# Statistical Learning Model of the Sense of Agency

**Shiro Yano[1]\*, Yoshikatsu Hayashi[2], Yuki Murata[1], Hiroshi Imamizu[3], Takaki Maeda[4,5] and Toshiyuki Kondo[1]**

[1] Division of Advanced Information Technology & Computer Science, Tokyo University of Agriculture and Technology, Tokyo, Japan, [2] Biomedical Science and Biomedical Engineering, School of Biological Science, University of Reading, Berkshire, United Kingdom, [3] Department of Psychology, The University of Tokyo, Tokyo, Japan, [4] Department of Neuropsychiatry, Keio University School of Medicine, Tokyo, Japan, [5] Center for Psychiatry and Behavioral Science, Komagino Hospital, Tokyo, Japan

A sense of agency (SoA) is the experience of subjective awareness regarding the control of one's actions. Humans have a natural tendency to generate prediction models of the environment and adapt their models according to changes in the environment. The SoA is associated with the degree of the adaptation of the prediction models, e.g., insufficient adaptation causes low predictability and lowers the SoA over the environment. Thus, identifying the mechanisms behind the adaptation process of a prediction model related to the SoA is essential for understanding the generative process of the SoA. In the first half of the current study, we constructed a mathematical model in which the SoA represents a likelihood value for a given observation (sensory feedback) in a prediction model of the environment and in which the prediction model is updated according to the likelihood value. From our mathematical model, we theoretically derived a testable hypothesis that the prediction model is updated according to a Bayesian rule or a stochastic gradient. In the second half of our study, we focused on the experimental examination of this hypothesis. In our experiment, human subjects were repeatedly asked to observe a moving square on a computer screen and press a button after a beep sound. The button press resulted in an abrupt jump of the moving square on the screen. Experiencing the various stochastic time intervals between the action execution (button-press) and the consequent event (square jumping) caused gradual changes in the subjects' degree of their SoA. By comparing the above theoretical hypothesis with the experimental results, we concluded that the update (adaptation) rule of the prediction model based on the SoA is better described by a Bayesian update than by a stochastic gradient descent.

Keywords: sense of agency, statistical learning, online learning, Bayes' rule, stochastic gradient descent

## 1. INTRODUCTION

The sense of agency(SoA) is the subjective awareness about "the experience of controlling one's own motor acts and, through them, the course of external events" (Haggard, 2017). The SoA is regarded as a fundamental role in maintaining continuous self-consciousness (Gallagher, 2000). Normally, we are unaware of the existence of the SoA because it exists behind our daily actions. An explicit awareness of the decreased SoA arises as a result of a striking conflict of expectations, i.e., a mismatch between the intended and the actual result of an action (Haggard, 2017).

The SoA has been studied in various disciplines, ranging from psychiatry (Maeda et al., 2012, 2013) to human computer interaction (HCI) domains and brain computer interface (BCI) domains (Coyle et al., 2012; Minohara et al., 2016; Cornelio Martinez et al., 2017). Measuring and understanding the SoA plays an important role in these domains. Disruptions in the SoA in movement disorders have major implications for quality of life (Haggard, 2017). The SoA can be the indicator for identifying the user's experience for an appropriate sense of operation in HCI and BCI domains, as the SoA is associated with the awareness of the response of one's own action.

Previous researchers have assumed that there should exist at least two different generative processes of the SoA: prospective and retrospective processes (Wolpe and Rowe, 2014; Haggard, 2017). The comparator model, which is our focus in this research, corresponds to the former, while the apparent mental causation model corresponds to the latter (Wegner and Wheatley, 1999). **Figure 1** shows a schematic picture of the comparator model. The comparator model was originally proposed to describe the process of executing motor controls (**Figure 1**) (Wolpert et al., 1995; Blakemore et al., 2001). For this process, motor command causes actual bodily dynamics and changes in the external environment. In **Figure 1**, these changes are represented by the environment block. On the other hand, it is assumed that the motor command also drives the forward model which simulates the dynamics of the body. The actual sensory feedback caused by environmental change and the predicted outcomes are compared following action execution. In the comparator model, the SoA is assumed to be generated only if the sensory feedback and the predicted outcomes match. Any mismatch, i.e., a prediction error, is believed to reduce the SoA (Synofzik et al., 2008; Haggard and Chambon, 2012; Haggard, 2017). Some research assumes that the forward model would be updated to minimize the prediction error (Synofzik et al., 2008; Haggard, 2017). The upward left arrow in **Figure 1** represents this update.
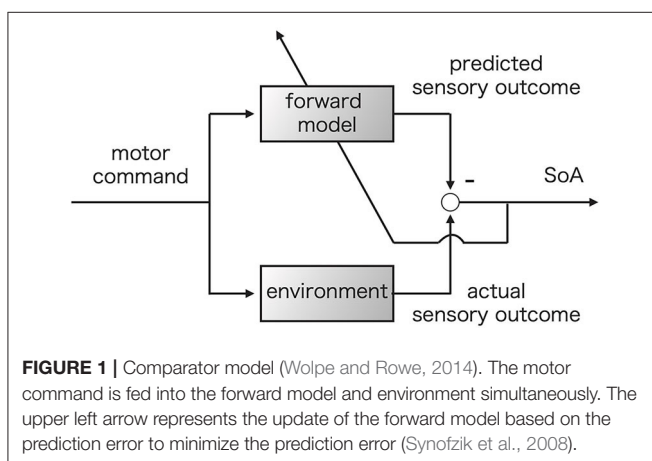
Various types of mathematical modeling approaches have been applied in the cognitive science field (Sun, 2008). One of the advantages of the mathematical modeling approach is that it enables the study of implications of the attractive hypotheses

with deep insight (McClelland, 2009). While this approach would definitely help the SoA studies, to the best of our knowledge, very few studies have proposed a mathematical formulation of the generative process of the SoA except for the work by Moore and Fletcher (2012) and Legaspi and Toyoizumi (2019). Legaspi and Toyoizumi (2019) proposed a concrete and analyzable formula for the "Bayesian cue integration model" proposed by Moore and Fletcher (2012), which had been described only in the abstract formula. In their model, cue integration corresponds to multi-modal sensory integration, e.g., audiovisual integration. As noted by the authors, their hypothesis is that Bayesian cue integration is the "general principle" behind the SoA generation. One of the important contributions by Legaspi and Toyoizumi (2019) was a proposal of a quantitative definition for the SoA. According to our understanding of their work, at least three important issues exist regarding their model to be discussed.

1. No evidence was shown for their assumption that humans solve the cue integration problem by Bayesian inference. Other well-known solution methods exist for the cue integration problem such as stochastic gradient descent (Morency and Baltrušaitis, 2017).
2. The comparator model does not consider multimodality as the essential assumption. The comparator model tells us that we can sense the SoA even from a single-modal sensory signal alone.
3. They proposed the maximum value of the posterior distribution as a possible candidate for the SoA. This proposal does not fit the comparator model, as the candidate value does not correspond to coincidence between the prediction and the actual outcome.

Many researchers have assumed that Bayesian inference is the fundamental and leading principle algorithm in cognitive brain function (Wolpert and Ghahramani, 2000; Körding, 2008; Sanborn et al., 2010; Pouget et al., 2013; Penny, 2015; Legaspi and Toyoizumi, 2019), known as a "Bayesian coding hypothesis" or a "Bayesian brain hypothesis" (Knill and Pouget, 2004). Gershman (2019) summarizes the hypothesis as follows: (1) the brain is equipped with an internal model, i.e., the likelihood function and prior distribution, and (2) Bayes' conditionalization or its approximations work to update the prior distribution. As we show below, Bayesian inference is only one of many solution methods for some inference problems.

One of the important criticisms for the Bayesian hypothesis in cognitive science is that the researchers blindly accept the hypothesis as the standard principle, i.e., the superiority of the Bayesian inference is rarely verified with respect to other inference rules (Bowers and Davis, 2012; Colombo et al., 2018)Pearl (1988, Chapter 2) (Halpern, 2017, Chapter 3). In particular, there are various proofs that Bayesian inference can be derived from general inference rules, i.e., that Bayesian inference is simply a special case of these rules. While the product rule of the probability is definitely the well-known mathematical foundation of Bayes' theorem (Griffiths et al., 2008), there are at least three theoretical foundations known as Kullback's principle of minimum cross-entropy (MINXENT), i.e., the principle of minimum discrimination information (Shore and Johnson, 1980;



**FIGURE 1** | Comparator model (Wolpe and Rowe, 2014). The motor command is fed into the forward model and environment simultaneously. The upper left arrow represents the update of the forward model based on the prediction error to minimize the prediction error (Synofzik et al., 2008).

Shu-Cherng and Tsao, 2001; Rao, 2011; Halpern, 2017, Chapter 3), the information conservation principle (Zellner, 1988, 2002; Soofi, 2000), and the mirror descent algorithm (Warmuth, 2006; Dai et al., 2016). As shown in Fang et al. (1997), the Bayesian inference can be derived from the MINXENT. The other important inference rules such as the maximum entropy principle (MAXENT) can be derived from the MINXENT (Shu-Cherng and Tsao, 2001). Zellner (2002) introduced a concept of the information processing rule, i.e., the rule that transfers prior information and current information into posterior information. Zellner showed that Bayes' theorem can be viewed as the optimal information processing rule under some specific constraint conditions (the information conservation principle). Despite the fact that it is possible to create numerous mathematical models that can potentially explain certain known phenomena, Bayesian inference was not explored in comparison with other candidates. Thus, in this study, we made a comparison of the Bayesian inference with the stochastic gradient descent.

In the first half of this paper, we describe a systematic derivation for the proposed mathematical model of the SoA on the basis of the comparator model , which is related to the prospective generative process of the SoA. We propose a perspective that the SoA is equal to the likelihood value for a given observation. Then, by focusing on the optimization algorithm to maximize the likelihood function, we introduce a perspective that stochastic gradient descent (SGD) (Bishop, 2006; Morency and Baltrušaitis, 2017, Chapter 5.2.4) is the alternative algorithm to the Bayes rule. This is because both algorithms are equally derived from the mirror descent algorithm (Warmuth, 2006; Dai et al., 2016). The difference between the two algorithms exists only in the constraint to the mirror descent algorithm. Bayes' rule uses the Kullback-Leibler divergence (KL divergence), while the SDG uses the Euclid distance as a constraint to the mirror descent algorithm. The stochastic gradient descent method can also be derived from the mirror descent (Bubeck, 2017, Chapter 6.1). If we adopt the Euclid distance as a constraint to the mirror descent algorithm, it follows the SGD. As both the KL divergence and the Euclid distance are essential components in statistical learning, we propose that it is reasonable to compare the SGD with the Bayes' rule. One of the most important ideas in this paper is that these two algorithms predict different behaviors of learners in an online learning problem setting. By focusing on the difference in the observed behavior, we can confirm which algorithm is more suitable for the explanation of the experimental data. In this paper, we do not provide the details of the mirror descent algorithm but refer to (Bubeck, 2017, Chapter 6.1) and (Warmuth, 2006) for further information on this subject.

In the second half of this paper, we validate whether the learning process of human subjects is based on the Bayes' rule or the stochastic gradient descent by performing participant experiments. The experimental task assigned to the subjects was a sense of agency task (Keio method) in which the subjects were asked to report the perception of the SoA during the repeated sequence of intentional action and the observation of resulting target motion (Maeda et al., 2012, 2013). For each trial in the sense of agency task, the presence or absence of the SoA was reported in response to a question. Our objective in performing the experiments was to measure the degree of the SoA that was gradually changed during the task and to determine whether this was caused by the Bayesian inference or the SGD algorithm. Although it had been reported by previous studies that the SoA adapts over several repeated experiences (Leotti et al., 2015), there are only a few studies that clarify the mechanism behind this adaptation (Legaspi and Toyoizumi, 2019; Di Plinio et al., 2020). We propose an indicator that can be used to test the hypothesis obtained from the proposed mathematical model, and we analyze the gradual changes of the perception of the SoA during the task.

This paper consists of three topics: First, we propose a novel mathematical model for the generative process of the SoA based on the comparator model. Second, we derive a scientifically testable hypothesis from the proposed model. Third, we verify our hypothesis of the generative process of the SoA by participant experiments.

The following is an explanation of the terms used in this paper. The term "statistical learning" is a technical term in cognitive science referring to the ability to extract general rules from a series of observations over time (Santolin and Saffran, 2018), carrying out the implications of some automatic learning processes (Schapiro and Turk-Browne, 2015). The term "forward model" is replaced by the term "predictive distribution" (Wolpert and Ghahramani, 2000) or "statistical model" conditioned by one's own action. As explained in Nguyen-Tuong and Peters (2011, section 2.1.1), the term "forward model" often means a deterministic prediction model. The terms "statistical model" and "predictive distribution" are used to emphasize the aspect of probabilistic models. The differences between these terms are described in the next section.

## 2. MATHEMATICAL MODELS AND HYPOTHESIS

### 2.1. The Sense of Agency and the Learning Process

#### 2.1.1. Likelihood Value as the Sense of Agency

In the following section, we explain our perspective regarding the mathematical formulation. We consider the situation where a human subject selects a specific action $a_k \in \mathcal{A}$ and then observes some realized *i.i.d.*-values $X_k \in \mathcal{X}$. The parameter $k$ indicates the number of trials ($k = 1, \ldots, K$). The value $X_k$ corresponds to the actual sensory outcome at the $k$-th trial. We assume that the human subject possesses the probability density function $p(x|\theta, a)$ to predict the sensory outcomes in the future, which is also called the statistical model. The statistical model is parameterized by the parameter $\theta \in \mathbb{R}^n$.

We focus on two major representations of the parameter: the deterministic manner $\theta = \theta_k$ and the probabilistic manner $p_k(\theta)$. The subject predicts the results of the action using the predictive distribution $p_k(x|a)$. The distribution with deterministic parameter representation is defined as Equation (1):

$$p_k(x|a) := p(x|a, \theta_k), \qquad (1)$$

and that with probabilistic parameter representation is Equation (2):

$$p_k(x|a) := \int p(x|a, \theta) p_k(\theta) d\theta. \tag{2}$$

The SoA is based on the degree of coincidence between the actual sensory outcome and its prediction. Thus, we propose the perspective of the SoA being the likelihood value for a given observation. The likelihood value at the $k$ step is

$$L_k := p_k(x = X_k | a = a_k). \tag{3}$$

Note that $L_k$ is the scalar-valued stochastic variable because of the stochastically generated observations $X_k$. The likelihood value $L_k$ indicates how likely the observed data occur from the viewpoint of the human subject's current predictive distribution for a given action. This means that the likelihood reflects the degree of the discrepancy between the actual observation and the prediction: a higher likelihood represents a lower discrepancy. The comparator model hypothesizes that the SoA is generated based on the discrepancies between the prediction and actual outcomes (**Figure 1**). Thus, we consider the likelihood value as the SoA.

### 2.1.2. Online Learning Algorithms: Bayes' Rule and Stochastic Gradient Descent

Data become available in sequential order in real life. Online learning algorithms are the class of algorithms used to update the predictive distribution every time new sensory data are observed (Hazan, 2016). The predictive distribution is updated by the improvement of $\theta_k$ or $p_k(\theta)$ toward a more precise one in an online manner. As in the standard problem setting, in this study we assume that the distribution is updated to minimize the negative log-likelihood function $L(\theta)$:

$$L_k(\theta) = -\log p(X_k | \theta, a_k). \tag{4}$$

Note that $L_k$ in Equation (3) and $L_k(\theta)$ in Equation (4) are different objects. $L_k(\theta)$ is a function of $\theta$.

This section is structured as follows. First, we will introduce a brief review on a derivation of the Bayes' rule (Equations 5–9). Next, we will introduce the stochastic gradient descent algorithm (Equations 10–11).

We will begin by introducing the derivation that states that the Bayes' rule is an algorithm for the minimization of problems. The fact that the Bayes' rule can be derived as a minimization algorithm is useful to understand that this rule is only one of the options for an optimization algorithm. There are two prominent representations of $\theta$ that can be used to minimize (Equation 4). One is to search for the optimal $\theta$ directly. The other is to search the optimal probability density function $p(\theta)$ which minimizes the expected loss $\int p(\theta)L_k(\theta) d\theta$. The former corresponds to the stochastic gradient descent and the latter corresponds to the Bayes' rule. Equation (5) belongs to the latter, representing an iterative algorithm for the expected loss minimization problem (Beck and Teboulle, 2003):

$$p_{k+1}(\theta) = \arg\min_{p \in \mathcal{P}} \left\{ \int p(\theta) L_k(\theta) d\theta + \beta_k \text{KL}[p(\theta)|p_k(\theta)] \right\}, \tag{5}$$

where $\beta_k > 0$. $\mathcal{P}$ is the set of probability density functions. $\text{KL}[p(\theta)|p_k(\theta)]$ is called the KL divergence, which is a measure of the similarity of two probability distributions. The posterior distribution $p_{k+1}(\theta)$ becomes the balanced solution between the prior $p_k(\theta)$ and the most optimal solution because the KL divergence works as a penalty function. It is notable that Equation (5) is equivalent to Equation (6), which is also known as the normalized exponentiated gradient algorithm, multiplicative weight algorithm, entropic mirror descent algorithm, and so on (Beck and Teboulle, 2003; Shalev-Shwartz, 2011).

$$p_{k+1}(\theta) = \frac{1}{Z} \exp\left(-\beta_k^{-1} L_k(\theta)\right) p_k(\theta), \tag{6}$$

where $Z$ is the normalization factor $Z = \int \exp\left(-\beta_k^{-1} L_k(\theta)\right) p_k(\theta) d\theta$. The equivalence is shown by solving the right-hand side of Equation (5) (Beck and Teboulle, 2003). In Equation (6), prior distribution $p_k(\theta)$ is weighted in responding to the value of objective function $L_k(\theta)$. **Figure 2** shows the process of minimizing some quadratic objective function with Equation (5). As that equation shows, the algorithm finds the posterior distribution near the prior distribution at each step. The probability distribution $p_k(\theta)$ finally converges to the Dirac's delta function which surrounds the smallest value of the objective function.

Bayes' rule can be obtained by substituting the negative log-likelihood function into Equation (6) (Warmuth, 2006).

$$p_{k+1}(\theta | \mathcal{X}) \propto p(X_k | \theta)^{\frac{1}{\beta_k}} p_k(\theta). \tag{7}$$

It is important to note that a similar equation is obtained in an iterative manner:

$$p_{k+1}(\theta | \mathcal{X}_k) \propto p(X_k | \theta)^{\frac{1}{\beta_k}} p_k(\theta | \mathcal{X}_{k-1}) \propto \cdots \propto \prod_{l=1}^{k} p(X_l | \theta)^{\frac{1}{\beta_l}} p(\theta) \tag{8}$$
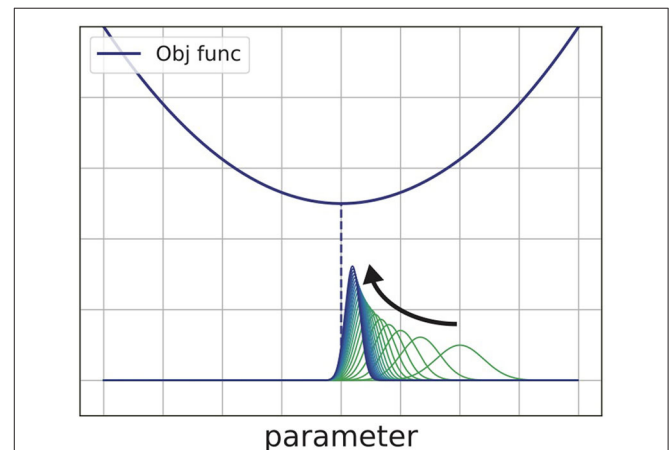
**FIGURE 2 |** Schematic view of the exponentiated gradient method. This method repeatedly updates the distribution of the parameter to find the minimum value of the objective function.

where $\mathcal{X}_k = \{X_1, ..., X_k\}$. In this procedure, the prior distribution is updated with each observation $X_k$. This is the online learning algorithm (Shalev-Shwartz, 2011). In the Bayesian framework, the predictive distribution at each step is defined as

$$p_k(x|\mathcal{X}_{k-1}) = \int p(x|\theta)p_k(\theta|\mathcal{X}_{k-1})d\theta, \qquad (9)$$

so the likelihood of the observation $X_k$ is $p_k(X_k|\mathcal{X}_{k-1})$.

Next, we introduce the stochastic gradient descent (SGD). SGD is also known as an online algorithm that directly updates $\theta \in \mathbb{R}^n$ such that

$$\theta_{k+1} = \theta_k - \eta_k^{-1}g(\theta_k), \qquad (10)$$
$$g(\theta) = -\nabla_\theta \log p(X_k|\theta), \qquad (11)$$

where $\eta_k > 0, \sum_{k=1}^{\infty} \eta_k^{-1} = \infty$, and $\sum_{k=1}^{\infty} \eta_k^{-2} < \infty$ (Bishop, 2006, Chapter 5.2.4). The parameter is updated by each additional observation. In this model, the parameter $\theta_k$ holds the prior knowledge. Since the predictive distribution is $p(x|\theta_k)$, the likelihood is represented as $p(X_k|\theta_k)$. Both the entropic mirror descent Equation (5) and the stochastic gradient descent Equation (10) are equally derived from the mirror descent algorithm (Beck and Teboulle, 2003), and the Bayes' rule can be derived from the entropic mirror descent as described by Equation (7). This indicates that it is natural to compare the Bayes' rule Equation (7) and stochastic gradient descent as expressed in Equations (10) and (11).

Our major argument in this section is that we should test the Bayesian hypothesis against another method such as the stochastic gradient descent. In our view, Bayes' rule is not an oracular theorem but just one of the optimization algorithms for the likelihood maximization problem. We tested the Bayesian hypothesis against the stochastic gradient descent, which is reported below.

## 2.2. Scientifically Testable Hypothesis for the SoA Attribution Task

The type of learning algorithm, for instance Bayes' rule or stochastic gradient descent, has strong effects on the learning curve. As we will see in this section, these methods show the different learning curves of each. Thus, the analysis of the learning curve enables us to test whether the actual learning process is based on the Bayes' rule or the stochastic gradient descent.

We consider the situation in which humans are exposed to an environment in which the consequences of their actual actions are not free from a probabilistic temporal bias. The following description assumes the experimental protocol described in the experimental setup section in the second half of this paper (cf. section 2.3). There exists a temporal bias $x \in \mathbb{R}$ [ms] between the timing of "press button (action)" and "the occurrence of the event (the actual outcome of an action)." We used a one-dimensional Gaussian distribution as a model of human subjects' internal representation regarding the distribution of the temporal bias $x$ for the mathematical simplicity of the data analysis. Specifically,

the subject infers the mean value of the Gaussian distribution and does not infer its standard deviation. Note that this simple assumption requires future validation.

Each subject possesses the statistical model $p(x|\mu)$, which is a one-dimensional Gaussian as expressed in Equations (12) and (13).

$$p(x|\mu) \propto \exp(L(x, \mu)) \qquad (12)$$
$$L(x, \mu) = -\frac{(x - \mu)^2}{2\sigma^2}, \qquad (13)$$

where $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}^+$. The symbol $\theta$ used in the previous section corresponds to the symbol $\mu$ in this section.

The statistical model is used for predicting the temporal bias $x$ between the button press and actual outcome. There are various types of human subjects, some trying to make exact predictions and some trying to make approximate predictions. The individual difference is reflected by the parameter $\sigma$.

The statistical model $p(x|\mu)$ is updated by the improvement of $\mu$ or $p(\mu)$ toward a more precise one. By applying Equations (10) and (11), we derive the learning dynamics of the SGD-type subject in section 2.2.1. By applying Equation (7), we derive the learning dynamics of the Bayesian-type subject in section 2.2.2. Finally, we compare these dynamics and provide the scientifically testable hypothesis in section 2.2.3.

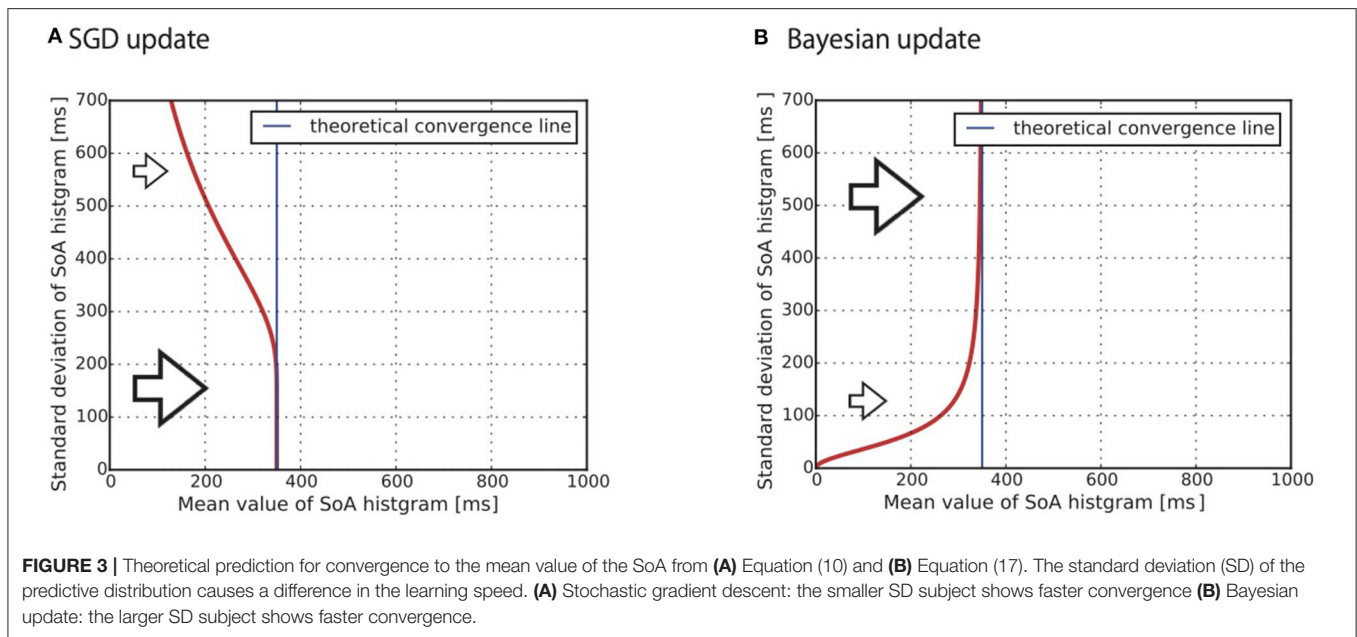### 2.2.1. Learning Dynamics of the SGD-type Subject
Assume that the human subject observes the i.i.d. temporal biases $X_k$ [ms] at each trial $k$.

By applying the SGD algorithm (Equations 10, 11) to the statistical model in Equations (12) and (13), the learning dynamics of the SGD-type subject are derived as Equation (14).

$$\mu_{k+1} = \mu_k + \frac{\eta_k^{-1}(X_k - \mu_k)}{\sigma^2}. \qquad (14)$$

It is important to note that the parameter $\sigma$ determines the learning speed in Equation (14). More precisely, it predicts that a subject with a larger $\sigma$ shows a slower learning speed. **Figure 3A** shows a snapshot of $k$-step of the numerical simulation of Equation (14). The vertical axis shows the $\sigma$ and the horizontal axis shows the $\mu$. As we previously mentioned, each subject is represented by a different parameter $\sigma$. The red line indicates that a subject with a larger $\sigma$ shows the slower convergence. The blue line is the $\mu = \sum_{i=1}^{m} X_i$ where Equation (14) converges regardless of the parameter $\sigma$. Since the true distribution that generates $X_i$ was assumed to be uniform between –300 and 1,000 ms in this simulation, the equation converges to a mean value 350 ms.

As discussed in section 2.1.1, the predictive distribution for the SGD-type subject is modeled as $p(x|a, \mu_k)$ (see Equation 1). If the SGD hypothesis is correct, the predictive distribution $p(x|a, \mu_k)$ behaves slower when its standard deviation is larger. This means that the distributions of the likelihood values, i.e., the histogram of the SoA, also behaves in the same way.

**FIGURE 3 |** Theoretical prediction for convergence to the mean value of the SoA from **(A)** Equation (10) and **(B)** Equation (17). The standard deviation (SD) of the predictive distribution causes a difference in the learning speed. **(A)** Stochastic gradient descent: the smaller SD subject shows faster convergence **(B)** Bayesian update: the larger SD subject shows faster convergence.

### 2.2.2. Learning Dynamics of the Bayesian-Type Subject

Next, we introduce the learning dynamics of the Bayesian-type subject. We assume that this subject holds the prior distribution described by Equations (15) and (16).

$$p_0(\mu) \propto \exp\left(\frac{1}{2\sigma_0^2} G_0(\mu)\right) \quad (15)$$

$$G_0(\mu) = -\mu^2 \quad (16)$$

In this model, the individual difference is reflected by both $\sigma$ and $\sigma_0$.

By applying Equation (7) to the model expressed in Equations (12), (13), (15), and (16), the learning dynamics of the Bayesian-type subject are derived as Equation (17).

$$p_{k+1}(\mu) := p(\mu|\mathcal{X}_{k+1}) \propto \exp\left(\sum_{i=1}^{k} \frac{1}{\beta_i} L(X_i, \mu) + \frac{1}{2\sigma_0^2} G(\mu)\right) \quad (17)$$

The posterior distribution gradually converges to $\lim_{t\to\infty} p(\mu|\mathcal{X}_{t+1}) \propto \exp(\lim_{t\to\infty} \sum_{i=1}^{t} \frac{1}{\beta_i} L(X_i, \mu))$ as the subject obtains observations. The parameters $\sigma$ and $\sigma_0$ determine the learning speed in Equation (17). The way $\sigma$ works is the same as in the SGD case: a larger $\sigma$ results in a slower learning speed. On the other hand, a larger $\sigma_0$ results in a faster learning speed. The smaller $\sigma_0$ becomes, the greater the importance of $G(\mu)$; this makes the posterior distribution difficult to shift from the prior distribution. This is a major property of Bayesian inference that we would like to highlight in this study. Due to this property, the learning dynamics of the Bayesian-type subject can behave as shown in **Figure 3B**. A subject who has a larger standard deviation of the predictive distribution shows faster convergence to the convergence line. Specifically, the larger standard deviation of the prior distribution $\sigma_0$ is the key factor causing this behavior. **Figure 3B** shows a snapshot of the numerical simulation of Equation (17). The vertical axis shows the standard deviation of the predictive distribution. It is important to note that the SGD hypothesis cannot explain the phenomenon that the curve is rising to the right shoulder as shown in **Figure 3B**.

As discussed in section 2.1.1, the predictive distribution for the Bayesian-type subject is modeled as $\int p(x|a, \mu)p_k(\mu)d\mu$ (see Equation 2). The discussion in this section indicates that the distribution of the likelihood values, i.e., the histogram of the SoA, responds faster when its standard deviation is large.

### 2.2.3. Proposed Hypothesis for the SoA Attribution Task

To summarize sections 2.2.1 and 2.2.2, our theory suggests the following:

1. An experimentally measured histogram of the SoA would show the learning behavior in response to the time interval between motor execution and the actual sensory outcome.
2. If the experimentally measured histogram of the SoA with a larger standard deviation were to show faster convergence in the learning processes, it would be supportive evidence for Bayesian learning.

In the second half of this paper, we demonstrate two points through experimentation: (1) that the histogram of the SoA shows the learning behavior, and (2) that the learning behavior is due to Bayesian inference.

### 2.3. Experimental Protocol to Quantify the Sense of Agency

In previous SoA research, there are two types of tasks for measuring the SoA. One is an explicit measure of the SoA in

which subjects verbally report how much they feel a sense of control over external events (an agency attribution task). The other is an implicit measure of the SoA, in which the SoA is evaluated by the "intentional binding effect" (Haggard, 2017), which refers to the subjective binding in time of voluntary actions to their sensory consequences.

In this study, to investigate the adaptation process of a prediction model on the SoA, we used our original agency attribution task (Keio method). A task was established in which human subjects were asked to report their SoA based on their perception of the causal relationship between an intentional action and the visual outcome (Maeda et al., 2012, 2013). The experimental protocol in this study is based exactly on the protocol described in the studies by Maeda et al. (2012, 2013).

The experimental stimuli were presented on a 14-inch computer monitor. A 5-mm square shape appeared from the bottom of the screen and moved straight upwards at a uniform speed (22 [mm/s]). The human subjects were instructed to push a button as quickly as possible when they heard a beep. After they pushed the button, the square jumped 35 mm upwards after a random temporal bias, i.e., at [0, 100, 200, 300, 400, 500, 600, 700, 800, 900, 1,000] [ms]. (**Figure 4A**) Then, they were instructed to respond orally whether they felt that they had caused the square to jump upward as intended by giving a "Yes" or "No" response. A "Yes" response meant that they attributed the jump of the square to their button press, i.e., they felt an SoA during the action. Each condition was conducted 10 times (i.e., 11 conditions × 10 times = 110 trials). In addition to these trials, "event prior to action" (EPA) trials were included in which the square jumped when the beep occurred instead of when the button was pressed. The three EPA conditions were as follows: the square jumped at 100 ms before the beep, at the

time of the beep, or at 100 ms after the beep. They were also instructed to respond whether they felt an SoA during these trials. Each EPA condition was also conducted 10 times (i.e., 3 conditions × 10 times = 30 trials). **Figure 4B** shows the timing at which the temporally biased event occurred. Therefore, we obtained 140 yes/no responses per subject, as represented in **Figure 5A**.

## 2.4. Participants

Twenty-one healthy volunteers were enrolled in this study (8 males and 13 females). Their mean age was 22.0 ± 1.4SD. years old. They were confirmed to have no psychiatric or neurological disorders. This study was approved by the Ethics Committee at Komagino Hospital and the Tokyo University of Agriculture and Technology. All subjects gave written informed consent prior to participation.

## 2.5. Statistical Analysis

As shown in **Figure 5A**, we obtained 140 responses per subject. Our purpose of the analysis was to verify the two hypotheses shown in section 2.2.3:

1. The experimentally measured histogram of the SoA would show the learning behavior in response to the temporal bias.
2. If the experimentally measured histogram of the SoA with a larger standard deviation were to show faster convergence in the learning processes, it would be supportive evidence for the Bayesian learning.

### 2.5.1. Dividing Data by Time-Window

We focused on analyzing the histogram consisting of Yes answers. It is reasonable to assume that the histogram consisting of Yes answers reflects the histogram of likelihood, i.e., the
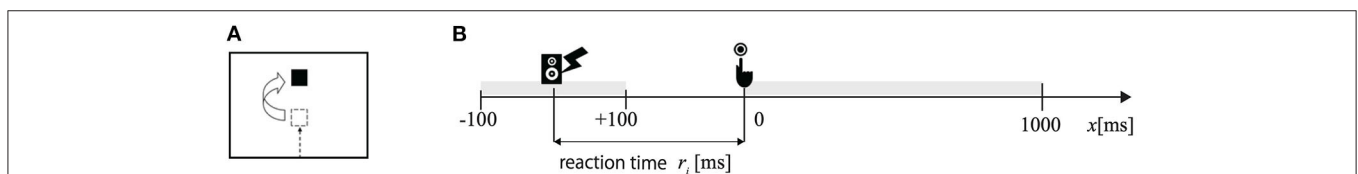


**FIGURE 4 | (A)** Each trial started with a dark computer screen. A square shape appeared at the bottom of the screen and moved straight upwards at a uniform speed (22 [mm/s]). **(B)** The human subjects were instructed to press a button when they heard a beep. When the button was pressed, the square jumped 35 [mm] upward, with various temporal biases. The jump of the square had action-linked conditions and event-prior-to-action (EPA) conditions. In the action-linked conditions, temporal biases were introduced from 0 to 1,000 [ms] in 100-ms increments. In the EPA conditions, the movement of the square was based on the beep and not on the button press, where the movement of the square was programmed to precede the subjects' intentional actions. There were three EPA conditions in which the square jumped 100 [ms] before the beep, at the time of the beep, or at 100 [ms] after the beep. The subjects answered "Yes" or "No" about whether they felt that the square jumped as they intended.
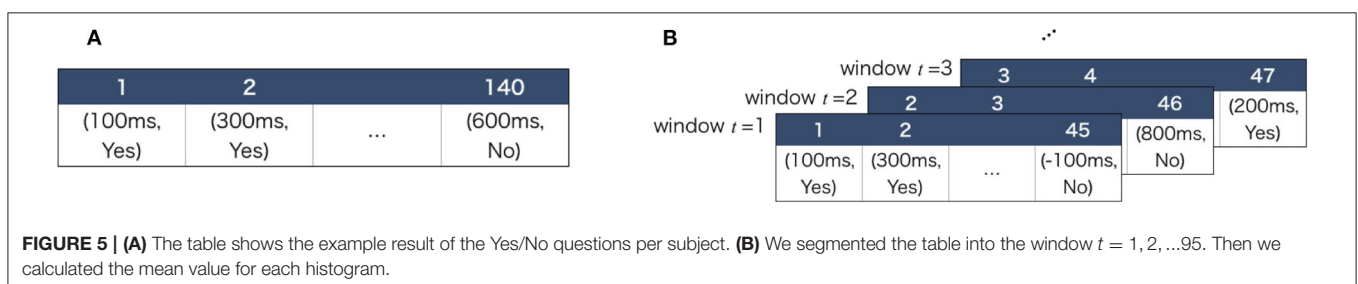


**FIGURE 5 | (A)** The table shows the example result of the Yes/No questions per subject. **(B)** We segmented the table into the window $t = 1, 2, ...95$. Then we calculated the mean value for each histogram.

histogram of the SoA. In the following, the term "histogram" refers to the histogram consisting of Yes answers.

To verify these two hypotheses, we first divided the sequence of YES/NO answers using the moving window which consisted of 45 trials (**Figure 5B**). By dividing the data in this way, it was possible to analyze how the histogram changes over time. We decided on a value of 45 as a criterion to obtain the shape of each of the histograms because a majority of the responses with No would not allow us to analyze the histogram.

### 2.5.2. Gaussian Curve Fit

By dividing the data into the windows, we obtained 95 histograms per subject (**Figure 5B**). As noted in section 2.2, we assumed that the subjects employ a one-dimensional Gaussian distribution as a statistical model. This means that the histograms were also assumed to have the shape of a one-dimensional Gaussian distribution, corresponding to the histogram of the likelihood.

We performed the following procedure to estimate the mean values and standard deviations of the histograms:

First, we quantitatively defined the center of each of the bins of the histogram. Regarding the timing of visual stimuli presentation, while no individual differences occurred for the positive temporal biases (0 to 1,000 [ms]), we needed to consider the individual differences due to the reaction time for the negative temporal biases (cf. **Figure 1**). We calculated the averaged reaction time during the 140 trials for each subject. Then we defined the center of each bin as

$$\mathcal{B}_i := (-r_i - 100, -r_i, -r_i + 100, 0, 100, \ldots, 1000) \quad (18)$$

per subject. Here, $r_i > 0$ represents the $i$-th subject's averaged reaction time during the 140 trials.

Second, we performed the Gaussian curve fitting to the histograms according to the following procedure: We introduce the notation $(\mathcal{B}_i, \mathcal{H}_i^t)$ to represent the $i$-th subject's $t$-th window histogram. $\mathcal{B}_i$ is the 14-dimensional vector that we defined in Equation (18). $\mathcal{H}_i^t \in \mathbb{N}^{14}$ is another 14-dimensional vector that consists of the number of the YES answers corresponding to the temporal biases $\mathcal{B}_i$. We introduce $b_i(j)$ to represent the $j$-th element of the vector $\mathcal{B}_i$ ($j = 1, \ldots, 14$). We also introduce $h_i^t(j)$ to represent the $j$-th element of the vector $\mathcal{H}_i^t$. Although the standard procedure to estimate the Gaussian distribution is to solve the maximum likelihood problem, we cannot apply this approach in this case. This is because the histogram $(\mathcal{B}_i, \mathcal{H}_i^t)$ lacks the data corresponding to the three temporal bias domains $(-\infty, -r_i - 100]$, $[-r_i + 100, 0]$, and $[1000, +\infty)$. Instead, we solved a non-linear least square problem such that

$$Z_i(t), \mu_i(t), \sigma_i(t) = \arg\min_{Z, \mu, \sigma} \sum_{j=1}^{14} (f(b_j; Z, \mu, \sigma) - h_i^t(j))^2, \quad (19)$$

where $f(\cdot; Z, \mu, \sigma)$ is the Gaussian function:

$$f(x; Z, \mu, \sigma) = Z \exp(-\frac{(x - \mu)^2}{2\sigma^2}). \quad (20)$$

We applied the Levenberg-Marquardt algorithm[1] to problem (19) and obtained $(\mu(t), \sigma(t), Z(t))$ which are the estimations of the mean value, the standard deviation, and the scaling factor, respectively.

### 2.5.3. Learning Curve Analysis

By analyzing $(\mu_i(t), \sigma_i(t))_{t=1,\ldots,95}$ obtained from the Gaussian curve fitting, we can clarify how the predictive distribution behaved. The aim of this analysis was to determine whether the learning behavior is occurring for the temporal bias and whether the learning algorithm is the Bayesian or the SGD algorithm. The main idea was to confirm whether the indicators $(\mu_i(t), \sigma_i(t))_{t=1,\ldots,95}$ show a similar movement to either **Figure 3A** or **Figure 3B**.

We focused on two features of **Figure 3**. The first feature is convergence toward the "theoretical convergence line." The red lines gradually converge to the "theoretical convergence line" which is also shown in the figure. The theoretical convergence line is settled as the solution to the KL divergence minimization problem of the true distribution and the statistical model. As we assumed that the statistical model is the Gaussian distribution described by Equations (12) and (13), the line becomes the expectation of the observed values $E_{q(x)}[x] \simeq \frac{1}{K} \sum_{k=1}^{K} X_k$ where $q(x)$ is the true distribution of the observed values $X_k \sim q(x)$ (Wolpert et al., 2013, section 5.1.1). As noted in the previous section 2.5.2, the set of the temporal biases $\mathcal{B}_i$ is affected by the individually different reaction times $r_i(i = 1, \ldots, 21)$. This individual difference makes it difficult to calculate the convergence line $\frac{1}{K} \sum_{k=1}^{K} X_k$. To approximately estimate the line, we assumed the special case that all subjects equally showed the same reaction time $r_i = 200$ [ms]. Under this assumption, we calculated $\frac{1}{K} \sum_{k}^{K} X_k \simeq 350$ [ms]. Note that we adopted the assumption only for the approximate estimation of the line. We did not adopt the assumption for the other types of statistical analysis in this paper. All of the other statistical analyses were achieved using the estimation of the individually different reaction times (cf. section 2.5.2). The approximate estimated line was used only as a guide to confirm whether the data converged to the line. The convergence indicates that the subjects learned the temporal bias.

The second significant difference between the two subfigures **Figures 3A,B** appears in the slope of the red line. The red line shows a negative slope in **Figure 3A**. This reflects the implication of Equation (14) that, under the SGD hypothesis, a subject who has a larger standard deviation of the predictive distribution shows the slower convergence to the convergence line (cf. section 2.2.1). In contrast, the red line shows a positive slope in **Figure 3B**. This reflects the implication of Equation (17) that, under the Bayesian hypothesis, a subject who has a larger standard deviation of the predictive distribution shows faster convergence (cf. section 2.2.2).

We addressed two questions: (1) whether the learning of the temporal bias occurred and (2) whether the learning algorithm is the Bayesian or the SGD algorithm. To answer the first question,

---

[1]Our Python code for this analysis is available to all at https://github.com/yshlo/gf_curvefit.

we analyzed how the center of the scatter plot $(\mu_i(t), \sigma_i(t))$ changed as the number of trials increased. To answer the second question, we investigated whether the slope of the scatter plot was negative or positive, and how the slope changed as the number of trials increased. These analyses were applied to the scatter plots of $(\mu_i(t), \sigma_i(t))$ at the group level. To perform a quantitative analysis, we applied a linear regression on each scatter plot using Equation (21).

$$\sigma = b_t(\mu - a_t) + c_t. \qquad (21)$$

For the linear regression, we applied the Levenberg-Marquardt algorithm to the least square problem. We obtained a time sequence of the estimated parameters $(a_t, b_t, c_t)_{t=1,\dots,95}$, as we have the time sequence of the scatter plots. If the time evolution of the estimated parameter $a_t$ shows convergence to the neighbor of 350 [ms], which is the approximate estimation of the convergence line (see section 2.5.2), it is supportive evidence that the subjects showed learning behavior to the temporal bias. If the estimated slopes $b_t$ are positive at almost all the time steps, it implies that the standard deviation increases as the mean value increases (**Figure 3B**: Bayes' rule). Then, we rejected the SGD algorithm hypothesis (**Figure 3A**). We applied the binomial test for the set of the estimated slopes: the null hypothesis was that the observed proportion of the positive slopes $b_t > 0$ vs the negative slopes $b_t < 0$ is not different from 0.5. If the hypothesis is rejected, it indicates that positivity or negativity of the slope exists. We set the significance level as $\alpha = 0.05$.

## 3. RESULTS

**Figure 6A** shows a distribution of YES answers obtained from a subject (the $i$-th subject) during the $t$-th time-window (cf. **Figure 5B**). We have a total of $95 \times 21$ histograms, as there are 95 time-windows and 21 subjects. **Figure 6A** is simply an example of these histograms. Using a Gaussian curve fit (cf. section 2.5.3), we obtained the mean value and the standard deviation $(\mu_i(t), \sigma_i(t))$ of each histogram.

**Figure 6B** shows the scatter plot consisting of 21 points $(\mu_i(t), \sigma_i(t))_{i=1,\dots,21}$ for one of the 95 time-windows. Each point corresponds to subject. The horizontal axis represents the mean value, and the vertical axis represents the standard deviation. Since we have 95 time-windows, we obtained a time sequence of 95 scatter plots. **Figure 6B** is simply a snapshot of the time sequence. The red line in **Figure 6B** shows the result of the linear regression Equation (21). We obtained a time sequence of the estimated parameters $(a_t, b_t, c_t)_{t=1,\dots,95}$ by applying the linear regression to each of the 95 scatter-plots of $(\mu_i(t), \sigma_i(t))$.

**Figure 6C** shows the time evolution of the parameter $a_t$, i.e., $(a_1, \dots, a_{95})$ (blue line). The red curve in the figure is the exponential regression curve to the time evolution. It mostly converges to approximately 291 [ms] which is close to the average value of the uniform distribution of the temporal bias, which is approximately 350 [ms]. This result is supportive evidence for the existence of the learning process of subjects during the experiment. For the exponential curve regression, the Levenberg-Marquardt algorithm was applied to the least square problem.
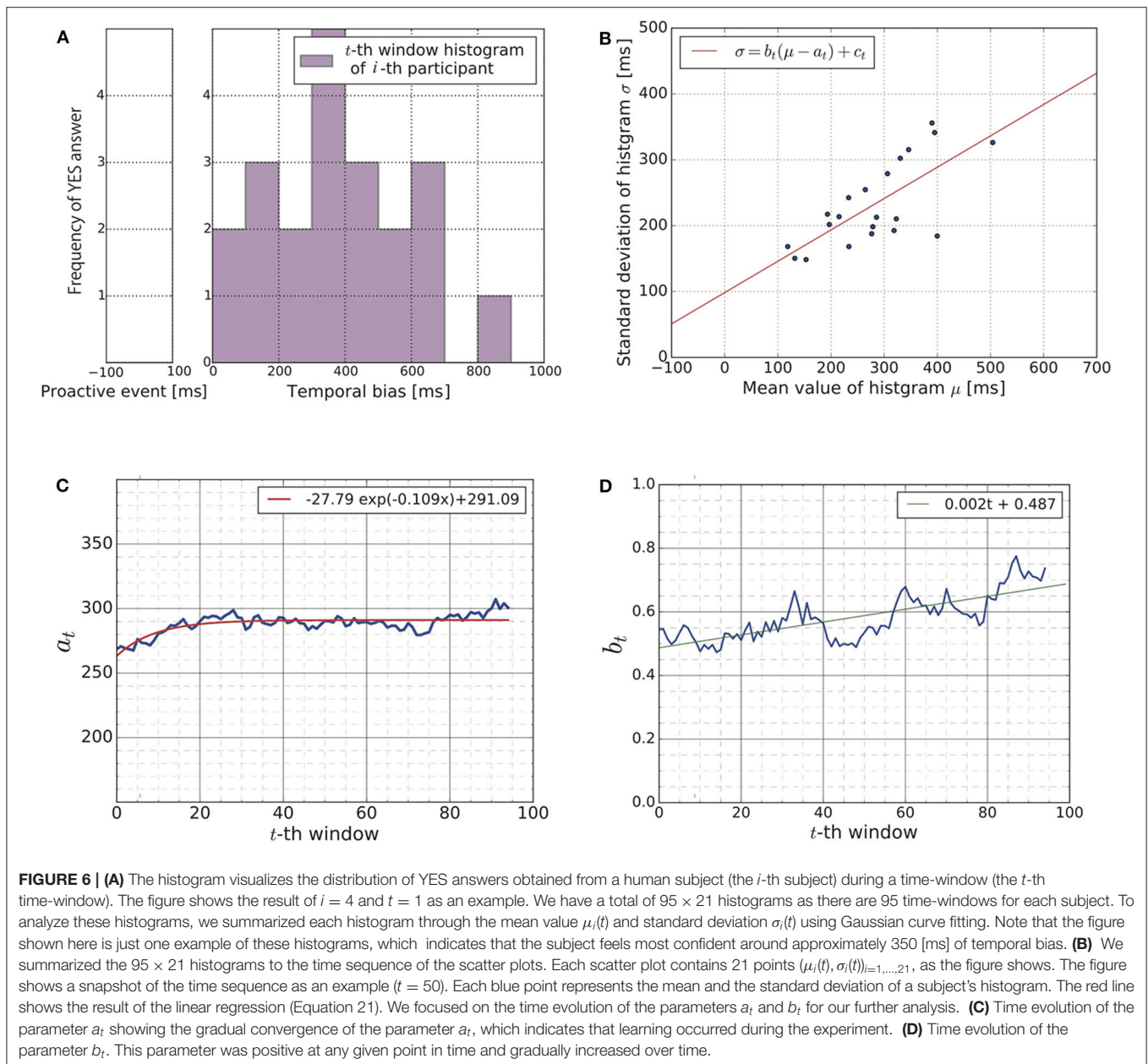
**Figure 6D** shows the time evolution of the parameter $b_t$, i.e., $(b_1, \dots, b_{95})$. As it shows, the parameter $b_t$ was positive and gradually increased across the 95 time-windows. As we explained in section 2.2.3, the positivity of the parameter $b_t$ supports the hypothesis that the updating rule of the subjects is the Bayesian update. As all 95 parameters $(b_1, \dots, b_{95})$ were positive, we applied a binomial test to confirm the statistical significance of the positivity of $b_t$ as explained in section 2.5.3. The null hypothesis was that the observed proportion of the positive slopes $b_t > 0$ vs. the negative slopes $b_t < 0$ was not different from 0.5. The null hypothesis was rejected ($p$-value $< 2.2e^{-16} < 0.05$). The 95 percent confidence interval of the ratio regarding the number of negative slopes and positive slopes was $(0.00, 0.04)$. These results indicate that the updating rule of the subjects should be the Bayesian update. For the linear regression shown in **Figure 6B**, the Leveberg-Marquardt algorithm was applied to the least square problem.

To conclude, our results in **Figure 6** suggest that the learning process of the SoA for temporal bias should be described by the update of the predictive distribution in Bayes' rule rather than the SGD algorithm.

## 4. DISCUSSION

The sense of agency is the experience of subjective awareness regarding the initiation and control of one's own actions. The comparator model indicates the process required to generate the SoA, i.e., a mismatch between the prediction and the consequent event simply causes the loss of the SoA. Therefore, discrepancies in timing between the prediction and event hinder the flow of information integration and thus the induction of the SoA (Wegner and Sparrow, 2004; Wegner et al., 2004). This means that the individual lack ownership of the motor actions if the visual stimuli occur after a short time delay (Saposnik et al., 2010). However, through interactions with the environment, humans have a natural tendency to adapt our control ability and to adapt our internal model to predict the environment. Along this line, the adaptive aspects of the SoA toward changes in the surrounding environment have been debated (Leotti et al., 2015; Haggard, 2017; Legaspi and Toyoizumi, 2019; Di Plinio et al., 2020). As pointed out by Di Plinio et al. (2020), understanding the mechanisms behind the generation and adaptation of SoA is an important emerging topic in this field. While Legaspi and Toyoizumi (2019) proposed the mathematical model for the generation of the SoA, their model includes some issues such as (i) inconsistency with the comparator model and (ii) blind acceptance of the hypothesis that humans solve inference problems by the Bayes' rule. We addressed these problems by building a mathematical model to provide the experimentally testable hypothesis and by verifying the hypothesis through participant experiments in this study.

In summary, our contribution is divided into two parts: (1) a contribution related to the construction of mathematical models and (2) a contribution related to experimental validation. The contributions related to the construction of mathematical models are (i) pointing out that statistical models and learning

**FIGURE 6 | (A)** The histogram visualizes the distribution of YES answers obtained from a human subject (the $i$-th subject) during a time-window (the $t$-th time-window). The figure shows the result of $i = 4$ and $t = 1$ as an example. We have a total of $95 \times 21$ histograms as there are 95 time-windows for each subject. To analyze these histograms, we summarized each histogram through the mean value $\mu_i(t)$ and standard deviation $\sigma_i(t)$ using Gaussian curve fitting. Note that the figure shown here is just one example of these histograms, which indicates that the subject feels most confident around approximately 350 [ms] of temporal bias. **(B)** We summarized the $95 \times 21$ histograms to the time sequence of the scatter plots. Each scatter plot contains 21 points $(\mu_i(t), \sigma_i(t))_{i=1,\ldots,21}$, as the figure shows. The figure shows a snapshot of the time sequence as an example ($t = 50$). Each blue point represents the mean and the standard deviation of a subject's histogram. The red line shows the result of the linear regression (Equation 21). We focused on the time evolution of the parameters $a_t$ and $b_t$ for our further analysis. **(C)** Time evolution of the parameter $a_t$ showing the gradual convergence of the parameter $a_t$, which indicates that learning occurred during the experiment. **(D)** Time evolution of the parameter $b_t$. This parameter was positive at any given point in time and gradually increased over time.

algorithms should be considered separately, (ii) the actual construction of a mathematical model using the SGD algorithm as a comparison object for the mathematical model using the Bayesian estimation algorithm, (iii) formulating the likelihood as the SoA, and (iv) the experimental hypothesis that the two proposed mathematical models behave differently in an experimental task where iterative learning is required. Through the participant experiments, we reported supportive evidence that the human subjects updated their prediction models and that the SoA transformed accordingly: the subjects adapted to the visual stimuli representing the result of the action execution even though with temporal biases (**Figure 6C**). Moreover, our results suggested that the learning algorithm behind the temporal bias adaptation should be described by the Bayes' rule rather than the SGD algorithm (**Figure 6D**).

We call attention to three limitations of the current study. First, our experimental verification of the theoretical hypothesis was done on only one experimental task (the sense of agency task; Keio method). We need further verification using the variety of the sense of agency tasks to indicate the generalizability of our hypothesis. As our experimental task with beep-referenced temporal biases causes individual differences in reaction time, we needed to take individual differences into account (cf. section 2.5.2). In addition, one could also address that the beep-referenced temporal biases and the button-press temporal biases are qualitatively different experiences for the subjects. Thus, further verification using the other SoA tasks is important. It would also be useful to note that the sense of agency task (Keio method) was originally proposed for evaluating the degree of the schizophrenia patient's SoA (Maeda et al., 2012, 2013). If

we could overcome the shortcomings related to the complicated procedures as noted above, we expect our results to be helpful for future schizophrenia research.

Second, we analyzed the experimental data under the assumption that the statistical model for the temporal bias is the Gaussian distribution. While the assumption seems reasonable, note that this is an assumption that simplifies the analysis. We do not exclude all other probability distribution possibilities. In particular, in order to apply our theory to the other tasks (e.g., higher-dimensional prediction problems), the other probability distribution would be suitable. We need to test our theory with other tasks to show that the theory is not limited to the specific probability distribution.

Third, we previously mentioned the limitation associated with this research based on its use of a mathematical model. As Box (1976) has pointed out, we should remember that "all models are wrong" because a model is just an approximation of reality. A traditional guideline for designing a useful model is to illustrate reality as a simple but evocative representation (Box, 1976). Although we followed this guideline to propose the Gaussian model in this paper, a drawback of our approach is that the model was not selected in any quantitative way. Specifically, we did not quantitatively assess the adequacy of the Gaussian assumption and several options for posterior inference, i.e., Markov chain Monte Carlo methods and variational inference methods (Bishop, 2006). Future planned work in this area includes an assessment of our model in comparison to other models using different options according to model selection methods such as information criteria, Bayesian model selection, cross-validation, and norm regularization. (See Ding et al., 2018 for an extensive overview of the selection methods.)

It is useful to discuss the relationship between our model and the model proposed by Legaspi and Toyoizumi (2019) who proposed the Bayesian cue integration model and defined the SoA on the basis of their model. Their definition of the SoA is the maximum value of the posterior distribution. There are at least two important differences between our model and their model. First, our definition of the SoA is the likelihood value that reflects the discrepancies between the prediction and the actual outcome. Second, our model is not intended for the cue integration problem. With respect to the first argument, namely the definition of SoA, it is not that either of the definitions are wrong, but both of them could be right. As Synofzik et al. (2008) pointed out, the SoA is not a single sense but a hierarchical and multilayered sense. Indeed, if we were to adopt a Bayesian model, their definition, i.e., the maximum value of the posterior distribution, could also be defined in our mathematical model. With respect to the second argument, our mathematical framework does not exclude the model for the cue integration problem. Our framework is easy to extend for modeling the cue integration problem if we replace the statistical model appropriately. We hope that the theoretical foundation will continue to be built in the future.

The way in which the probabilistic nature of the SoA can play a key role in the learning process deserves discussion. In its simplest form, the sense of agency originates from the time domain (Gallagher, 2000). In other words, the perception of the

onset of the intentional action and the external event gives rise to the perception of the time interval and the SoA is attributed to the perceived time intervals. However, the perception of the time interval is subject to the inevitable noise of sensing and observation. Thus, observable variables should be treated as stochastic variables rather than as deterministic ones. One of the realistic strategies for predicting an event under the uncertainty of sensory information is to infer its distribution from past experiences and thus make a probabilistic prediction.

Perception of time plays a fundamental role in human perception, cognition and action, which is essential for everyday activities and survival. Rhodes (2018) systematically discussed a variety of mathematical models that can predict that the brain uses temporal expectations to bias perception in a way that stimuli are "regularized," i.e., stimuli look like what has been seen previously. This indicates that the perception of time is subject to various contextual distortions. For example, when observers are presented with various intervals of different lengths and are subsequently instructed to reproduce each interval, they tend to overestimate the duration of short intervals and underestimate that of long intervals (Jazayeri and Shadlen, 2010; Shi et al., 2013). This was reported as a type of "central-tendency" effect, i.e., human subjects migrate their estimates of duration toward the mean of exposed intervals (Rhodes, 2018).

Time is perceived over multiple scales from millisecond scales to interval scales (seconds to minutes), and circadian scales, and these scales should be simulated by different computational and neural mechanisms (Rhodes, 2018). In contrast to the explicit perception of time intervals at a scale of a few seconds, our finding of Bayesian updates in the learning process focused on time perception at a subconscious level (the millisecond scale), which is directly associated with the generation of the SoA. It would be interesting to study the effects of time perception on the SoA at various time scales.

## DATA AVAILABILITY STATEMENT

The datasets presented in this article are not readily available because we need the approval of our institutional ethical committee for the transfer of the dataset to other institutions. Requests to access the datasets should be directed to the corresponding author (shiroyano@ieee.org).

## ETHICS STATEMENT

Studies involving human participants were reviewed and approved by Keio University, Tokyo University of Agriculture and Technology. The patients/participants provided their written informed consent to participate in this study.

## AUTHOR CONTRIBUTIONS

SY mainly designed the theoretical framework of this study. YH and TK supported the work. TM and HI designed the experimental setup. SY and YM performed the analysis of the experimental data. All authors contributed to writing the paper.

## REFERENCES

Beck, A., and Teboulle, M. (2003). Mirror descent and nonlinear projected subgradient methods for convex optimization. *Oper. Res. Lett.* 31, 167–175. doi: 10.1016/S0167-6377(02)00231-6

Bishop, C. M. (2006). *Pattern Recognition and Machine Learning.* New York, NY: Springer.

Blakemore, S. J., Frith, C. D., and Wolpert, D. M. (2001). The cerebellum is involved in predicting the sensory consequences of action. *Neuroreport* 12, 1879–1884. doi: 10.1097/00001756-200107030-00023

Bowers, J. S., and Davis, C. J. (2012). Bayesian just-so stories in psychology and neuroscience. *Psychol. Bull.* 138:389. doi: 10.1037/a0026450

Box, G. E. (1976). Science and statistics. *J. Am. Stat. Assoc.* 71, 791–799. doi: 10.1080/01621459.1976.10480949

Bubeck, S. (2017). Convex optimization: algorithms and complexity. *Found. Trends® Mach. Learn.* 8, 231–357. doi: 10.1561/2200000050

Colombo, M., Elkin, L., and Hartmann, S. (2018). Being realist about Bayes, and the predictive processing theory of mind. *Br. J. Philos. Sci.* doi: 10.1093/bjps/axy059. [Epub ahead of print].

Cornelio Martinez, P. I., De Pirro, S., Vi, C. T., and Subramanian, S. (2017). "Agency in mid-air interfaces," in *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems* (New York, NY: ACM), 2426–2439.

Coyle, D., Moore, J., Ola, P., Paul, K., and Alan, C. F. (2012). "I did that ! Measuring Users ' Experience of Agency in their own Actions," in *Proceedings of the 2012 ACM Annual Conference on Human Factors in Computing Systems CHI 12* (New York, NY).

Dai, B., He, N., Dai, H., and Song, L. (2016). "Provable Bayesian inference via particle mirror descent," in *Artificial Intelligence and Statistics* (Cadiz), 985–994.

Di Plinio, S., Arnò, S., Perrucci, M. G., and Ebisch, S. J. (2020). The evolving sense of agency: context recency and quality modulate the interaction between prospective and retrospective processes. *Conscious. Cogn.* 80:102903. doi: 10.1016/j.concog.2020.102903

Ding, J., Tarokh, V., and Yang, Y. (2018). Model selection techniques: an overview. *IEEE Signal Process. Magaz.* 35, 16–34. doi: 10.1109/MSP.2018.2867638

Fang, S.-C., Rajasekera, J. R., and Tsao, H.-S. J. (1997). *Extensions and Related Results,* Boston, MA: Springer, 285–323.

Gallagher, S. (2000). Philosophical conceptions of the self: implications for cognitive science. *Trends Cogn. Sci.* 4, 14–21. doi: 10.1016/S1364-6613(99)01417-5

Gershman, S. J. (2019). What does the free energy principle tell us about the brain? *arXiv preprint* arXiv:1901.07945.

Griffiths, T. L., Kemp, C., and Tenenbaum, J. B. (2008). "Bayesian models of cognition," in *Cambridge Handbook of Computational Cognitive Modeling,* ed R. Sun (Cambridge: Cambridge University Press), 59–100. doi: 10.1017/CBO9780511816772.006

Haggard, P. (2017). Sense of agency in the human brain. *Nat. Rev. Neurosci.* 18, 196–207. doi: 10.1038/nrn.2017.1

Haggard, P., and Chambon, V. (2012). Sense of agency. *Curr. Biol.* 22, R390–R392. doi: 10.1016/j.cub.2012.02.040

Halpern, J. Y. (2017). *Reasoning About Uncertainty.* Cambridge: MIT Press. doi: 10.7551/mitpress/10951.001.0001

Hazan, E. (2016). Introduction to online convex optimization. *Found. Trends® Optim.* 2, 157–325. doi: 10.1561/9781680831719

Jazayeri, M., and Shadlen, M. N. (2010). Temporal context calibrates interval timing. *Nat. Neurosci.* 13, 1020–1026. doi: 10.1038/nn.2590

Knill, D. C., and Pouget, A. (2004). The bayesian brain: the role of uncertainty in neural coding and computation. *Trends Neurosci.* 27, 712–719. doi: 10.1016/j.tins.2004.10.007

Körding, K. (2008). "Bayesian statistics (with particular focus on the motor system)," in *Encyclopedia of Neuroscience* (Berlin; Heidelberg: Springer-Verlag Berlin Heidelberg), 355–359.

Legaspi, R., and Toyoizumi, T. (2019). A bayesian psychophysics model of sense of agency. *Nat. Commun.* 10, 1–11. doi: 10.1038/s41467-019-12170-0

Leotti, L. A., Cho, C., and Delgado, M. R. (2015). "Chapter: 6: The neural basis underlying the experience of control in the human brain," in *The Sense of Agency,* eds P. Haggard and B. Eitam (Oxford: Oxford University Press), 145–176. doi: 10.1093/acprof:oso/9780190267278.001.0001

Maeda, T., Kato, M., Muramatsu, T., Iwashita, S., Mimura, M., and Kashima, H. (2012). Aberrant sense of agency in patients with schizophrenia: forward and backward over-attribution of temporal causality during intentional action. *Psychiatry Res.* 198, 1–6. doi: 10.1016/j.psychres.2011.10.021

Maeda, T., Takahata, K., Muramatsu, T., Okimura, T., Koreki, A., Iwashita, S., et al. (2013). Reduced sense of agency in chronic schizophrenia with predominant negative symptoms. *Psychiatry Res.* 209, 386–392. doi: 10.1016/j.psychres.2013.04.017

McClelland, J. L. (2009). The place of modeling in cognitive science. *Top. Cogn. Sci.* 1, 11–38. doi: 10.1111/j.1756-8765.2008.01003.x

Minohara, R., Wen, W., Hamasaki, S., Maeda, T., Kato, M., Yamakawa, H., et al. (2016). Strength of intentional effort enhances the sense of agency. *Front. Psychol.* 7:1165. doi: 10.3389/fpsyg.2016.01165

Moore, J. W., and Fletcher, P. C. (2012). Sense of agency in health and disease: a review of cue integration approaches. *Conscious. Cogn.* 21, 59–68. doi: 10.1016/j.concog.2011.08.010

Morency, L.-P., and Baltrušaitis, T. (2017). "Multimodal machine learning: integrating language, vision and speech," in *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics: Tutorial Abstracts* (Vancouver, BC), 3–5.

Nguyen-Tuong, D., and Peters, J. (2011). Model learning for robot control: a survey. *Cogn. Process.* 12, 319–340. doi: 10.1007/s10339-011-0404-1

Pearl, J. (1988). *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference.* Burlington, MA: Morgan Kaufmann. doi: 10.1016/C2009-0-27609-4

Penny, W. (2015). "Bayesian models in neuroscience," in *International Encyclopedia of the Social & Behavioral Sciences: Second Edition,* ed J. D. Wright (Amsterdam: Elsevier), 368–372. doi: 10.1016/B978-0-08-097086-8.56035-8

Pouget, A., Beck, J. M., Ma, W. J., and Latham, P. E. (2013). Probabilistic brains: knowns and unknowns. *Nat. Neurosci.* 16:1170. doi: 10.1038/nn.3495

Rao, C. (2011). "Entropy and cross entropy as diversity and distance measures," in *International Encyclopedia of Statistical Science,* ed M. Lovri? (Berlin; Heidelberg: Springer-Verlag Berlin Heidelberg), 440–446. doi: 10.1007/978-3-642-04898-2

Rhodes, D. (2018). On the distinction between perceived duration and event timing: towards a unified model of time perception. *Timing Time Percept.* 6, 90–123. doi: 10.1163/22134468-20181132

Sanborn, A. N., Griffiths, T. L., and Navarro, D. J. (2010). Rational approximations to rational models: alternative algorithms for category learning. *Psychol. Rev.* 117:1144. doi: 10.1037/a0020511

Santolin, C., and Saffran, J. R. (2018). Constraints on statistical learning across species. *Trends Cogn. Sci.* 22:53. doi: 10.1016/j.tics.2017.10.003

Saposnik, G., Teasell, R., Mamdani, M., Hall, J., McIlroy, W., Cheung, D., et al. (2010). Effectiveness of virtual reality using WII gaming technology in stroke rehabilitation. *Stroke* 41, 1477–1484. doi: 10.1161/STROKEAHA.110.584979

Schapiro, A., and Turk-Browne, N. (2015). "Statistical learning," in *Brain Mapping,* ed A. W. Toga (Waltham, MA: Academic Press), 501–506.

Shalev-Shwartz, S. (2011). Online learning and online convex optimization. *Found. Trends® Mach. Learn.* 4, 107–194. doi: 10.1561/2200000018

Shi, Z., Church, R. M., and Meck, W. H. (2013). Bayesian optimization of time perception. *Trends Cogn. Sci.* 17, 556–564. doi: 10.1016/j.tics.2013.09.009

Shore, J., and Johnson, R. (1980). Axiomatic derivation of the principle of maximum entropy and the principle of minimum cross-entropy. *IEEE Trans. Inform. Theor.* 26, 26–37.

Shu-Cherng, F., and Tsao, H. S. J. (2001). "Entropy optimization: shannon measure of entropy and its properties," in *Encyclopedia of Optimization*, eds C. A. Floudas and P. Pardalos (New York, NY: Springer US), 552–558. doi: 10.1007/0-306-48332-7_119

Soofi, E. S. (2000). Principal information theoretic approaches. *J. Am. Stat. Assoc.* 95, 1349–1353. doi: 10.1080/01621459.2000.10474346

Sun, R. (2008). "Chapter: 1: Introduction to computational cognitive modeling," in *Cambridge Handbook of Computational Psychology*, ed R. Sun (Cambridge: Cambridge University Press), 3–20. doi: 10.1017/CBO9780511816772.003

Synofzik, M., Vosgerau, G., and Newen, A. (2008). Beyond the comparator model: a multifactorial two-step account of agency. *Conscious. Cogn.* 17, 219–239. doi: 10.1016/j.concog.2007.03.010

Warmuth, M. K. (2006). "A bayes rule for density matrices," in *Advances in Neural Information Processing Systems*, 1457–1464.

Wegner, D. M., and Sparrow, B. (2004). *Authorship Processing*. Cambridge, MA: MIT Press, 1201–1209.

Wegner, D. M., Sparrow, B., and Winerman, L. (2004). Vicarious agency: experiencing control over the movements of others. *J. Pers. Soc. Psychol.* 86, 838–848. doi: 10.1037/0022-3514.86.6.838

Wegner, D. M., and Wheatley, T. (1999). Apparent mental causation: sources of the experience of will. *Am. Psychol.* 54, 480.

Wolpe, N., and Rowe, J. B. (2014). Beyond the "urge to move": objective measures for the study of agency in the post-Libet era TL - 8. *Front. Hum. Neurosci.* 8:450. doi: 10.3389/fnhum.2014.00450

Wolpert, D. H., Bieniawski, S. R., and Rajnarayan, D. G. (2013). "Chapter: 4: Probability collectives in optimization," in *Handbook of Statistics*, Vol. 31, eds C. R. Rao, and V. Govindaraju (Amsterdam: Elsevier), 61–99. doi: 10.1016/B978-0-444-53859-8.00004-7

Wolpert, D. M., and Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nat. Neurosci.* 3(Suppl.), 1212–1217. doi: 10.1038/81497

Wolpert, D. M., Ghahramani, Z., and Jordan, M. I. (1995). An internal model for sensorimotor integration. *Science* 269, 1880–1882.

Zellner, A. (1988). Optimal information processing and Bayes's theorem. *Am. Stat.* 42, 278–280.

Zellner, A. (2002). Information processing and Bayesian analysis. *J. Econometr.* 107, 41–50. doi: 10.1016/S0304-4076(01)00112-9