

Retrospective surprise: a computational component for active inference

Kentaro Katahira¹, Yoshihiko Kunisato²,
Tsukasa Okimura³, Yuichi Yamashita⁴

1. Department of Psychological and Cognitive Sciences, Nagoya University, Nagoya, Aichi, Japan
2. Department of Psychology, Senshu University, Kawasaki, Kanagawa, Japan
3. Department of Neuropsychiatry, Keio University School of Medicine, Shinanomachi, Shinjuku, Tokyo, Japan
4. Department of Information Medicine, National Institute of Neuroscience, National Center of Neurology and Psychiatry, Kodaira, Tokyo, Japan

Abstract

In the free energy principle (FEP) proposed by Friston, it is supposed that agents seek to minimize the “surprise”—the negative log (marginal) likelihood of observations (i.e., sensory stimuli)—given the agents’ current belief. This is achieved by minimizing the free energy, which provides an upper bound on the surprise. The FEP has been applied to action selection in a framework called “active inference,” where agents are supposed to select an action so that they minimize the “expected free energy” (EFE). While EFE can be decomposed to interpretable components such as epistemic value and extrinsic value, it is difficult to understand intuitively how EFE itself is directly related to “surprise” and what psychological construct is related to EFE itself (as a single quantity). To facilitate the discussion and interpretation of psychological processes underlying active inference, we introduce a computational component termed the “retrospective surprise,” which is the surprise of an observation after updating the belief given the observation itself. The predicted retrospective surprise (PRS) is mathematically derived from a special case of EFE and provides a lower bound on EFE. We illustrate the properties of

EFE and PRS using examples of inference for a binary hidden cause given a binary observation. We discuss how information-seeking behavior is accounted for by EFE and PRS. Our results highlight the role of prior distribution of future observation in EFE. By setting this prior distribution to be fixed irrespective of which action is selected, the epistemic value can exert an influence on EFE, leading to information-seeking behavior.

Keywords: Active inference; Surprise; Expected free energy; Retrospective surprise; Action selection

1 Introduction

The free energy principle (FEP) was proposed as a unifying theory for the computational principle of the brain function (Friston, Kilner, & Harrison, 2006; Friston, 2010; Bogacz, 2017; Buckley, Kim, McGregor, & Seth, 2017). FEP posits that agents seek to minimize *surprise*. The term surprise (or *surprisal*) is defined as the negative log (marginal) likelihood of an observation, given the agents' belief. This can be achieved by minimizing the variational free energy, which provides an upper bound on surprise. In the perceptual domain, the minimization can be performed by updating the current belief. FEP has also been applied to action selection within a framework called *active inference* (Friston, Daunizeau, & Kiebel, 2009; Friston, Mattout, & Kilner, 2011; Friston et al., 2015). In this framework, especially for selecting action in discrete time, the agents are supposed to select action such that they can minimize the expected free energy (EFE) (Friston et al., 2015).

While EFE has been related to some computational components (Friston et al., 2015), its interpretation as a psychological construct remains elusive. EFE can be decomposed into two components: the *epistemic value*, which is the value for obtaining information about the hidden state of the world, and the *extrinsic value* (Friston et al., 2015). These quantities are related to other computational constructs such as Bayesian surprise and KL (Kullback-Leibler; risk-sensitive) control (Friston et al., 2015). Nevertheless, how EFE *as a single quantity* corresponds to psychological or computational constructs is unclear. For example, it is not straightforward how “surprise,” which is a key quantity in FEP, is related to EFE. Given the notion that the FEP posits that agents

try to minimize surprise, how surprise is defined with regard to future observation is important for conceptual understanding of active inference within the framework of FEP. In addition, defining the free energy for future observation is not straightforward, as there may be different variants of free energy for future observation, *predicted free energy* (Schwöbel, Kiebel, & Marković, 2018) and *generalized free energy* (Parr & Friston, 2018), have been proposed and applied to active inference. These facts have made a comprehensive understanding of active inference based on EFE difficult, especially for those who are not familiar with related mathematical concepts.

The aim of the present study is to propose a natural quantity, for which EFE is related to (upper bound on) and which gives better intuition into what minimizing EFE is actually doing. Specifically, we introduce a novel computational component for active inference, called *retrospective* (or *residual*) surprise (RS). RS is the surprise from an observation after updating the belief given the observation itself. The predicted value of RS prior to the observation is termed the *predicted retrospective surprise* (PRS). PRS can be mathematically derived from a special case of EFE in a similar way to which (ordinally) surprise appeared from variational free energy (Friston et al., 2006). We show that EFE has a close relationship with PRS. Specifically, EFE provides an upper bound on PRS (i.e., EFE cannot be smaller than PRS). Using a simple example of selecting an action to inspect the binary hidden cause (e.g., disease or not) given binary observation (e.g., positive or negative), we illustrate the properties of EFE and PRS. As we mentioned above, FEP has been interpreted in terms of surprise. Though this is in an information theoretic sense, it helps conceptual understanding. Likewise, PRS will help conceptual understanding of active inference based on EFE.

In the remainder of this paper, we first formalize the problem setting considered in this paper (Section 2). Next, we introduce the key concept of FEP including surprise and variational free energy (Section 3). We then introduce active inference based on EFE in Section 4. In Section 5, we introduce RS and PRS. In Section 6, we provide specific examples for four cases to illustrate the properties of active inference based on EFE and PRS. Finally, in Section 7, we summarize our findings and discuss their implications and limitations.

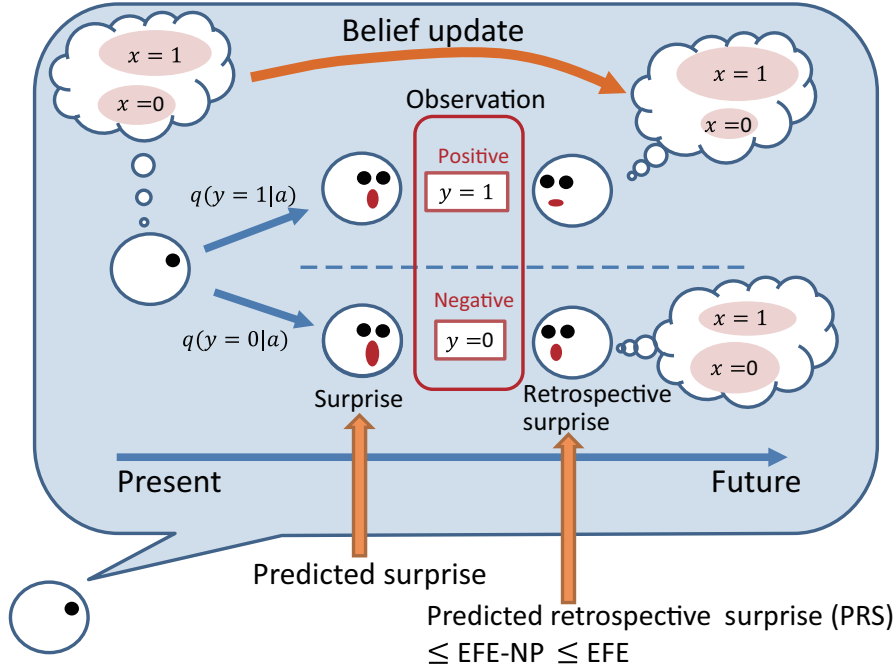


Figure 1: Schematic of problem setting (Case 1a) for active inference and related quantities discussed in this paper. EFE, expected free energy with no preference; EFE-NP, expected free energy with no preference.

2 Problem setting

Here we formalize the general problem setting considered in this paper (Figure 1). In this study, we are only concerned with a very simple case where there is only one hidden discrete variable, denoted by x (e.g., the presence or absence of a specific disease in a patient), which can influence the binary observation (e.g., positive or negative result of the test). An agent selects only one discrete action a (e.g., test or treatment). After the action occurs, the agent observes a single outcome y . After that, the agent may update the belief about x . In the original formulation of active inference based on Markov decision processes (Friston et al., 2015; Friston, FitzGerald, Rigoli, Schwartenbeck, & Pezzulo, 2017), an action is assumed to influence only the subsequent hidden state. Here we take a more general formulation, in which an action may directly affect both (or either of) the state and the observation ¹. That is, we assume that an action can

¹The case where an action directly changes observations has been also considered in previous ideas of active inference where actions are implemented as a reflex (Friston, Daunizeau, Kilner, & Kiebel, 2010).

directly change the observation (without changing the hidden state). In the specific examples, Case 1 assumes that an action directly affects only the observation, while Case 2 assumes that an action directly affects only the hidden state.

The agent has a generative model (in his brain), which is defined as the joint probability of x and y given action a :

$$p(y, x|a) = p(y|x, a)p(x|a), \quad (1)$$

where $p(x|a)$ represents prior belief about x before observing y , which may depend on action a .

The current belief about x is denoted by a distribution $q(x|a)$. In our problem setting, this corresponds to the prior distribution of x before observing y (i.e., $q(x|a) = p(x|a)$). The belief about x and y is denoted by²

$$q(y, x|a) = p(y|x, a)q(x|a). \quad (2)$$

The posterior belief about x after observing y is denoted by $q(x|y, a)$.

3 Surprise and free energy

For our setting, surprise is defined as the negative log-marginal likelihood, in which the hidden variable x is marginalized out:

$$\text{surprise} = -\log p(y|a) = -\log \sum_x p(y|x, a)p(x|a).$$

Because we assume that x is a discrete variable, marginalization is performed by taking the summation over x . Our formulation applies to continuous variable cases if the summation is replaced with integration. When x is binary, the marginalization (summation) over x is not difficult. However, when x is a multidimensional variable or takes multiple values, the marginalization becomes difficult. FEP proposes that the brain minimizes

²Because $p(y|x, a)$ is defined as part of the generative model, we use “ p ” instead of “ q ” for $p(y|x, a)$.

the (variational) free energy, which provides an upper bound on the surprise, rather than directly minimizing the surprise.

Before introducing the (variational) free energy for future observation, we introduce the free energy given an observation y after taking action a , which depends on the belief $q(x|a)$ (denoted by FE):

$$\text{FE} = \sum_x q(x|a) \log \frac{q(x|a)}{p(y, x|a)}. \quad (3)$$

Using the identity $p(y, x|a) = p(x|y, a)p(y|a)$, Eq. (3) is rewritten as

$$\text{FE} = \underbrace{\sum_x q(x|a) \log \frac{q(x|a)}{p(x|y, a)}}_{=D_{KL}(q(x|a)||p(x|y, a))} - \log p(y|a), \quad (4)$$

where $D_{KL}(q(x|a)||p(x|y, a))$ denotes the Kullback–Leibler (KL) divergence between $q(x|a)$ and $p(x|y, a)$. The KL divergence is non-negative, and zero only when $q(x|a) = p(x|y, a)$. From this fact, the free energy provides an upper bound on surprise:

$$\text{FE} \geq -\log p(y|a). \quad (5)$$

4 Active inference

Within the framework of active inference, prior to taking the action a , and before observing y , the agent is supposed to seek an action that minimizes free energy regarding the future observation. Here we introduce EFE (Friston et al., 2015) as a method for evaluating free energy for the upcoming future observation. In a more general formulation of FEP, we consider a policy π , which specifies the set of rules for taking an action and is considered as a random variable (Friston et al., 2009, 2015; Schwöbel et al., 2018). In Friston’s formulation of active inference, the *negative* EFE for each policy is proportional to the prior probability of the policy (Friston et al., 2015): the policy with lower EFE has a higher probability of being chosen. Because in our problem setting, a policy is deemed as selecting a discrete action from a possible candidate action set, we are mainly concerned with the value of EFE for each action.

4.1 Expected free energy

In the free energy for a given observation (Eq. (3)), the unknown hidden variable is only x . In active inference, the future observation y is also dealt with as an unknown hidden variable. EFE is given by incorporating the likelihood of observation y given x and action a , $q(y|x, a)$, into Eq. (3):

$$\text{EFE} = \sum_x q(x|a) \sum_y q(y|x, a) \log \frac{q(x|a)}{p(y, x|a)} \quad (6)$$

$$= \sum_y \sum_x q(y, x|a) \log \frac{q(x|a)}{p(y, x|a)}. \quad (7)$$

We further assume that $p(y, x|a)$ is decomposed as $p(y, x|a) = q(x|y, a)\bar{p}(y)$, where $q(x|y, a)$ denotes the belief about x , given the observation y . $\bar{p}(y)$ encodes prior preference (the utility) on the outcome (Friston et al., 2009, 2015; Schwöbel et al., 2018)³. The reason for associating preference (utility) with probability is that a belief that a given outcome is highly likely adds plausibility to any hypothesized course of action that leads to that outcome being realized. In this sense, preferred outcomes are simply those that appear to be sought through action (i.e., are highly probable under the selected action plans).

In general, the prior preference $\bar{p}(y)$ is not necessarily equal to the marginal likelihood function of y given the current belief and specific action a (i.e., $q(y|a) = \sum_x p(y|x, a)q(x)$). However, in the case without the desired outcome, the marginal likelihood $q(y|a)$ may substitute to $\bar{p}(y)$. This can be regarded as a special case of EFE. We call this quantity “EFE-no preference” (abbreviated as “EFE-NP”). As we see later, for a given belief about hidden states and any action, EFE takes the minimum value when $\bar{p}(y) = q(y|a)$ among other possible prior distributions. Thus, under FEP, this substitution may be natural when there is no prior preference⁴. It should be noted that using EFE-NP differs from using EFE with a uniform, non-informative prior. Using a uniform prior may be a natural choice when no prior preference exists. Setting uniform prior means that the prior does not depend on the action that agent chooses. In contrast, the prior prefer-

³We use the notation “ \bar{p} ” following Schwöbel et al. (2018).

⁴Comparing EFE-NP among different action sets deviates from the active inference framework by Friston because Friston’s framework compares the candidate actions with a single fixed prior preference.

ence in EFE-NP does depend on the selected action. This distinction makes a critical difference in the action selection as discussed later (see Section 7, Discussion).

EFE can be arranged as follows (see Appendix A):

$$\text{EFE} = - \underbrace{\sum_y q(y|a) \sum_x q(x|y, a) \log \frac{q(x|y, a)}{q(x)}}_{= D_{KL}(q(x|y, a)||q(x))} - \underbrace{\sum_y q(y|a) \log \bar{p}(y)}_{= \text{Extrinsic value}}. \quad (8)$$

The KL divergence $D_{KL}(q(x|y, a)||q(x))$ evaluates the degree to which the belief about x is updated by observing y . The expected value of this about y is termed the “epistemic value” (the informativeness of an action). The higher the epistemic value, the lower is the free energy. Thus, the action having a higher epistemic value is likely to be selected. The second term is called “extrinsic value.” This corresponds to the predicted value of surprise. When no prior preference is assigned and the marginal likelihood of y is used for an extrinsic value ($\bar{p}(y) = q(y|a)$), we call the negative of the second term the “predicted surprise.” A smaller predicted surprise or a larger extrinsic value leads to a smaller EFE and the corresponding action is likely to be selected.

EFE can also be arranged in the following form (Friston et al., 2015; see Appendix B for the derivation):

$$\text{EFE} = \underbrace{\sum_x q(x|a) \left\{ - \sum_y q(y|x, a) \log q(y|x, a) \right\}}_{\text{Predicted uncertainty}} + \underbrace{\sum_y q(y|a) \log \frac{q(y|a)}{\bar{p}(y)}}_{\text{Predicted divergence}} \quad (9)$$

The first term, the predicted value of the conditional entropy of y given x and a , represents uncertainty and is called “predicted uncertainty” or “ambiguity.” The second term is KL divergence from $q(y|a)$ to $\bar{p}(y)$. Because KL divergence is a non-negative quantity and takes zero only when two distributions are identical, EFE takes the minimum value when $\bar{p}(y) = q(y|a)$. Thus, we have the inequity

$$\text{EFE} \geq \text{EFE-NP}. \quad (10)$$

Eq. (9) indicates that the agent can likely select an action that brings the marginal distribution $q(y|a)$ closer to the prior preference $\bar{p}(y)$. EFE-NP is always equal to ambiguity

or predicted uncertainty, that is, the first term in Eq. (9), because the second term in Eq. (9) vanishes when $\bar{p}(y) = q(y|a)$. However, to emphasize that this quantity is a special case of EFE that may be used for action selection, we use the term EFE-NP, highlighting its equivalence to ambiguity when pertinent to the matter.

5 Retrospective surprise (RS)

We define RS as a re-evaluated surprise to the observation y after updating the belief given the observation y itself:

$$\text{RS} = -\log \sum_x p(y|x, a)q(x|y, a).$$

PRS is defined as the expected value of RS with respect to y :

$$\text{PRS} = -\sum_y q(y|a) \log \sum_x p(y|x, a)q(x|y, a).$$

Note that because RS and PRS are calculated after obtaining an observation y , these quantities do not depend on the prior preference $\bar{p}(y)$.

When the prior belief about x equals the prior of x (i.e., $q(x|a) = p(x|a)$), which corresponds to the assumption in our problem setting, PRS provides a lower bound on EFE-NP (ambiguity) (see Appendix C for proof):

$$\text{EFE-NP} \geq \text{PRS}. \tag{11}$$

In another view, PRS can be derived from EFE-NP by using Jensen's inequality (see Appendix C). From Eq. (10) and (11), we have the relation

$$\text{EFE} \geq \text{EFE-NP} \geq \text{PRS}. \tag{12}$$

Thus, PRS is a lower bound also on (original) EFE.

One condition of the equality in Eq (11) is that x and y are independent given a , i.e., $q(y, x|a) = q(y|a)q(x|a)$. In this case, the epistemic value is zero and we have

Table 1: List of the binary tests and related quantities for Case 1a. The prior belief about x is set as $p_D = 0.6$. EFE-NP, expected free energy with no preference; PRS, predicted retrospective surprise.

	Hit rate (p_H)	False alarm rate (p_F)	Predicted surprise	Epistemic value	EFE-NP	PRS
Test 1	0.8	0.2	0.686	0.186	0.5	0.387
Test 2	0.8	0.8	0.5	0	0.5	0.5

$$\text{EFE-NP} = \text{PRS} = - \sum_y q(y|a) \log q(y|a).$$

This coincides with the predicted surprise. Another condition for the equality is that the relation between x and y is one-to-one, i.e., in the binary case, possible combinations of x and y are (i) $(x, y) \in \{(0, 0), (1, 1)\}$ (i.e., y always equals x) or (ii) $(x, y) \in \{(0, 1), (1, 0)\}$ (i.e., y always differs from x). In these cases, PRS and EFE-NP are zero.

6 Illustrative examples

By using simple examples of diagnostic inference from a single patient (Kunisato, Katahira, Okimura, & Yamashita, 2019), we illustrate the properties of the quantities discussed above. Let us consider the situation in which an agent infers the presence ($x = 1$) or absence ($x = 0$) of a disease in a patient from a test result, which can be positive ($y = 1$) or negative ($y = 0$). The agent has prior belief about x , which is denoted by $q(x|a)$. For brevity, we represent this as $q(x = 1|a) = p_D$ and $q(x = 0|a) = 1 - p_D$ (“ D ” indicates “disease”). The agent knows the properties of the candidate tests, which are specified by the hit rate, $p_H = p(y = 1|x = 1)$, and the false alarm rate, $p_F = p(y = 1|x = 0)$. This knowledge corresponds to a generative model the agent has and is not necessarily true.

The first examples (Case 1a and Case 1b) consider the situation where the agent chooses a test that produces a binary outcome. In this case, the action (selecting a test) affects the outcome (positive or negative) but not the hidden state (disease or not). In

the second example (Case 2), the agent chooses a treatment that can affect the hidden state, but the outcome of the test is independent of the action given the state.

6.1 Test selection with no prior preference (Case 1a)

Case 1a considers the problem of selecting a test among tests that differ in the hit rate p_H and false alarm rate p_F . Here we assume that there is no prior preference and set the extrinsic value to the marginal likelihood of y ($\bar{p}(y) = q(y|a)$) for each test (i.e., we assume the agent choose a test with smallest EFE-NP).

First, we discuss how active inference works in this case, along with specific scenarios based on the schematic in Figure 1. An agent is about to choose an action (test). He is now considering Test 1 as a candidate (Table 1). He has prior probability about x , for example, $q(x = 1) = p_D = 0.6$ and $q(x = 0) = 1 - p_D = 0.4$ (Here, because we assume that x does not depend on a , the condition “ $|a$ ” is omitted). He knows that the hit rate of Test 1 is $p_H = 0.8$ and the false alarm rate is $p_F = 0.2$. Thus, under his prior belief and this generative model, he expects that the test results will be positive with the probability of $q(y = 1|a) = p(y = 1|x = 1) \times q(x = 1) + p(y = 1|x = 0) \times q(x = 0) = p_H p_D + p_F(1 - p_D) = 0.8 \times 0.6 + 0.2 \times 0.4 = 0.56$ (See Appendix D for the details of calculation). Thus, he predicts that he will experience a surprise after observing a positive result with the degree $-\log(0.56) = 0.58$. A negative result is predicted to occur with the probability $q(y = 0|a) = 1 - q(y = 1|a) = 0.44$. When the result is negative, he will be more surprised with the degree $-\log(0.44) = 0.82$. Thus, the extrinsic value (negative predicted surprise) is $0.56 \times 0.58 + 0.44 \times 0.82 = 0.686$.

After observing the result, he will update his belief about the hidden variable x . If the update is according to the Bayesian inference, when the result is positive ($y = 1$), the belief will be $q(x = 1|y = 1) = p_H p_D / q(y = 1|a) = 0.8 \times 0.6 / 0.56 = 0.857$, while $q(x = 0|y = 1) = 0.143$ (see Appendix D). With this (updated) belief, the RS to the positive result is $-\log(p_H \times 0.857 + p_F \times 0.143) = 0.336$. When the result is negative ($y = 0$), the belief will be updated as $q(x = 1|y = 0) = (1 - p_H) p_D / q(y = 0|a) = 0.2 \times 0.6 / 0.44 = 0.273$ and $q(x = 0|y = 0) = 0.727$. The RS to the negative result is $-\log((1 - p_H) \times 0.273 + (1 - p_F) \times 0.727) = 0.452$. Thus, the PRS is $0.56 \times 0.336 + 0.44 \times 0.452 = 0.387$. This is greatly reduced from the predicted

surprise (= 0.664).

Next, consider Test 2, which is not informative at all (Table 1) and whose hit rate and false alarm rate are both 0.8, i.e., it produces a result irrespective of the hidden variable x . In this case, the belief about x does not change even after seeing the result of the test, because the observation provides no information. Thus, PRS has the same value as that of the predicted surprise (= 0.50, see Appendix D). This is lower than the predicted surprise of Test 1 (= 0.664) but higher than the PRS of Test 1. Thus, if the agent just wants to minimize the predicted surprise, he should select the non-informative test, Test 2. If he wants to minimize PRS, he should choose the informative test, Test 1.

In contrast, EFE-NP (ambiguity) is insensitive to the informativeness of the tests: Test 1 and Test 2 provide the same EFE-NP, 0.5 (Table 1). This means that the active inference based on EFE-NP may fail to select a more informative action. Let us consider the reason for this. From Eq. (9), when there is no prior preference and $\bar{p}(y) = q(y|a)$, only the first term, the predicted uncertainty (conditional entropy) matters for EFE-NP. When $q(x|a)$ does not depend on a (i.e., $q(x|a) = q(x)$) and the entropy given x , $-\sum_y q(y|x, a) \log q(y|x, a)$, is identical (= $-0.2 \log 0.2 - 0.8 \log 0.8$) for both $x = 0$ and $x = 1$, the conditional entropy is identical for both tests. Thus, EFE-NP takes the same value for both Test 1 and Test 2.

Next, let us check the epistemic and extrinsic values (predicted surprise; Eq. (8)). The epistemic value of Test 1 is 0.186, while that of Test 2 is zero. Because EFE-NP is a negative predicted surprise minus the negative epistemic value (Eq. (8)), the effect of the epistemic value of Test 1 is canceled by the larger predicted surprise of Test 1 (0.686, which is larger than Test 2 by 0.186).

In Figure 2, we systematically vary the hit rate and false alarm rate. We can confirm that the epistemic value increases as the information in the test increases (Figure 2C): In the region where the hit rate p_H is larger than the false alarm rate p_F , the epistemic value increases with p_H . In another region where $p_H < p_F$, the smaller p_H is, the larger is the epistemic value. The negative predicted surprise shows an opposite tendency (Figure 2B). These two effects completely cancel out each other: the advantage of the large epistemic value in an informative test is abolished by its large surprise. As a result, EFE-NP is insensitive to the informativeness of the test. For example, the tests with $p_F = 0.2$ and 0.8 yield exactly the same EFE-NP, irrespective of the value of p_H

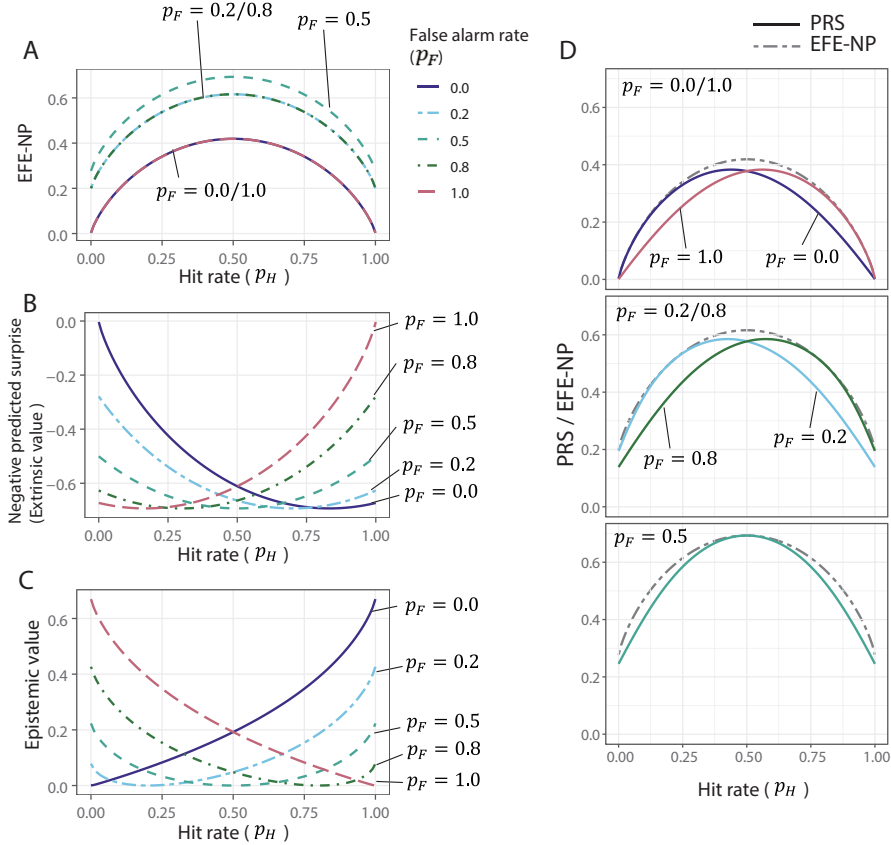


Figure 2: How the hit (p_H) rate and false alarm rate (p_F) influence the quantities for selecting the binary test in active inference (Case 1a). (A): EFE-NP. (B): The negative predicted surprise (extrinsic value). (C): The epistemic value. (D): PRS and EFE-NP.

(Figure 2A).

In accordance with the inequality Eq. (11), PRS is upper-bounded by (is equal to or smaller than) EFE-NP (Figure 2D). In addition, these quantities coincide only when (i) $p_H = p_F$ or (ii) p_H is zero or one and p_F is zero or one. Importantly, PRS has a lower value for a more informative test except for the case in which p_H and p_F are exactly zero or one. In the region where p_H is larger than 0.5, the smaller p_F is, the smaller is the PRS. In contrast, when p_H is smaller than 0.5, the opposite occurs. These results indicate that agents can choose a more informative test based on PRS.

Interestingly, even with defined preferences, EFE of Eq. (6) would perform similarly to EFE-NP in this task⁵. The reason for this is that this task is designed in such

⁵We thank an anonymous reviewer for this suggestion.

Table 2: List of binary tests and related quantities for Case 1b. The prior belief about x is set as $p_D = 0.6$ and prior preference about y is set as $\bar{p}(y = 1) = 0.6$.

	Hit rate (p_H)	False alarm rate (p_F)	Extrinsic value	Epistemic value	EFE	PRS
Test 1	0.8	0.2	-0.689	0.186	0.50	0.387
Test 2	0.8	0.8	-0.591	0	0.59	0.5

a way as to preclude actions influencing x . When Eq. (6) is rearranged into its ‘risk’ and ‘ambiguity’ form, it has the same appearance as Eq. (9), but with the predicted divergence expressed in terms of the divergence between states given actions and a prior over states. If actions do not influence states, this divergence is invariant to the choice of action, leaving only the ambiguity term (as in EFE-NP) to disambiguate between alternative choices. The importance of this observation is that Eq. (6) could be approximated by Eq. (9) in applications of active inference. This has the advantage of allowing for preferences in terms of outcomes without having to assume a form for the rest of the generative model. The results here highlight an additional advantage, as they show that Eq. (9) (and the PRS) enable epistemic behavior even when states are invariant to action.

6.2 Effect of prior preference (Case 1b)

In the next case (Case 1b), we consider the effects of prior preference for the observation y . That is, we suppose that the agent assumes that a certain observation (positive or negative) is more probable. Although such situation seems to be strange in this (inference on disease) example, this setting should be taken for an illustration purpose. We set $p_D = 0.6$, as in Case 1a. We also set prior preference as $\bar{p}(y = 1) = 0.6$. Again, let us compare Test 1 and Test 2 (Table 2). Here, we set a prior preference about y as $\bar{p}(y = 1) = 0.6$: the agent expects that the positive outcome is slightly more probable. The extrinsic value is now larger for Test 2, which more likely produces positive results (when $x = 0$). Because the epistemic value does not depend on the prior preference about y (also see Figure 3C), these are the same as Case 1a (Table 1) and Test 1 has higher epistemic value compared to Test 2. As a result, Test 1 has a lower EFE value

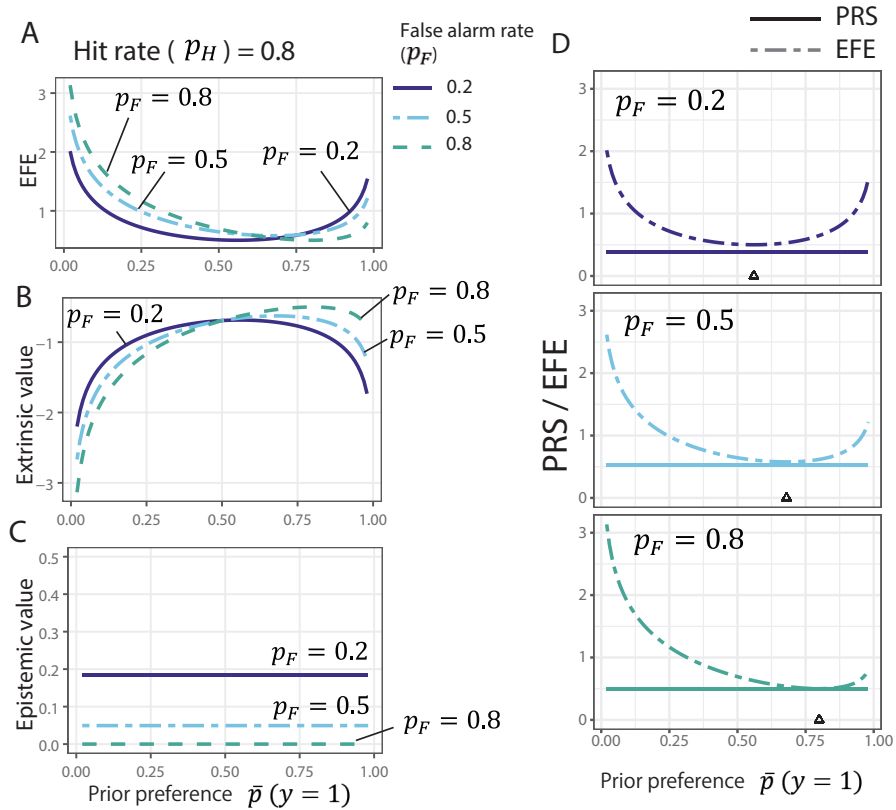


Figure 3: Effect of prior preference (Case 1b). (A): EFE. (B): The extrinsic value. (C): The epistemic value. (D): PRS and EFE. The triangles indicate the marginal likelihood $q(y = 1|a)$.

compared to Test 2 and is favored.

Figure 3 shows how EFE and PRS depend on the prior preference $\bar{p}(y = 1)$. We compare Test 1 ($p_H = 0.8, p_F = 0.2$), Test 2 ($p_H = 0.8, p_F = 0.8$), and a test with intermediate false positive rate ($p_H = 0.8, p_F = 0.5$). When $y = 1$ is preferred over $y = 0$ (i.e., $\bar{p}(y = 1) > 0.5$), the extrinsic value is larger for the tests that have a higher false alarm rate. In the opposite case ($\bar{p}(y = 1) < 0.5$), the test that has a smaller false alarm rate has a higher extrinsic value (Figure 3B). As the informative test, Test 1, has the highest epistemic value that decreases EFE the most, when the preference is around neutral ($\bar{p}(y = 1) \approx 0.5$), Test 1 yields the smallest EFE, and thus, is favored. Test 2 is favored (given a smaller EFE) when $\bar{p}(y = 1)$ is more than around 0.7, where the effect of the extrinsic value overcomes that of the epistemic value (Figure 3A). This confirms the notion that action selection based on EFE is driven by the trade-off between

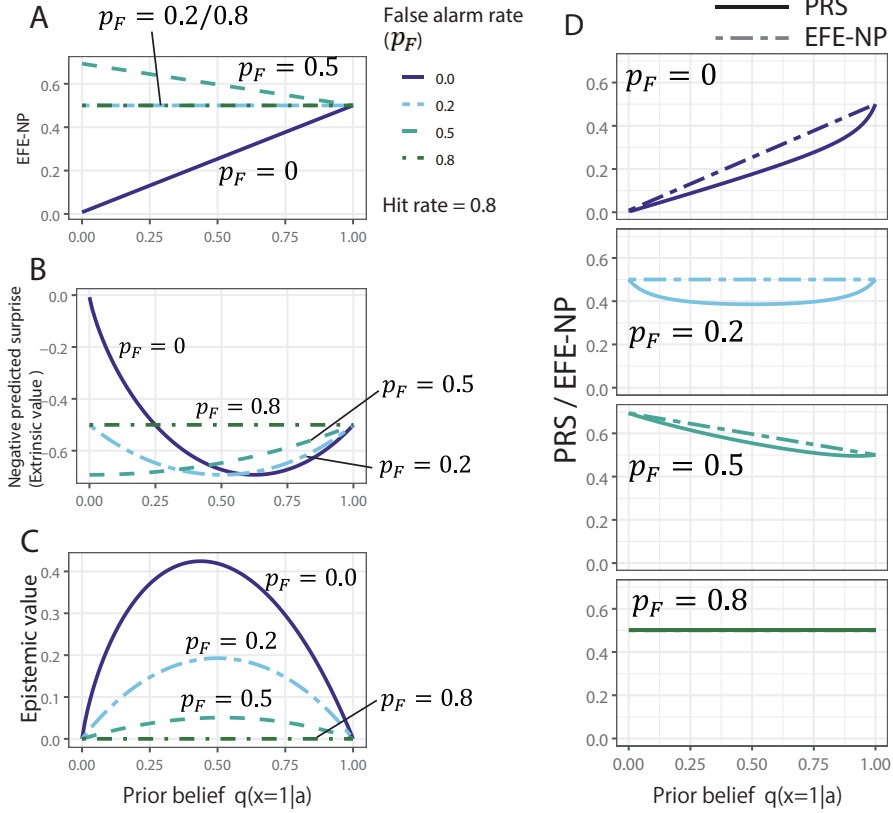


Figure 4: Effect of prior belief (which may depend on action) about hidden state x with no preference (Case 2a). (A): EFE-NP, (B): The negative predicted surprise (extrinsic value). (C): The epistemic value, (D): PRS and EFE.

exploration (maximizing epistemic value) and exploitation (maximizing extrinsic value) (Friston et al., 2015).

Meanwhile, PRS does not depend on prior preference by construction. From Figure 3D), we can confirm that EFE is lower-bounded by PRS and minimum when $\bar{p}(y) = q(y|a)$ (marked by triangles in the figure).

6.3 Cases where an action influences the hidden state (Case 2a and Case 2b)

In Case 2a and Case 2b, we assume that an action a influences the hidden state x rather than the observation y . This assumption corresponds to that adopted in the original active inference framework based on EFE (Friston et al., 2015).

To clarify the problem setting, let t be the index of discrete time steps. The action at the current time t , a_t , influences the state at the next time step x_{t+1} . The agent assumes that x changes depending on action a_t with the probability $p(x_{t+1}|x_t, a_t)$. Based on this generative model and the current belief about x_t , the prior belief about the next state is given by $q(x_{t+1}|a_t) = \sum_{x_t} p(x_{t+1}|x_t, a_t)q(x_t)$. EFE and PRS at time t are conditionally independent of x_t given x_{t+1} . Thus, checking how EFE and PRS depend on $q(x_{t+1} = 1|a_t)$ would suffice. Hereafter, we omit the index of time step, t (i.e., $q(x|a) = q(x_{t+1}|a_t)$).

Figure 4 shows how EFE-NP and PRS vary depending on the prior belief $q(x = 1|a)$ when no prior preference is set (Case 2a). Because in this case, the predicted divergence in Eq. (9) is zero and the entropy terms are fixed, EFE-NP is a linear function of $q(x = 1|a)$. Based on EFE-NP, an agent favors assigning a larger probability to x , which produces a smaller entropy. For example, when the false alarm rate p_F is 0 while the hit rate $p_H = 0.8$, $x = 0$ minimizes the entropy⁶, and thus, $q(x = 1|a) = 0$ produces the smallest EFE-NP. When p_F is 0.5 while the hit rate $p_H = 0.8$, the entropy with $x = 1$ is smaller than that with $x = 0$; thus, $q(x = 1|a) = 1$ produces the smallest EFE-NP. When both values of x produce the same entropy ($p_F = 0.2$ or $p_F = 0.8$, and $p_H = 0.8$), EFE-NP is independent of $q(x|a)$.

In contrast, PRS is a nonlinear function of $q(x = 1|a)$, except for the case with a non-informative test ($p_H = p_F = 0.8$; Figure 4D). Especially, when $p_F = 0.2$, PRS and EFE-NP yield quantitatively different predictions: PRS is smaller for an action that makes $q(x|a)$ more ambiguous ($q(x = 1|a)$ is close to the intermediate value), while EFE-NP is neutral regarding action selection.

Finally, we consider how the prior preference works for this setting. We assume that the agent has a prior belief (preference) that the negative result is more likely than the positive result: $\bar{p}(y = 1) = 0.2$. Figure 5 shows the behaviors of EFE and PRS. Because PRS does not depend on prior preference, the results are identical to Case 2a (Figure 4D). Basically, EFE has a smaller value for $q(x = 1|a)$, which provides a larger extrinsic value (Figure 5B). Because the epistemic value is larger for the intermediate value of $q(x = 1|a)$, the minimum of EFE shifts toward the intermediate value of

⁶ $-\sum_y q(y|x=0, a) \log q(y|x=0, a) = -q(y=1|x=0, a) \log q(y=1|x=0, a) - q(y=0|x=0, a) \log q(y=0|x=0, a) = -0 \log 0 - 1 \log 1 = 0$.

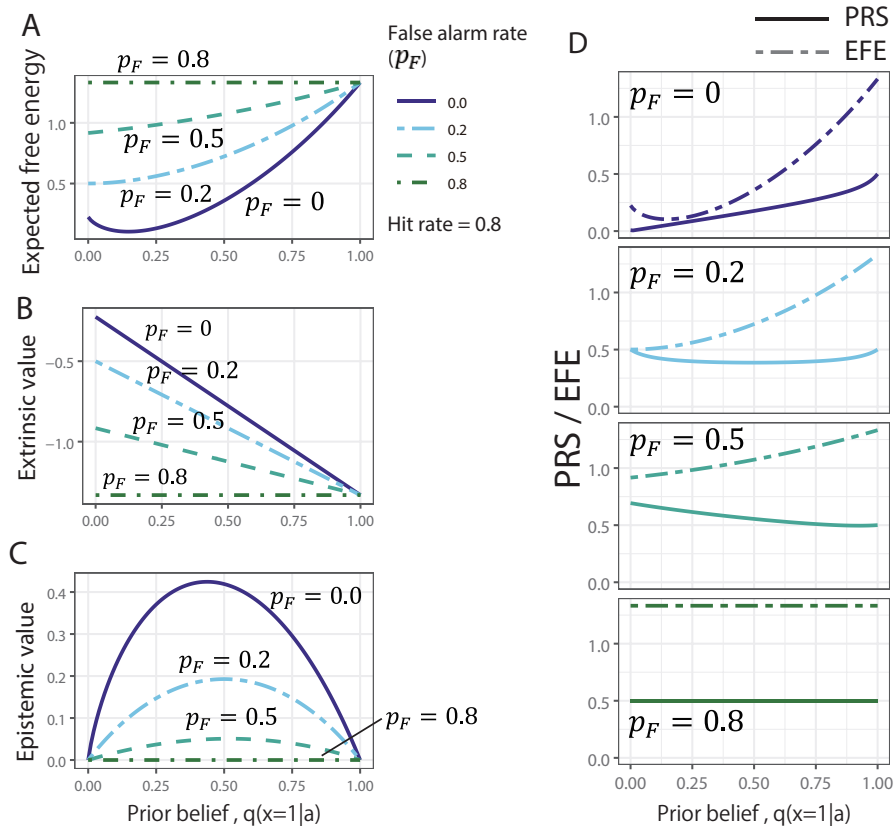


Figure 5: Effect of prior belief about hidden state x with fixed prior preference ($\bar{p}(y = 1) = 0.2$). The conventions are the same as those in Fig. 3

$q(x = 1|a)$ when the false alarm rate is small ($p_F = 0$).

7 Discussion

In this paper, we proposed a novel computational construct, the predicted retrospective surprise (PRS), and discussed its relation to expected free energy (EFE), which is a core component for active inference in free energy principle (FEP). We clarified one computational aspect of EFE: it provides an upper bound on PRS. Below, we discuss the contributions and implications of this study.

This study provides new insights on the principle of action selection, i.e., action selection that minimizes EFE, which leads to the minimization of surprise that remains after the belief is updated from the consequence of the action. As mentioned in the In-

roduction, FEP posits that living systems seek to minimize surprise (in the information theory sense); that is, minimizing free energy leads to minimizing surprise. A naïve application of a notion that action selection in consideration of future observation leads to the prediction that agents seek action that produces less informative but more stable outcomes because such action leads to smaller surprise (i.e., in the diagnostic example, this corresponds to the prediction that an agent prefers a noninformative test to an informative test). However, sometimes reality might be different; that is, animals often favor actions from which they can gain information (Hampton, Zivin, & Murray, 2004; Castro & Wasserman, 2013). If one assumes that living systems seek an action that minimize PRS (such an action often minimizes EFE) rather than predicted surprise, one can naturally explain such information-seeking (exploration) behavior. The explanation of such behavior is also possible based on EFE itself, as Friston et al. (2015) did. However, because interpreting EFE as a single psychological quantity is difficult, the intuitive understanding of how the information seeking behavior arises from EFE is less straightforward: specifically, although its mathematical derivation is clear, the intuitive understanding of how epistemic value arises from EFE is not easy. PRS provides an intuitive account of how curiosity arises: informative action makes the hidden state less ambiguous, thus reduces PRS.

This study also contributes to the understanding of how active inference based on EFE works. First, our results highlight the role of prior preference (utility) in EFE. By setting this preference as the prior distribution, which is independent of action, the epistemic value can exert an influence on EFE value and thus action selection. We demonstrate that if the negative predicted surprise, which depends on action, is used as the extrinsic value, the effect of the epistemic value is diminished by the negative predicted surprise. Indeed, in certain situations it is completely canceled. If the prior preference is fixed across candidate actions, this reduction or cancelation does not occur. It should be emphasized that this is the case even with a uniform, non-informative prior, which may be a natural choice when no prior preference exists. Thus, the use of prior preference irrespective of the action is a key premise of active inference based on EFE. Second, we show how the prior preference influences EFE. Given a specific value of belief about the hidden state x , EFE is minimum when the prior preference $\bar{p}(y)$ coincides with the predicted marginal likelihood of observation, $q(y|a)$ (cf. Figure 3D;

for this case, EFE is called EFE-NP and is equivalent to ambiguity). This means that as the prior preference departs from $q(y|a)$, EFE increases. Thus, when there is no preference regarding the outcome, one natural choice under FEP is to assume that $\bar{p}(y) = q(y|a)$ for each candidate action. Active inference based on EFE can be regarded as an interplay between desire for the future outcome, which increases EFE and any effort to reducing EFE (including compromise about future outcome; too unrealistic desire may be avoided). EFE-NP provides the most compromised value of EFE.

Although PRS and EFE-NP show similar behavior, in some cases, the actions that minimize them are different. This indicates that although EFE-NP provides an upper bound on PRS, the bound is not tight in some situations. One such case is the selection of tests that have same ambiguity (conditional entropy) but different information (Test 1 vs. Test 2 in Case 1a). In other cases where the action directly influences the hidden state, PRS tends to favor the action that makes the belief about the hidden state more ambiguous.

Nevertheless, we do not intend to claim that the principle of action selection of humans or other animals is better described by PRS rather than EFE or EFE-NP. Although their predictions are different in certain situations, the prediction of any Bayesian model including active inference models depends on a generative model that is difficult to infer from the experiment. Therefore, a rigorous test of the frameworks of active inference based on the experimental data is a difficult task. We consider that the contribution of the present study is to propose a psychological interpretation of EFE as a single quantity. It should be noted that PRS is just one quantity for which EFE is upper bound on: there may be different quantities, which may provide a tighter bound than PRS. While we focused on PRS because its analytical expression is psychologically interpretable, we cannot exclude the existence of another quantity that provides a different interpretation of EFE (as a single quantity). Future research may be required to explore this quantity.

The findings of this study should be evaluated considering some limitations. We have only considered a very simple situation of action selection—only single discrete action is chosen and there is a single hidden cause. In a future study, we should consider whether these results can be applied to more general problems such as Markov decision processes, where future multiple actions should be considered, and to action selection

in continuous time (e.g., Buckley et al., 2017). Implementation of active inference for a realistic problem is still under debate. Computing EFE or predicted free energy on large-scale Markov decision problems needs some approximation, e.g., mean-field approximation (Friston et al., 2015) or Bethe approximation (Schwöbel et al., 2018). While action selection based on PRS may have potential for improving in a certain situation, implementing this in a manner that is adjustable to a realistic problem also remains an issue for future work.

While PRS is interpretable as a psychological construct, its function from the normative perspective is not clear: Does action selection based on PRS increase the survival rate of living systems? If this is the case, under what circumstance? To address these questions, realistic model settings are needed because this question depends on the environment of the agents. Elucidating the functional role of PRS may lead to an understanding of the evolutionary origin of curiosity or intrinsic motivation.

In conclusion, this paper proposed a novel computational component, PRS, for action selection. We indicated that Friston’s expected free energy is closely related to (provides an upper bound on) PRS. This study may contribute to understanding and refining the computational theory of active inference, and its related processes such as learning, attention, behavioral control, and homeostatic regulation (Pezzulo, Rigoli, & Friston, 2015).

Acknowledgments

This study was supported by JSPS KAKENHI Grant Numbers JP16H05957, JP17H05946, JP17H06039, JP18KT0021, JP18K03173, JP18K07597, JP25330301, and JP26120002, and JST CREST Grant Number JPMJCR16E2. The authors thank Asako Toyama and Michiyo Sugawara for their helpful discussion and Masatoshi Yoshida for his thoughtful comments on our study and kind introduction of active inference on his blog⁷. The authors also thank two anonymous reviewers for providing critical and constructive comments on the earlier version of the manuscript.

⁷<http://pooneil.sakura.ne.jp/archives/permalink/001663.php> (in Japanese)

A Derivation of Eq (8)

The transformation of EFE given in Eq (8) is conducted as follows (Friston et al., 2015):

$$\begin{aligned}
\text{EFE} &= \sum_y \sum_x q(y, x|a) \log \frac{q(x|a)}{p(y, x|a)} \\
&= \sum_y \sum_x q(y, x|a) \log \frac{q(x|a)}{q(x|y, a)\bar{p}(y)} \\
&= \sum_y q(y|a) \sum_x q(x|y, a) \log \frac{q(x|a)}{q(x|y, a)} - \sum_y \sum_x q(y, x|a) \log \bar{p}(y) \\
&= - \sum_y q(y|a) \sum_x q(x|y, a) \log \frac{q(x|y, a)}{q(x|a)} - \sum_y q(y|a) \log \bar{p}(y).
\end{aligned}$$

B Derivation of Eq. (9)

The EFE rearrangement given in Eq. (9) is conducted as follows (Appendix A in Friston et al., 2015):

$$\text{EFE} = \sum_y \sum_x q(y, x|a) \log \frac{q(x|a)}{p(y, x|a)} \quad (13)$$

$$= \sum_y \sum_x q(y, x|a) [\log q(x|a) - \log \bar{p}(y) - \log q(x|y, a)] \quad (14)$$

$$= \sum_y \sum_x q(y, x|a) \left[\log q(x|a) - \log \bar{p}(y) - \log \frac{q(y|x, a)q(x|a)}{q(y|a)} \right] \quad (15)$$

$$= \sum_y \sum_x q(y, x|a) \left[-\log q(y|x, a) + \log \frac{q(y|a)}{\bar{p}(y)} \right] \quad (16)$$

$$= \sum_x q(x|a) \left\{ - \sum_y q(y|x, a) \log q(y|x, a) \right\} + \sum_y q(y|a) \log \frac{q(y|a)}{\bar{p}(y)}. \quad (17)$$

To obtain Eq. (15), we use the Bayes' theorem.

C Proof of the inequality in Eq. (11)

By Jensen's inequality, we can find that EFE-NP (ambiguity) provides an upper bound on PRS as follows:

$$\begin{aligned}
\text{EFE-NP} &= - \sum_y q(y|a) \sum_x q(x|y, a) \log \frac{p(y|x, a)q(x|a)}{q(x|a)} \\
&= - \sum_y q(y|a) \sum_x q(x|y, a) \log p(y|x, a) \\
&\geq - \sum_y q(y|a) \log \sum_x p(y|x, a)q(x|y, a) = \text{PRS}.
\end{aligned}$$

D Calculation for the binary test

Here we detail the specific calculation for the diagnostic inference example. Although this might be redundant, we think it is useful to grasp the calculation procedure of active inference for readers who are not familiar with such calculations.

By setting specific values for p_H (hit rate), p_F (false alarm rate), and p_D (prior probability of disease), and by substituting these into the following equations, one can obtain specific values of the quantities discussed in this paper.

Here we assume that the belief is updated according to Bayes's theorem. Then, the posterior of x given observation y , $q(x|y, a)$, is computed as follows:

$$\begin{aligned}
q(x = 1|y = 1, a) &= \frac{p(y = 1|x = 1, a)q(x = 1)}{p(y = 1)} \\
&= \frac{p_H p_D}{p_H p_D + p_F(1 - p_D)}, \\
q(x = 1|y = 0, a) &= \frac{p(y = 0|x = 1, a)q(x = 1)}{p(y = 0)} \\
&= \frac{(1 - p_H)p_D}{(1 - p_H)p_D + (1 - p_F)(1 - p_D)}.
\end{aligned}$$

Using these quantities, we have $q(x = 0|y = 1, a) = 1 - q(x = 1|y = 1, a)$ and $q(x = 0|y = 0, a) = 1 - q(x = 1|y = 0, a)$. If the test is noninformative ($p_H = p_F$), we can confirm that the posterior probabilities do not change from the prior probabilities (e.g., $q(x = 1|y = 1, a) = p_D$ and $q(x = 0|y = 1, a) = 1 - p_D$).

The marginal likelihood of y is computed as

$$\begin{aligned}
q(y = 1|a) &= \sum_x p(y = 1|x, a)q(x|a) \\
&= p(y = 1|x = 0, a)q(x = 0) + p(y = 1|x = 1, a)q(x = 1) \\
&= p_F(1 - p_D) + p_H p_D,
\end{aligned}$$

and $q(y = 0|a) = 1 - q(y = 1|a)$. Using these values, the epistemic value is calculated as

$$\begin{aligned}
\text{epistemic value} &= \sum_y p(y|a) \sum_x q(x|y, a) \log \frac{q(x|y, a)}{q(x|a)} \\
&= q(y = 0|a)q(x = 0|y = 0, a) \log \frac{q(x = 0|y = 0, a)}{q(x = 0)} \\
&\quad + q(y = 0|a)q(x = 1|y = 0, a) \log \frac{q(x = 1|y = 0, a)}{q(x = 1)} \\
&\quad + q(y = 1|a)q(x = 0|y = 1, a) \log \frac{q(x = 0|y = 1, a)}{q(x = 0)} \\
&\quad + q(y = 1|a)q(x = 1|y = 1, a) \log \frac{q(x = 1|y = 1, a)}{q(x = 1)}.
\end{aligned}$$

The extrinsic value is computed as

$$\begin{aligned}
\text{extrinsic value} &= \sum_y q(y|a) \log \bar{p}(y) \\
&= q(y = 0|a) \log \bar{p}(y = 0) + q(y = 1|a) \log \bar{p}(y = 1).
\end{aligned}$$

One can obtain EFE by

$$\text{EFE} = -\text{epistemic value} - \text{extrinsic value}.$$

Using the above quantities, the explicit form of PRS is given as

$$\begin{aligned}
\text{PRS} &= - \sum_y q(y|a) \log \sum_x p(y|x, a)q(x|y, a) \\
&= - (p_H p_D + p_F(1 - p_D)) \log \left(\frac{p_H^2 p_D + p_F^2 p_D}{p_H p_D + p_F(1 - p_D)} \right) \\
&\quad - ((1 - p_H)p_D + (1 - p_F)(1 - p_D)) \log \left(\frac{(1 - p_H)^2 p_D + (1 - p_F)^2 (1 - p_D)}{(1 - p_H)p_D + (1 - p_F)(1 - p_D)} \right).
\end{aligned}$$

References

- Bogacz, R. (2017). A tutorial on the free-energy framework for modelling perception and learning. *Journal of Mathematical Psychology*, *76*, 198–211.
- Buckley, C. L., Kim, C. S., McGregor, S., & Seth, A. K. (2017). The free energy principle for action and perception: A mathematical review. *Journal of Mathematical Psychology*, *81*, 55–79.
- Castro, L., & Wasserman, E. A. (2013). Information-seeking behavior: Exploring metacognitive control in pigeons. *Animal Cognition*, *16*(2), 241–254.
- Friston, K. (2010). The free-energy principle: a unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127.
- Friston, K., Daunizeau, J., & Kiebel, S. J. (2009). Reinforcement learning or active inference? *PLOS ONE*, *4*(7), e6421.
- Friston, K., Daunizeau, J., Kilner, J., & Kiebel, S. J. (2010). Action and behavior: a free-energy formulation. *Biological Cybernetics*, *102*(3), 227–260.
- Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. (2017). Active inference: a process theory. *Neural Computation*, *29*(1), 1–49.
- Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology-Paris*, *100*(1-3), 70–87.
- Friston, K., Mattout, J., & Kilner, J. (2011). Action understanding and active inference. *Biological Cybernetics*, *104*(1-2), 137–160.
- Friston, K., Rigoli, F., Ognibene, D., Mathys, C., Fitzgerald, T., & Pezzulo, G. (2015). Active inference and epistemic value. *Cognitive Neuroscience*, *6*(4), 187–214.
- Hampton, R. R., Zivin, A., & Murray, E. A. (2004). Rhesus monkeys (*macaca mulatta*) discriminate between knowing and not knowing and collect information as needed before acting. *Animal Cognition*, *7*(4), 239–246.
- Kunisato, Y., Katahira, K., Okimura, T., & Yamashita, Y. (2019). *Keisanron-teki Seishin Igaku (Computational Psychiatry)*. Tokyo: Keiso Shobo (in Japanese).
- Parr, T., & Friston, K. (2018). Generalised free energy and active inference: can the future cause the past? *BioRxiv*, 304782.
- Pezzulo, G., Rigoli, F., & Friston, K. (2015). Active inference, homeostatic regulation and adaptive behavioural control. *Progress in Neurobiology*, *134*, 17–35.

Schwöbel, S., Kiebel, S., & Marković, D. (2018). Active inference, belief propagation, and the bethe approximation. *Neural Computation*, 30(9), 2530–2567.