Kiyohiko Nakamura*

Information seeking criteria: artificial intelligence, economics, psychology, and neuroscience

<https://doi.org/10.1515/revneuro-2020-0137>

Received November 26, 2020; accepted March 19, 2021;

published online April 14, 2021

Abstract: There has been an enormous amount of interest in how the brain seeks information. The study of this issue is a rapidly growing field in neuroscience. Information seeking is to make informative choices among multiple alternatives. A central issue in information seeking is how the value of information is assessed in order to choose informative alternatives. This issue has been studied in psychology, economics, and artificial intelligence. The present review is focused on information assessment and summarizes the psychological and computational criteria with which humans and computers assess information. Based on the summary, neurophysiological findings are discussed. In addition, a computational view of the relationships between these criteria is presented.

Keywords: active learning; curiosity; decision-making; information assessment; reward; uncertainty.

Introduction

Faced with unknown environments, humans attempt to gather information as effectively as possible. For example, medical doctors attempt to make diagnoses using as few observations and tests as possible. To achieve such effective diagnoses, doctors have to assess the degree to which possible observations and tests are informative. How do they assess their informativeness? This question is a central issue in diagnoses and other information seeking behavior. This issue was initially addressed in research on subjective judgments and weighting evidence (Good 1950) and comparison of scientific experiments (Lindley 1956). Subsequently, the issue has been studied in various research fields, such as psychology, economics, artificial intelligence, and neuroscience. The present article is focused on describing relationships between criteria for information

assessment in these fields and implications from the viewpoints of different fields, rather than a synthesis of each criterion. First, an overview of studies in artificial intelligence is presented that provides a mathematical formulation of information seeking behavior and computational findings about this behavior. Second, research on how humans assess information is reviewed in economics and psychology. Finally, recent advances in neuroscience related to information seeking are summarized and a possible neural mechanism for information assessment is presented.

A subfield of artificial intelligence, i.e., active learning, addresses a learning framework in which computers aim to learn user interest by making as few queries as possible (Fu et al. 2013; Settles 2010). Consider the task of learning which web pages a user finds interesting. The computer asks the user to label media files (e.g., documents and images) as “interesting” or “not interesting”. The computer attempts to predict labels of unlabeled webpages using as few labeled files as possible. A key issue in the learning is how to select the webpages to be labeled in order to achieve high-accuracy prediction. Different selection criteria have been proposed and applied to real-world datasets. These criteria, the results of the applications, and their computational implications are summarized herein.

Human information seeking behavior has been studied in psychology (Coenen et al. 2019; Loewenstein 1994). Such behavior can be divided into two classes: instrumental and non instrumental information seeking. Instrumental information seeking is performed in obtaining extrinsic rewards, i.e., immediate benefits. Medical observations and tests are included in this class because they are useful for diagnoses. Several criteria for human information assessment have been proposed and experimentally tested (Nelson 2005). Non instrumental information seeking behavior is evoked intrinsically by curiosity. Curiosity is measured by subjective rating (Marvin and Shohamy 2016) and the time that subjects are willing to wait to obtain the information (Kang et al. 2009). This article mainly addresses instrumental information seeking in relation to the computational results in active learning. In addition, an overview of research on non instrumental information seeking is presented.

Study of information seeking is rapidly growing in neuroscience (Gottlieb and Oudeyer 2018; Kidd and Hayden

*Corresponding author: Kiyohiko Nakamura, School of Computing, Tokyo Institute of Technology, Tokyo, 152-8550, Japan, E-mail: nakamurakiyohiko@gmail.com

2015). Previous and recent studies have shown that different brain regions are related to information seeking, including the premotor (Nakamura 2006), lateral prefrontal (Nakamura and Komatsu 2019), orbital (Blanchard et al. 2015), and parietal cortices (Horan et al. 2019), and the midbrain (Bromberg-Martin and Hikosaka 2009). The criteria for information assessment by the nervous system have been examined in the lateral prefrontal cortex (Nakamura and Komatsu 2019). These studies are reviewed and a possible neural mechanism for information assessment is presented based on computational implications in artificial intelligence.

Information seeking in artificial intelligence

In active learning, the learner aims to find a rule that separates a target subset C from the entire set X of all “instances” x_i ($i = 1, \dots, n$). The learner tries to achieve high accuracy of the rule using as few labeled instances as possible, thereby minimizing the cost of obtaining labeled data. In the aforementioned example of learning for webpage searching, instances x_i are media files, and the target subset C is the set of files that are interesting to the user. The computer selects an “informative” instance and asks the user whether the instance is interesting to the user. According to the response of the user, the computer updates its *belief*, i.e., estimates of the probabilities $p(x_i|D_{k-1})$ that x_i is included in C given the condition D_{k-1} into $p(x_i|D_k)$, where D_k ($k = 0, 1, \dots$) is the set of k previous responses. This question-and-answer is repeated until the learning reaches some threshold of accuracy. This learning is composed of two operations: selecting query instances and estimating the conditional probabilities $p(x_i|D_k)$. The latter operation is required for a wide variety of probabilistic inferences and active learning. The present article is focused on a review of studies on the criteria for selecting queries.

A simple criterion is to choose the instance x_i^* with the smallest margin (Campbell et al. 2000; Fu et al. 2013; Settles 2010). Consider classification of multiple categories C_j ($j = 1, \dots, m$). The instance x_i^* is given by

$$x_i^* = \operatorname{argmin}_{x_i} [P(C_1^+|x_i) - P(C_2^+|x_i)], \quad (1)$$

where $P(C_j^+|x_i)$ is the probability of category C_j given x_i , and C_1^+ and C_2^+ are the first and second most probable category labels of x_i under the current classifier, respectively. This criterion searches the instances x_i for which the difference between the largest probabilities of two categories is the smallest, and therefore queries of these instances are expected to improve the resolution of the classifier. For a binary classification task, this criterion chooses the instance x_i^*

having a $P(C_j^+|x_i^*)$ closest to 0.5, i.e., the largest uncertainty, because the margin is given as the difference between $P(C_j^+|x_i^*)$ and $1 - P(C_j^+|x_i^*)$. The intuitive justification of this criterion is that the most uncertain instances are closest to the classification boundary and that labeling uncertain instances is expected to produce a good classifier. Figure 1A shows an illustrative example of active learning, where instances distribute in the two-dimensional space. The instances to be labeled (filled circles) were chosen around the boundary between the three categories (magenta, cyan, and blue circles), whereas instances chosen by random sampling distributed uniformly (Figure 1B). The classifier produced by logistic regression effectively separates the three categories.

Figure 1: Illustrative examples of active learning by uncertainty and random sampling. Circles in the two-dimensional space represent instances x_i . Instances x_i of three categories C_j ($j = 1, 2, 3$) and probabilities $P(C_j|x_i)$ were generated from Gaussian distributions that were centered at $\mu_j = (-2, 0)$, $(2, 0)$, and $(0, 2\sqrt{3})$ with standard deviation $\sigma = 1$. Initial values of probabilities $P(C_j|x_i)$ were given as $f_j(x_i)/\sum_j f_j(x_i)$, where $f_j(x_i)$ were probability density functions of Gaussian distributions with means μ_j and $\sigma = 1$. In active learning, values $P(C_j|x_i)$ were updated only for labeled instances. The rules (classifiers) that separated the three categories were produced using logistic regression of labeled instances. The three categories are indicated by colors; magenta, cyan, and blue. The instances labeled by active learning are indicated by filled circles. The lines indicate the classifiers. The two colors of each line indicate the two categories separated by the line. (A) Classification by minimum margin sampling. The instances sampled by the minimum margin criterion are located close to the classification boundary. The misclassified instances are indicated by crosses. (B) Classification by random sampling from the same distributions of instances as in (A). The labeled instances distribute uniformly.

The numbers of the misclassified instances are 11 for the smallest margin and 14 for random sampling.

Another criterion using a statistic value is minimizing the expected entropy over the instances $-\sum_{x_i} P(x_i) \sum_{C_j} P(C_j|x_i) \log P(C_j|x_i)$. Several studies have shown that the entropy-based criteria performed well for real-world problems. Guo and Greiner (2007) conducted a set of experiments on 17 UCI datasets (e.g., Breast Cancer Wisconsin [Diagnostic] Data Set) and showed that their approach using the entropy-based criterion worked effectively. Krause and Guestrin (2009) presented an algorithm for scheduling observation selection in multisensor networks and demonstrated that their algorithm using the entropy-based criterion improved selection of sensor observations for energy conservation in buildings.

Variance reduction has also been proposed for choice of queries (Saar-Tsechansky and Provost 2001), which selects instance x_i with the minimum variance $\sum_{C_j} (P(C_j|x_i) - \sum_{C_j} P(C_j|x_i)/m)^2/m$. Schein and Ungar (2007) evaluated several criteria by applying active learning algorithms with these criteria for real-world datasets, including OptDigits (handwritten digits from 43 people) and TIMIT (speech of American English speakers of different sexes and dialects), and found that sampling with variance reduction performed best and minimizing entropy performed worst.

Numerous studies have presented different criteria, including expectation maximization (McCallum and Nigam 1998), clustering (Nguyen and Smeulders 2004), and other criteria (Gu et al. 2014; Tong and Koller 1998), and showed that these criteria performed well for different datasets. As mentioned before, active learning is composed of two operations: selecting query instances and estimating conditional probabilities. For the latter, the reviewed studies used different methods that were fit for the datasets (e.g., a chain graphical model in Krause and Guestrin 2009; the Bayesian net in Nguyen and Smeulders 2004). The obtained results suggest that the performance of active learning depends on the methods for estimating conditional probabilities and the criteria for selecting queries and that the pair of the two must match the datasets in order to effectively learn the datasets. This conclusion is known as the “no free lunch” theorem in machine learning, i.e., there is no computational method that works well for any class of problems, but a certain method can match a certain class of problems to solve these problems (Wolpert and Macready 1996). The reviewed results indicate that the theorem also holds for active learning.

Although machine learning research indicates that addressing selection of informative queries without specifying the class of datasets to be learned is of no use, investigating how humans select informative queries or assess the value of information is worthwhile because their criterion is supposed to have been developed by evolution and therefore

should match real-world problems, which are supposed to compose a certain class. The next section presents a review of studies on human criteria for selecting queries and assessing information.

Human criteria for information assessment

Human criteria for selecting queries have been studied in economics and psychology.

Economic value of information

In economics, decision theory provided a formula for assessing the value of information (Hubbard 2010). This formula gives the value of information as the price that one would be willing to pay in order to gain access to the information. Suppose that one searches n mailboxes for a letter and the cost of opening each mailbox is V . There are n cases in each of which the letter will be found by opening i mailboxes ($i = 1, \dots, n$). Each case occurs with probability $1/n$ (Appendix A). Therefore, the number of mailboxes expected to be opened until finding the letter is $(1 + 2 + \dots + n)/n = (n+1)/2$ and the expected cost is $(n+1)V/2$. If information indicating which mailbox stores the letter is offered, the cost is reduced to V . Consequently, if the price of the information is less than $(n-1)V/2$, one would pay for the information. In general, the economic value of information is given by the difference between the expected costs or opportunity losses before and after the information is offered.

Psychological criteria for information assessment

Psychological research has proposed several human criteria for assessing information. Skov and Sherman (1986) and Slowiaczek et al. (1992) conducted an experiment in which subjects performed a hypothetical scientific inference of identifying a novel creature on a faraway planet. The subjects were told the distribution of the two types of creatures on the planet and the distribution of various features of the creatures. Then, they were asked to choose the feature they would ask about to identify the novel creature. The result showed that the diagnosticity (Edwards 1968; Good 1950, 1975, 1983) predicted the ordering of features f_i that subjects would choose to ask about. The diagnosticity of feature f_i is given as

$$\sum_h P(a_{ih}) \max(P(a_{ih}|C_1)/P(a_{ih}|C_0), P(a_{ih}|C_0)/P(a_{ih}|C_1)), \quad (2)$$

where C_0 and C_1 are two categories of the creatures, a_{ih} ($h = 1, \dots, l$) is the h -th value of the feature f_i , $P(a_{ih})$ is the probability that the creatures have feature value a_{ih} and $P(a_{ih}|C_j)$ is the conditional probability of feature value a_{ih} given category of the creature C_j ($j = 0, 1$). This criterion implies that people prefer evidence of f_i that is most differentially probable under hypothesis C_j and alternative C_j' .

Another criterion is the probability gain, which measures the degree to which medical tests f_i improves the expected probability of making a correct guess (Baron 1985):

$$(\sum_h P(a_{ih}) \max_{C_j} P(C_j|a_{ih})) - \max_{C_j} P(C_j) \quad (3)$$

Baron et al. (1988) asked subjects to indicate prior probabilities $P(C_j)$ and $P(a_{ih})$ of three diseases C_j ($j = 1, 2, 3$) and results a_{ih} of medical tests f_i , respectively, and conditional probabilities $P(C_j|a_{ih})$ of disease C_j given test result a_{ih} . The subjects were then asked to rate the usefulness of the medical tests f_i for diseases C_j . The obtained data showed that subjects' ratings of tests were consistent with the values of the probability gain.

The expected reduction in uncertainty of beliefs is also widely used for information assessment, which is the information gain quantified using Shannon entropy (Lindley 1956):

$$-\sum_{C_j} P(C_j) \log P(C_j) - (-\sum_h P(a_{ih}) \sum_{C_j} P(C_j|a_{ih}) \log P(C_j|a_{ih})). \quad (4)$$

Oaksford and Chater (1994) and Oaksford et al. (1997) examined the results of the abstract selection task (Wason 1968) in which subjects were shown the top faces of four cards, showing A, T, 4, and 9 (Figure 2), and were asked which cards needed to be flipped to falsify the rule that "If there is a vowel on one side, then there is an even number on the other side." They showed that the ordering of the information gain of each card matched the frequencies of card selection under the assumption that subjects considered As and 4s to be rare. Benish (1999, 2003) used the information gain to quantify diagnostic test performance and showed that the information gain has some advantages over receiver operating characteristic (ROC) curves.

Figure 2: Four cards for the experiment by Oaksford and Chater (1994) and Oaksford et al. (1997).

The absolute change in beliefs is another criterion for assessing information (Wells and Lindsay 1980), which is given by

$$\sum_h P(a_{ih}) \sum_{C_j} |P(C_j|a_{ih}) - P(C_j)|/m. \quad (5)$$

Klayman and Ha (1987) presented an analysis of cognitive biases in hypothesis testing. When subjects are provided with one triple (2, 4, 6) and guess that the rule is "three consecutive even numbers," they prefer testing whether a triple (6, 8, 10) fits the rule to testing whether (2, 4, 7) does not fit the rule (Wason 1960). The preference agreed with the value of the absolute change in beliefs that was larger for the former confirmation test than for the latter disconfirmation test under some common conditions, e.g., three consecutive even numbers were rare in all three consecutive numbers.

As reviewed before, different psychological criteria have been supported by different experimental data. Nelson (2005) reanalyzed these data and examined which of these criteria accounted for each experimental datum. The obtained results are as follows. First, the experimental results of Skov and Sherman (1986) and Slowiaczek et al. (1992) were not only predicted by the diagnosticity, but also by the probability gain, the information gain, and the absolute change in beliefs. Second, the data of Baron et al. (1988) are consistent with the information gain and the absolute change in beliefs, and the probability gain. Third, the data of Oaksford and Chater (1994, 2003) were predicted by the probability gain and the absolute change in beliefs, but not by the diagnosticity. These findings suggest that human subjects assess information primarily according to the probability gain, the information gain, or the absolute change in beliefs, but not the diagnosticity.

Based on this analysis, Nelson et al. (2010) examined which of the probability gain, the information gain, or the absolute change in beliefs best describes human information search. His experiments involved classifying simulated plankton specimens as species C_1 or C_2 , where the species was a probabilistic function of two two-valued features, colors of eyes f_1 (yellow or black) and claws f_2 (dark or light green). Subjects first learned the values of probabilities $P(C_j)$ and $P(f_i|C_j)$ by classifying the plankton specimens and receiving immediate feedback. The probabilities were designed so as to maximize disagreement about which feature was more useful for categorization between the criteria. After achieving high mastery of these probabilities, the subjects were asked which of the two features was considered to be most useful for categorization. The results indicated that the probability gain explained the subjects' choices better than the other criteria.

Crupi et al. (2018) showed that several psychological criteria, including the probability gain and the information gain, arose as special cases in a unified formalism, the Sharma–Mittal framework:

$$\left(1 - \left(\sum_j P(C_j)^r\right)^{(t-1)/(r-1)}\right) / (t-1), \quad (6)$$

where r and t are parameters. The framework is reduced to $1 - \max_{C_j} P(C_j)$ in the limit for $r \rightarrow \infty$ when $t = 2$, and to $-\sum_{C_j} P(C_j) \log P(C_j)$ for $r = t = 1$. Therefore, increments in their expected values by obtaining information correspond to the probability gain and the information gain, respectively. Since $\max_{C_j} P(C_j)$ gives the best guess of C_j , $1 - \max_{C_j} P(C_j)$ quantifies error of the guess. The larger the error, the greater the “surprise” expected after obtaining information. Surprise has also been defined as $-\log P(C_j)$ (Shannon 1948; Strange et al. 2005). These considerations by Crupi et al. (2018) illustrate that the psychological criteria can be viewed as the expected surprise, in which the surprise and the expectation were defined independently.

Advantages and limitations of human criteria for information seeking

The economic value of information is optimal in the sense that it maximizes the expected rewards or minimizes the expected costs for reaching goals. However, to compute the economic value of information, it is necessary to know all future rewards, costs, and their probabilities. This makes the computation generally combinatorial and intractable from the viewpoint of computational complexity. In the aforementioned example of searching mailboxes for a letter, probabilities of all future states were initially predictable (Appendix A) and therefore the economic value of information could be computed. In contrast, the psychological criteria only require probability values of rewards/costs in the current and next environmental states. This allows humans to assess information in most situations where complete knowledge of future rewards/costs is not available. Instead, psychological criteria work well only for a certain class of problems, which may consist of real-world problems.

Absolute change in beliefs as surprise

The absolute change in beliefs is also viewed as another index of surprise as follows. In the example of medical diagnosis, the absolute change in beliefs is maximized in the case in which the disease of the patient is believed

with certainty to be one disease before test i , and the obtained result a_{ih} of the test indicates that the disease of the patient is definitely a different disease (for proof, see Appendix B). This case implies that the initial belief is entirely changed by the result of the test, expecting that surprise should be evoked.

Non-instrumental information seeking: curiosity

Curiosity is evoked by an intrinsic desire for information and knowledge. Human intrinsic states are difficult to examine. Curiosity is investigated using questionnaires (Gruber et al. 2014) and measuring the willingness of subjects to wait for an outcome. The willingness to wait has been shown to be a measure of the motivational value (Frederick et al. 2002). Kang et al. (2009) asked subjects to read trivia questions, such as “What is the name of the galaxy that Earth is a part of?” and to guess the answers. The subjects were then asked to indicate their curiosity about the correct answers. After guessing, subjects had to wait until the answers appeared. The wait time varied from trial to trial, and subjects could quit waiting and skip to the next question at any time. Their results showed that the reported curiosity reached its maximum when the level of confidence was moderate and that subjects waited longer for answers that they were more curious about, suggesting that curiosity increases if people are more uncertain about their guesses and makes people patient to wait for the correct answers.

Van Lieshout et al. (2018) have also shown that curiosity increases with uncertainty and willingness to wait. They designed a lottery task in which subjects were presented with an image of a vase containing 20 marbles (Figure 3). Each marble could be either red or blue. Both colors were associated with points subjects could earn. The subjects indicated the rate of their curiosity regarding the outcome of a lottery in Experiment 1 (Figure 3A) and whether they wanted to wait to see the outcome, which was presented after an additional jittered delay in Experiment 2 (Figure 3B). The obtained data showed that the rating and the willingness to wait increased with the outcome uncertainty (OU), which is the product of Shannon entropy and the absolute difference between the points of red and blue marbles:

$$OU = \left(-\sum_{i=1,2} p_i \log_2 p_i\right) |z_1 - z_2|, \quad (7)$$

where p_1 and p_2 are the probabilities of drawing red and blue marbles, respectively, and z_1 and z_2 indicate the points associated with red and blue marbles, respectively.

Neural substrates of information seeking

Functional magnetic resonance imaging (fMRI) studies showed that different brain areas were involved in information seeking. The level of curiosity was correlated with activity in caudate regions that were previously suggested to be involved in anticipated reward (Kang et al. 2009). The midbrain and the nucleus accumbens were activated during states of high curiosity (Gruber et al. 2014). The left inferior parietal lobe activity increased with increasing rate of curiosity, while activity in a large network of regions decreased with increasing curiosity, including the hippocampus, precuneus, and some clusters within the temporal and frontal lobes (van Lieshout et al. 2018). The left lateral ventral striatum was more activated by a visual cue that predicted high information than another cue that predicted low information (Filimon et al. 2020). An

electroencephalogram (EEG) study showed that subjects wanted to sample more evidence before making a final decision if their confidence in the decision was low and that predecisional and postdecisional event-related potential components were similarly modulated by the level of confidence and by information seeking choices (Desender et al. 2019).

Horan et al. (2019) showed that parietal cortical neurons encoded reduction in uncertainty, which was measured with the information gain. Monkeys were trained with an oculomotor task in which the animals made two consecutive saccades to obtain a reward. All trials appeared in INF or unINF blocks. Each trial started with fixation of a central point. The color and shape of the point informed the monkeys about the reward size and the block type. After a short delay, two targets and a cue appeared. After an additional central fixation, the fixation point disappeared, instructing the monkey to make the first saccade to the cue. Fixation of the cue initiated the cue's move. In INF blocks, the monkey received a reward if it made the second saccade to the target that had been cued by the cue's motion. In contrast, in unINF blocks, a single target was correct regardless of the cue's motion. Therefore, the first saccade provided information about the correct target in INF blocks, but not in unINF blocks. During the fixation preceding the first saccade, parietal neurons delivered greater activity in INF blocks than in unINF blocks, suggesting that the neurons might encode the information gain. To examine whether the neural activity could be explained by the economic value of information (VOI), another type of trial was introduced to both blocks, in which no cue's motion was initiated by the first saccade. For these trials, the VOI is 0 in unINF blocks, while the VOI is positive and increases with reward size in INF blocks. These changes in the VOI were not reflected by the recorded neural activity.

The neural activity recorded by Horan et al. (2019) reflected the reduction in uncertainty, and therefore the activity could be explained with the probability gain and the information gain. Nakamura and Komatsu (2019) examined which of the probability gain, the information gain, or the VOI might be used for information seeking in the nervous system. They also trained monkeys with oculomotor tasks in which the three criteria predicted different information values carried by saccades (Figure 4A). The monkeys fixated on a central cross, and six white dots were subsequently illuminated around the cross. One dot was randomly assigned as a reward target. The monkeys searched by eye movements. The six dots also included an "informative" target, which was always the lower right dot. The choice of the informative target revealed the reward target in task A, whereas two dots remained as reward target candidates in task B. In task C, the reward target was always the lower right dot. In task D, the

Figure 3: (A) Schematic diagram of Experiment 1 (van Lieshout et al. 2018). Subjects saw a screen on which a vase with 20 marbles was depicted, either of which could be red or blue, and the points associated with these marbles. Subjects were told that one of the marbles would be selected for them and that they would be awarded with the points associated with this marble. Next, subjects indicated how curious they were about seeing the outcome of the vase (1–4). There was a 50% chance of seeing the outcome, regardless of the curiosity response of the subject. Importantly, a marble was selected in every trial, and subjects were awarded with the points associated with this marble, regardless of seeing the outcome of a trial. (B) Schematic diagram of Experiment 2. The task was similar to Experiment 1, except that the subjects indicated whether they wanted to see the outcome of a trial. If the subject responded “Yes,” then he/she had to wait an additional 3–6 s before the outcome was presented to them. If the subject responded “No,” then the outcome was not presented to him/her. Still, a marble was selected in every trial and the subjects were awarded with the points associated with this marble, regardless of seeing the outcome of a trial.

A

B

	Descending order of information value			
VOI	A	B	C	D
Information gain	A	B	D	C
Probability gain	A	D	B	C

Figure 4: Behavioral tasks and neural population responses (Nakamura and Komatsu 2019). (A) Examples of task performance. Yellow arrows indicate the eye movements of the monkey. Task A: In the upper row, the monkey first chose the top dot, which was not the informative target. The monkey then chose one dot after another until choosing the reward target. In the lower row, the monkey first chose the lower right informative target, followed by the reward target. Task B: After the monkey first chose the informative target, two dots remained white. Task C: The lower right dot was always the reward target. Task D: After the first choice, two dots remained at fixed locations, while the other dots disappeared. (B) Task orders with descending information values. (C through E) Projections of the population average responses of monkey S onto the axes that account for the variance in the population response due to the variation in the information values of the value of information (VOI), the information gain, and the probability gain, respectively. The data for tasks A, B, C, and D are shown in magenta, cyan, blue, and green, respectively. The horizontal bars indicate the time intervals during which the population responses could correctly encode the information values of the corresponding criteria. The vertical lines indicate the time 700 ms after the onset of six dots. (F through H) Regression lines that were obtained when the center of a 400 ms sliding window was at 700 ms after the onset of six dots (indicated by the vertical lines in (C through E)). The two-sided Wilcoxon signed rank test revealed that the component scores differed significantly from the values predicted by the regression lines of the information gain or the probability gain in task D ((G and H), asterisks).

reward target was the upper right or lower right dot. Each task was administered in a trial block. The information value carried by the first saccade to the informative target decreases in the order of tasks A, B, and C, regardless of which information measure is used to assess information (Figure 4B). However, the value in task D lies in different positions depending on the used criteria, i.e., the value in task D is the same as in task C if the VOI is used, is between the values in tasks B and C if the information gain is used, and is between the values in tasks A and B if the probability gain is used. They showed that neuronal population responses of the lateral prefrontal cortex were negatively correlated with the values of the VOI and the probability gain (Figure 4C and E), and that the duration of the correlation was longer for the VOI than for the probability gain.

Their results suggest that the lateral prefrontal cortex is involved in assessing information and that the assessment is primarily performed according to the VOI. As mentioned above, the VOI is optimal to minimize the expected costs, but the VOI needs knowledge of all future rewards, costs, and their probabilities for its computation. In this experiment, the monkeys were well trained to perform the four tasks before unit recording. Therefore, they could learn estimates of all future rewards and their probabilities and compute the VOI. Furthermore, they also computed the probability gain. Animals are often required to make decisions based on limited knowledge about the future, such as the current and next states of the environment. Computing the probability gain requires only the values of the reward probability for the current and next choices. Although their result suggests that the VOI might be primarily used for the tasks in their experiment, the nervous system may generally compute the probability gain to assess information in a wide range of circumstances.

Switch between the criteria and neural mechanism for computation of the probability gain

Neurophysiological and psychological studies (Nakamura and Komatsu 2019; Nelson 2010) have suggested that the probability gain might be used for information seeking. However, human behavior that is not predicted by the probability gain has been reported, which is choice by “split-half heuristic.” Suppose that one searches for a target item from a set of equally probable items and that he/she chooses its subset and asks whether the subset includes the target item. In this case, people prefer choosing a subset that includes nearly half of the items in order to exclude half of the

candidate items. In this task, any subset has the same value of the probability gain, whereas the information gain is maximum for a subset of half of the items, suggesting that the information gain may be used in the heuristic (Crupi et al. 2018).

An explanation of these inconsistent results is that the subjects and the animals learned the probabilities of stimuli and outcomes by iterative trials in the experiments of Nelson (2010) and Nakamura and Komatsu (2019), whereas, by initial instruction, the subjects know that all items are equally probable in the split-half heuristic task. Learning by iterative trials is unconsciously performed, and therefore computation of the probability gain may more directly stem from neural mechanisms than computation of the information gain. Computation of the probability gain consists of summation and maximization (Eq. 3). Neural circuitry can implement the two operations in convergent projection and lateral inhibition respectively, because lateral inhibition performs a winner-take-all competition to find the maximum (Coultrip et al. 1992). The two circuits are common in the neural circuitry. In contrast, the information gain needs logarithm operation to be computed (Eq. 4). This operation may require rather more complicated neural circuits than summation and maximization.

In some experimental conditions, human information seeking using learned probabilities is better explained by the information gain than by the probability gain (Meder and Nelson 2012). Artificial intelligence research (Schein and Ungar 2007) has also shown that the performance of active learning differs between criteria that reduce uncertainty, but these criteria perform roughly equally well in comparison to each other. These findings suggest that switching between the two criteria does not strictly depend on the manner of acquiring probability values, but may fluctuate.

Integration of information and primary reward

Information functions as a reward, just as primary rewards. Humans and animals are often required to process such different types of rewards in parallel to make decisions. Blanchard et al. (2015) showed that monkeys sacrificed primary rewards (water) to view advance information about the outcome of choices. In their experiment, monkeys chose between two gambles that yielded either a water reward or no reward with equal probability. The water amount of each gamble was randomly set and indicated by the height of a bar on a computer monitor (Figure 5). The contour of the two bars representing two gambles were colored by cyan or

Figure 5: Task design and recording location (Blanchard et al. 2015). (A) Basic task design. Two offers were presented in sequence, followed by a blank period. The monkey then had to fixate a central target. The two options then reappeared and the monkey chose one with a gaze shift. Then, a cue appeared, which was either informative (indicating whether the trial would be rewarded) or uninformative (leaving the monkey in a state of uncertainty). Following a 2.25-s delay, the monkey obtained the outcome. Cyan and magenta bars indicated informative and uninformative options, respectively. An inscribed white rectangle indicated gamble stakes. An inscribed red or green circle was the cue. (B) Magnetic resonance image indicating the position of 13 m.

magenta, which indicated that the gambling outcome was presented 2.25 s before the time of reward delivery or that there was no advance notice, respectively. Although monkeys preferred both greater water amounts and informative cues, they were indifferent when the difference in the water amount between informative and uninformative gambles was a substantial fraction of the water they were offered (33 or 22%). This indicates that the monkeys integrated the advance information value and water reward value into a single dimension of subjective value. Furthermore, they reported that neurons in the orbitofrontal cortex (OFC) encoded informativeness and water amount in orthogonal manners during the gambling tasks, suggesting that the values of information and water reward were not integrated in the OFC and that the OFC might be a relatively early stage of reward integration.

Midbrain dopamine neurons and lateral habenula neurons encode errors in the prediction of both information and primary reward (Bromberg-Martin and Hikosaka 2009, 2011). In their experiments, monkeys chose between two targets that provided different information about upcoming water rewards. The neuronal responses encoded errors in the prediction of the expected information at target onset and then replicated the same responses to cues that revealed the reward amount and to reward delivery, as reported in previous studies (reviewed by Schultz 2007). This indicates that the information value and the water reward value were integrated into a single dimension of the responses of single neurons.

Conclusion

In this article, studies on information seeking and information assessment were highlighted and the following was shown:

- (1) Artificial intelligence research provided a mathematical formulation of information seeking and showed that effective criteria for information seeking vary with the datasets to be learned.
- (2) Economics defines the value of information as the price that one would be willing to pay in order to gain access to information, and the price is given by the difference between the expected costs or opportunity losses before and after the information is offered.
- (3) Psychological studies suggested that people might use criteria that reduced uncertainty in beliefs, including the probability gain and the information gain. The former may be mainly used for information seeking based on implicitly learned probabilities.
- (4) The economic value of information is optimal in the sense that it minimizes the expected costs but its computation is intractable in general, because it requires complete knowledge concerning all future benefits and costs. In contrast, psychological criteria require only limited knowledge about the future, such as the current and next states of the environment. This makes psychological criteria effective for a limited class of problems that may consist of real-world datasets or problems.

- (5) Neurophysiological studies showed that activity of cortical neurons encoded the probability gain during information seeking, whereas the activity could reflect the economic value of information depending on the information seeking tasks. The probability gain is computed with summation and maximization, which may be implemented in convergent projection and lateral inhibition of neural circuitry, respectively.
- (6) Finally, present findings suggest that although humans and animals use different criteria depending on the information seeking tasks, they mostly use the probability gain in cases in which probabilities used in the tasks are implicitly learned by iterative trials. This suggests that the probability gain may be the most plausible measure of information value in animal research.

Information seeking behavior still has a number of issues to be addressed and requires much future research. Artificial intelligence research shows that information seeking behavior consists of two procedures: selecting queries and updating beliefs according to responses to the queries. The present article focused on the criteria for query selection. Belief update is performed by computing conditional probabilities $p(x_i|D_k)$ given the previous k responses D_k , which serves to propagate the responses to neighboring instances x_i . Conditional probabilities are used to formulate a wide variety of cognitive behaviors and information seeking. The method used to compute the conditional probabilities depends on the cognitive behavior and is a fundamental issue in cognitive science and neuroscience. Further research on belief update and integrating its results with the indications concerning query selection criteria in the present article will provide a consistent understanding of information seeking behavior.

Author contributions: The author has accepted responsibility for the entire content of this submitted manuscript and approved submission.

Research funding: None declared.

Conflict of interest statement: The author declares no conflicts of interest regarding this article.

Appendix A

In the example of searching mailboxes, the probabilities of all future states were initially predictable as follows. The probability that the mailbox opened in the first choice includes the letter is $1/n$. The probability that the mailbox opened in the second choice includes the letter is the product of the probability that the mailbox opened in the first choice does not include the letter and the probability

that the mailbox opened in the second choice includes the letter, that is, $((n-1)/n)(1/(n-1)) = 1/n$. Similarly, the probability that the mailbox opened in the n -th choice includes the letter is $((n-1)/n)((n-2)/(n-1)) \cdots (1/2) = 1/n$.

Appendix B

Assuming that $P(C_j)$, $P(a_{ih})$, and $P(C_j|a_{ih})$ are mutually independent, the absolute change in beliefs is maximized under the following conditions:

- (1) $P(C_j) = 1$ for $j = \exists j' \in \{1, \dots, m\}$ or 0 for $j \neq j'$.
- (2) $P(C_j|a_{ih}) = 1$ for $j = \exists j_h^* (\neq j_h')$ or 0 for $j \neq j_h^*$.

In the example of medical diagnosis, the first condition implies that the disease of the patient is initially believed with certainty to be disease $C_{j'}$, and the second condition means that the obtained result a_{ih} of test i indicates that the disease of the patient is definitely a different disease C_{jh^*} .

Proof: In the absolute change in beliefs (Eq. 5), $|P(C_j|a_{ih}) - P(C_j)|$ is maximum at $P(C_j|a_{ih}) = 0$ if $P(C_j) > 0.5$ or at $P(C_j|a_{ih}) = 1$ otherwise, which is given by the second condition. Suppose $P(C_{j'}) > 0.5$. Then, $P(C_j) \leq 0.5$ for $j \neq j'$, and Eq. (5) is not greater than $\sum_h P(a_{ih}) (P(C_{j'}) + \sum_{j \neq j'} (1 - P(C_j))) / m = (P(C_{j'}) + m - 1 - \sum_{j \neq j'} P(C_j)) / m = (m - 2(1 - P(C_{j'}))) / m$. Suppose $P(C_j) \leq 0.5$ for $\forall j$. Then, Eq. (5) is reduced to $\sum_h P(a_{ih}) (\sum_j (1 - P(C_j))) / m = (m - 1) / m$, which is smaller than $(m - 2(1 - P(C_{j'}))) / m$ because $1 - P(C_{j'}) < 0.5$. This value $(m - 2(1 - P(C_{j'}))) / m$ is maximum if $P(C_{j'}) = 1$, which is given as the first condition. Consequently, the absolute change in beliefs is maximized under the two conditions.

References

- Baron, J. (1985). *Rationality and intelligence*. Cambridge, England: Cambridge University Press.
- Baron, J., Beattie, J., and Hershey, J.C. (1988). Heuristics and biases in diagnostic reasoning: II. Congruence, information, and certainty. *Organ. Behav. Hum. Decis. Process.* 42: 88–110.
- Benish, W.A. (1999). Relative entropy as a measure of diagnostic information. *Med. Decis. Making* 19: 202–206.
- Benish, W.A. (2003). Mutual information as an index of diagnostic test performance. *Methods Inf. Med.* 42: 260–264.
- Blanchard, T.C., Hayden, B.Y., and Bromberg-Martin, E.S. (2015). Orbitofrontal cortex uses distinct codes for different choice attributes in decisions motivated by curiosity. *Neuron* 85: 602–614.
- Bromberg-Martin, E.S. and Hikosaka, O. (2009). Midbrain dopamine neurons signal preference for advance information about upcoming rewards. *Neuron* 63: 119–126.
- Bromberg-Martin, E.S. and Hikosaka, O. (2011). Lateral habenula neurons signal errors in the prediction of reward information. *Nat. Neurosci.* 14: 1209–1216.
- Campbell, C., Cristianini, N., and Smola, A. (2000). Query learning with large margin classifiers. In: *Proc. ICML*, pp. 111–118 (CA).

- Coenen, A., Nelson, J.D., and Gureckis, T.M. (2019). Asking the right questions about the psychology of human inquiry: nine open challenges. *Psychon. Bull. Rev.* 26: 1548–1587.
- Coultrip, R., Granger, R., and Lynch, G. (1992). A cortical model of winner-take-all competition via lateral inhibition. *Neural Netw* 5: 47–54.
- Crupi, V., Nelson, J.D., Meder, B., Cevolani, G., and Tentori, K. (2018). Generalized information theory meets human cognition: introducing a unified framework to model uncertainty and information search. *Cognit. Sci.* 42: 1410–1456.
- Desender, K., Murphy, P., Boldt, A., Verguts, T., and Yeung, N. (2019). A postdecisional neural marker of confidence predicts information-seeking in decision-making. *J. Neurosci.* 39: 3309–3319.
- Edwards, W. (1968). Conservatism in human information processing. In: Kleinmuntz, B. (Ed.), *Formal representation of human judgment*. New York: Wiley, pp. 17–52.
- Filimon, F., Nelson, J.D., Sejnowski, T.J., Sereno, M.I., and Cottrell, G.W. (2020). The ventral striatum dissociates information expectation, reward anticipation, and reward receipt. *Proc. Natl. Acad. Sci. USA* 117: 15200–15208.
- Frederick, S., Loewenstein, G., and O'Donoghue, T. (2002). Time discounting and time preference: a critical review. *J. Econ. Lit.* 40: 351–401.
- Fu, Y., Zhu, X., and Li, B. (2013). A survey on instance selection for active learning. *Knowl. Inf. Syst.* 35: 249–283.
- Good, I.J. (1950). *Probability and the weighing of evidence*. New York: Griffin.
- Good, I.J. (1975). Explicativity, corroboration, and the relative odds of hypotheses. *Synthese* 30: 39–73.
- Good, I.J. (1983). *Good thinking*. Minneapolis: University of Minnesota.
- Gottlieb, J. and Oudeyer, P.Y. (2018). Toward a neuroscience of active sampling and curiosity. *Nat. Rev. Neurosci.* 19: 758–770.
- Gruber, M.J., Gelman, B.D., and Ranganath, C. (2014). States of curiosity modulate hippocampus-dependent learning via the dopaminergic circuit. *Neuron* 84: 486–496.
- Gu, Y., Jin, Z., and Chiu, S.C. (2014). Active learning with maximum density and minimum redundancy. In: *Proc. ICONIP*, pp. 103–110.
- Guo, Y. and Greiner, R. (2007). Optimistic active learning using mutual information. In: *Proc. IJCAI*, pp. 823–829.
- Horan, M., Daddaoua, N., and Gottlieb, J. (2019). Parietal neurons encode information sampling based on decision uncertainty. *Nat. Neurosci.* 22: 1327–1335.
- Hubbard, D.W. (2010). *How to measure anything: finding the value of intangibles in business*. Hoboken: John Wiley & Sons Inc.
- Kang, M.J., Hsu, M., Krajbich, I.M., Loewenstein, G., McClure, S.M., Wang, J.T., and Camerer, C.F. (2009). The wick in the candle of learning: epistemic curiosity activates reward circuitry and enhances memory. *Psychol. Sci.* 20: 963–973.
- Kidd, C. and Hayden, B.Y. (2015). The psychology and neuroscience of curiosity. *Neuron* 88: 449–460.
- Klayman, J. and Ha, Y.-W. (1987). Confirmation, disconfirmation, and information. *Psychol. Rev.* 94: 211–228.
- Krause, A. and Guestrin, C. (2009). Optimal value of information in graphical models. *J. Artif. Intell. Res.* 35: 557–591.
- Lindley, D.V. (1956). On a measure of the information provided by an experiment. *Ann. Math. Stat.* 27: 986–1005.
- Loewenstein, G. (1994). The psychology of curiosity: a review and reinterpretation. *Psychol. Bull.* 116: 75–98.
- Marvin, C.B. and Shohamy, D. (2016). Curiosity and reward: valence predicts choice and information prediction errors enhance learning. *J. Exp. Psychol. Gen.* 145: 266–272.
- McCallum, A.K. and Nigam, K. (1998). Employing EM and pool-based active learning for text classification. In: *Proc. ICML*, pp. 350–358 (Madison, Wisconsin).
- Meder, B. and Nelson, J.D. (2012). Information search with situation-specific reward functions. *Judgm. Decis. Mak.* 7: 119–148.
- Nakamura, K. (2006). Neural representation of information measure in the primate premotor cortex. *J. Neurophysiol.* 96: 478–485.
- Nakamura, K. and Komatsu, M. (2019). Information seeking mechanism of neural populations in the lateral prefrontal cortex. *Brain Res.* 1707: 79–89.
- Nelson, J.D. (2005). Finding useful questions: on Bayesian diagnosticity, probability, impact, and information gain. *Psychol. Rev.* 112: 979–999.
- Nelson, J.D., McKenzie, C.R.M., Cottrell, G.W., and Sejnowski, T.J. (2010). Experience matters: information acquisition optimizes probability gain. *Psychol. Sci.* 21: 960–969.
- Nguyen, H.T. and Smeulders, A. (2004). Active learning using pre-clustering. In: *Proc. ICML*, pp. 79–86 (Banff, Canada).
- Oaksford, M. and Chater, N. (1994). A rational analysis of the selection task as optimal data selection task. *Psychol. Rev.* 101: 608–631.
- Oaksford, M. and Chater, N. (2003). Optimal data selection: revision, review, and reevaluation. *Psychon. Bull. Rev.* 10: 289–318.
- Oaksford, M., Chater, N., Grainger, B., and Larkin, J. (1997). Optimal data selection in the reduced array selection task (RAST). *J. Exp. Psychol. Learn. Mem. Cogn.* 23: 441–458.
- Saar-Tsechansky, M. and Provost, F. (2001). Active learning for class probability estimation and ranking. In: *Proc. IJCAI*, pp. 911–920.
- Schein, A.L. and Ungar, L.H. (2007). Active learning for logistic regression: an evaluation. *Mach. Learn.* 68: 235–265.
- Schultz, W. (2007). Behavioral dopamine signals. *Trends Neurosci.* 30: 203–210.
- Settles, B. (2010). *Active learning literature survey*. Computer sciences technical report 1648. Univ. of Wisconsin-Madison.
- Shannon, C.E. (1948). A mathematical theory of communication. *BSTJ* 27: 379–423.
- Skov, R.B. and Sherman, S.J. (1986). Information-gathering processes: diagnosticity, hypothesis-confirmatory strategies, and perceived hypothesis confirmation. *J. Exp. Soc. Psychol.* 22: 93–121.
- Slowiczzek, L.M., Klayman, J., Sherman, S.J., and Skov, R.B. (1992). Information selection and use in hypothesis testing: what is a good question, and what is a good answer? *Mem. Cognit.* 20: 392–405.
- Strange, B.A., Duggins, A., Penny, W., Dolana, R.J., and Friston, K.J. (2005). Information theory, novelty and hippocampal responses: unpredicted or unpredictable? *Neural Netw* 18: 225–230.
- Tong, S. and Koller, D. (1998). Support vector machine active learning with applications to text classification. In: *Proc. ICML*, pp. 287–295 (Stanford, CA).
- van Lieshout, L.L.F., Vandenbrouke, A.R.E., Muller, N.C.J., Cools, R., and de Lange, F.P. (2018). Induction and relief of curiosity elicit parietal and frontal activity. *J. Neurosci.* 38: 2597–2588.
- Wason, P.C. (1960). On the failure to eliminate hypotheses in a conceptual task. *Q. J. Exp. Psychol.* 12: 129–140.
- Wason, P.C. (1968). Reasoning about a rule. *Q. J. Exp. Psychol.* 20: 273–281.
- Wells, G.L. and Lindsay, R.C.L. (1980). On estimating the diagnosticity of eyewitness nonidentifications. *Psychol. Bull.* 88: 776–784.
- Wolpert, D.H. and Macready, W.G. (1996). No free lunch theorems for optimization. *IEEE Trans. Evol. Comput.* 1: 67–82.