and we restricted our analysis to those judgments for which people made an effortful response. However, it remains a possibility that probability judgments were biased in some way through the self-report procedure, and a goal of ongoing work is to verify the validity of margin sampling using alternative measures of subjective uncertainty.

### *5.2. Relation to other modeling approaches*

Although normative models are generally intended to be computational accounts of sampling behavior, it is useful to consider their implications for process-level models of how those decisions are made. A model like *information gain* implies a prospective process in which the expected outcomes of an observation are combined to estimate its effect on one's current beliefs. We have argued that it is unlikely that people perform this prospective evaluation, particularly in problems that are unfamiliar or complex, but that people may use simpler forms of uncertainty that can lead to similar gains. This trade-off between computational costs and optimality is shared by work on AML, which has shown that uncertainty sampling is a widely applicable means for improving the efficiency of training (Settles, 2012).

Specifically, we have shown that when hypotheses are deterministic, *information gain* is equivalent to sampling by *label entropy*, according to which people need only assess their uncertainty about how to classify an item. Thus, uncertainty sampling may be a general-purpose strategy that is less demanding but often consistent with normative principles. A similar point has been made in recent analyses showing that positive testing can be equivalent to *information gain* under certain conditions (Austerweil & Griffiths, 2011; Klayman & Ha, 1987; Navarro & Perfors, 2011). For example, Austerweil and Griffiths (2011) showed that when hypotheses make deterministic predictions about the next event in a sequence, testing the prediction of the most probable hypothesis (e.g., asking "is the next event A?" and receiving yes/no feedback) will maximize *information gain*. Our analysis is complementary in that it applies to the selection of queries without making a specific prediction (e.g., simply asking "what is this?" and receiving a label). When

hypotheses make deterministic predictions about the category membership of different items, uncertainty about an item's label is directly related to the amount of information it is expected to convey in this setting.

Within the framework of uncertainty sampling, we provided an example of an intermediate sampling process in the *label margin* model. Our approach is similar to examples of "rational process models" that embody optimal decision-making with respect to an approximate representation of the hypothesis space (Griffiths et al., 2012; Sanborn et al., 2010), the fidelity of which may range from a single-point estimate up to the full distribution of alternatives. This kind of graded representation could result from a limited capacity for storing alternatives in memory, leading to information search decisions that aim to reduce uncertainty about the "local" set of alternatives currently being considered (Dougherty, Thomas, & Lange, 2010; Thomas et al., 2008). However, further work is necessary to determine whether this kind of preference is caused by limits on memory or hypothesis generation rather than being learned over the course of experience with similar problems.

Although we have focused on margin sampling in the present report, people might pool information about any subset of alternatives when evaluating potential samples. For example, another model used in AML research, known as *least confident*, evaluates selections based on confidence in the single most likely label. Although its predictions are relatively similar to those of *label margin* in our study, *least confident* might represent another intermediate form of sampling that is efficient in many problems that people face. In general, in rule-based reasoning it is likely that the number of alternatives people consider at a given time is relatively low, and that forms of local uncertainty sampling provide the best account of how people decide to collect information.

## 6. Acknowledgments

This work was supported by grant number BCS-1255538 from the National Science Foundation and the Intelligence Advanced Research Projects Activity (IARPA) via Department of the Interior (DOI) contract D10PC20023 to T.M.G. The U.S. Government is authorized to reproduce and distribute reprints for Governmental purposes notwithstanding any copyright annotation thereon. The views and conclusions contained herein are those of the authors and should not be interpreted as necessarily representing the official policies or endorsements, either expressed or implied, of IARPA, DOI, or the U.S. Government.

## Notes

1. The more general term *confirmation bias* has been used to refer to both confirmatory forms of sampling (such as the PTS) as well as biased responses to evidence (e.g., overweighting data consistent with a favored hypothesis). Note that a

confirmatory *sampling* process on its own does not imply a bias in the learner's beliefs, but it can cause learning to be less efficient than other strategies (see Klayman, 1995).

2. Note that the unit of measure for entropy depends on the base of the logarithm in Eq. 1, but different units are related by a constant multiple (e.g., 1 *nat* is equal to 1.44 *bits* of information).
3. No differences were found between the four variations of the category rule at any point in the analysis (G tests of independence, all $p > .05$)
4. Although the non-response criterion was chosen prior to the experiment, we conducted additional analyses with different criteria to check whether our results depended on the original value. All results were highly similar across different criteria, with one important change that the difference in proportion of *LM* samples was marginal at other criteria (e.g., when the non-response criterion is doubled to .1, the resulting $p$-value is .056). These additional analyses are available upon request.

## References


Austerweil, J., & Griffiths, T. (2011). Seeking confirmation is rational for deterministic hypotheses. *Cognitive Science*, *35*, 499–526.

Avrahami, J., Kareev, Y., Bogot, Y., Caspi, R., Dunaevsky, S., & Lerner, S. (1997). Teaching by examples: Implications for the process of category acquisition. *The Quarterly Journal of Experimental Psychology*, *50*(3), 586–606.

Beattie, J., & Baron, J. (1988). Confirmation and matching biases in hypothesis testing. *The Quarterly Journal of Experimental Psychology*, *40*(2), 269–297.

Boekaerts, M. (1997). Self-regulated learning: A new concept embraced by researchers, policy makers, educators, teachers, and students. *Learning and instruction*, *7*(2), 161–186.

Bruner, J. (1961). The act of discovery. *Harvard Educational Review*, *31*(1), 21–32.

Castro, R., Kalish, C., Nowak, R., Qian, R., Rogers, T., & Zhu, X. (2008). Human active learning. *Advances in Neural Information Processing Systems*, *21*, 241–248.

Cohn, D., Ghahramani, Z., & Jordan, M. (1996). Active learning with statistical models. *Journal of Artificial Intelligence Research*, *4*, 129–145.

Cox, J., & Griggs, R. (1982). The effects of experience on performance in Wason's selection task. *Memory and Cognition*, *10*(5), 496–502.

Dougherty, M., & Hunter, J. (2003). Hypothesis generation, probability judgment, and individual differences in working memory capacity. *Acta Psychologica*, *113*(3), 263–282.

Dougherty, M., Gettys, C., & Thomas, R. (1997). The role of mental simulation in judgments of likelihood. *Organizational Behavior and Human Decision Processes*, *70*, 135–148.

Dougherty, M., Thomas, R., & Lange, N. (2010). Toward an integrative theory of hypothesis generation, probability judgment, and hypothesis testing. *Psychology of learning and motivation*, *52*, 299–342.

Evans, J. (2007). *Hypothetical thinking: Dual processes in reasoning and judgement*. New York: Psychology Press.

Fedorov, V. V. (1972). *Theory of optimal experiments*. New York: Elsevier.

Glimcher, P., & Rustichini, A. (2004). Neuroeconomics: The concilience of the brain and decision. *Science*, *306*, 447–452.

Griffiths, T., Vul, E., & Sanborn, A. (2012). Bridging levels of analysis for probabilistic models of cognition. *Current Directions in Psychological Science*, *21*(4), 263–268.

Gureckis, T. M., & Markant, D. (2009). Active learning strategies in a spatial concept learning game. In N. Taatgen, H. van Rijn, L. Schomaker, & J. Nerbonne (Eds.), *Proceedings of the 31st Annual Conference of the Cognitive Science Society* (pp. 3145–3150). Austin, TX: Cognitive Science Society.

Gureckis, T., & Markant, D. (2012). Self-directed learning: A cognitive and computational perspective. *Perspectives in Psychological Science*, *7*(5), 464–481.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decisions under risk. *Econometrica*, *47* (2), 263–291.

Klahr, D., & Dunbar, K. (1988). Dual space search during scientific reasoning. *Cognitive Science*, *12*, 1–48.

Klayman, J. (1995). Varieties of confirmation bias. *Psychology of Learning and Motivation*, *32*, 385–418.

Klayman, J., & Ha, Y. (1987). Confirmation, disconfirmation, and information in hypothesis testing. *Research on Judgment and Decision Making: Currents*, *94*(2), 211–228.

Klayman, J., & Ha, Y. (1989). Hypothesis testing in rule discovery: Strategy, structure, and content. *Journal of Experimental Psychology: Learning*, *15*(4), 596–604.

Kuhn, D., & Dean, D. (2005). Is developing scientific thinking all about learning to control variables? *Psychological Science*, *16*(11), 866.

Lindley, D. (1956). On a measure of the information provided by an experiment. *The Annals of Mathematical Statistics*, *27*(4), 986–1005.

Luce, R. D. (1959). *Individual choice behavior a theoretical analysis*. Oxford, England: Wiley and Sons.

Mackay, D. (1992). Information-based objective functions for active data selection. *Neural Computation*, *4*, 590–604.

Markant, D., & Gureckis, T. M. (2012). Does the utility of information influence sampling behavior? In N. Miyake, D. Peebles, & R. Cooper (Eds.), *Proceedings of the 34th Annual Conference of the Cognitive Science Society* (pp. 719–724). Austin, TX: Cognitive Science Society.

Markant, D., & Gureckis, T. M. (2014). Is it better to select or to receive? Learning via active and passive hypothesis testing. *Journal of Experimental Psychology: General*, *143*(1), 94–122. doi: 10.1037/a0032108.

McKenzie, C. (2006). Increased sensitivity to differentially diagnostic answers using familiar materials: Implications for confirmation bias. *Memory and Cognition*, *34*(3), 577.

Najemnik, J., & Geisler, W. (2005). Optimal eye movement strategies in visual search. *Nature*, *434*(7031), 387–391.

Navarro, D., & Perfors, A. (2011). Hypothesis generation, sparse categories, and the positive test strategy. *Psychological Review*, *118*(1), 120.

Nelson, J. (2005). Finding useful questions: On Bayesian diagnosticity, probability, impact, and information gain. *Psychological Review*, *114*(3), 677.

Nelson, J., & Movellan, J. (2001). Active inference in concept learning. *Advances in Neural Information Processing Systems*, *13*, 45–51.

Nelson, J., McKenzie, C., Cottrell, G., & Sejnowski, T. (2010). Experience Matters: Information acquisition optimizes probability gain. *Psychological Science*, *21*(7), 960.

von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton, NJ: Princeton University Press.

Newstead, S., Thompson, V., & Handley, S. (2002). Generating alternatives: A key component in human reasoning? *Memory and Cognition*, *30*(1), 129–137.

Nickerson, R. (1998). Confirmation bias: A ubiquitous phenomenon in many guises. *Review of General Psychology*, *2*, 175–220.

Oaksford, M., & Chater, N. (1994). A rational analysis of the selection task as optimal data selection. *Psychological Review*, *101*(4), 608–630.

Olsson, F. (2009). A literature survey of active machine learning in the context of natural language processing (Tech. Rep. No. T2009:06). Kista, Sweden: Swedish Institute of Computer Science.

Rottman, B., & Keil, F. (2012). Causal structure learning over time: Observations and interventions. *Cognitive Psychology*, *64*(1), 93–125.

Rubens, N., Kaplan, D., & Sugiyama, M. (2011). Active learning in recommender systems. In Ricci, F., Rokach, L., Shapira, B and Kantor, P. B. (Eds.), *Recommender systems handbook* (pp. 735–767). New York: Springer.

Sanbonmatsu, D. M., Posavac, S. S., Kardes, F. R., & Mantel, S. P. (1998). Selective hypothesis testing. *Psychonomic Bulletin and Review*, *5*(2), 197–220.

Sanborn, A., Griffiths, T., & Navarro, D. (2010). Rational approximations to rational models: Alternative algorithms for category learning. *Psychological Review*, *117*(4), 1144–1167.

Settles, B. (2012). Active learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, *6*(1), 1–114.

Settles, B., & Craven, M. (2008). An analysis of active learning strategies for sequence labeling tasks. In Lapata, M. and Ng, H. T. (Eds.), *Proceedings of the conference on empirical methods in natural language processing* (pp. 1070–1079). Stroudsburg, PA: Association for Computational Linguistics.

Sprenger, A., Dougherty, M., Atkins, S., Franco-Watkins, A., Thomas, R., Lange, N., & Abbs, B. (2011). Implications of cognitive load for hypothesis generation and probability judgment. *Frontiers in Psychology*, *2*, 129.

Steyvers, M., Tenenbaum, J., Wagenmakers, E., & Blum, B. (2003). Inferring causal networks from observations and interventions. *Cognitive Science*, *27*(3), 453–489.

Sutton, R., & Barto, A. (1998). *Introduction to reinforcement learning*. Cambridge, MA: MIT Press.

Thomas, R., Dougherty, M., Sprenger, A., & Harbison, J. (2008). Diagnostic hypothesis generation and human judgment. *Psychological Review*, *115*(1), 155–185.

Trope, Y., & Mackie, D. (1987). Sensitivity to alternatives in social hypothesis-testing. *Journal of Experimental Social Psychology*, *23*(6), 445–459.

Wason, P. (1960). On the failure to eliminate hypotheses in a conceptual task. *The Quarterly Journal of Experimental Psychology*, *12*(3), 129–140.

Wason, P. (1968). Reasoning about a rule. *The Quarterly Journal of Experimental Psychology*, *20*(3), 273–281.