Bayesian Filtering with Multiple Internal Models: Toward a Theory of Social Intelligence

Takuya Isomura
takuya.isomura@riken.jp
Laboratory for Neural Computation and Adaptation, RIKEN Center
for Brain Science, Wako, Saitama 351-0198, Japan

Thomas Parr
thomas.parr.12@ucl.ac.uk
Karl Friston
k.friston@ucl.ac.uk
Wellcome Centre for Human Neuroimaging, Institute of Neurology,
University College London, London, WC1N 3AR, U.K.

To exhibit social intelligence, animals have to recognize whom they are communicating with. One way to make this inference is to select among internal generative models of each conspecific who may be encountered. However, these models also have to be learned via some form of Bayesian belief updating. This induces an interesting problem: When receiving sensory input generated by a particular conspecific, how does an animal know which internal model to update? We consider a theoretical and neurobiologically plausible solution that enables inference and learning of the processes that generate sensory inputs (e.g., listening and understanding) and reproduction of those inputs (e.g., talking or singing), under multiple generative models. This is based on recent advances in theoretical neurobiology—namely, active inference and post hoc (online) Bayesian model selection. In brief, this scheme fits sensory inputs under each generative model. Model parameters are then updated in proportion to the probability that each model could have generated the input (i.e., model evidence). The proposed scheme is demonstrated using a series of (real zebra finch) birdsongs, where each song is generated by several different birds. The scheme is implemented using physiologically plausible models of birdsong production. We show that generalized Bayesian filtering, combined with model selection, leads to successful learning across generative models, each possessing different parameters. These results highlight the utility of having multiple internal models when making inferences in social environments with multiple sources of sensory information.

1 Introduction

One of the most notable abilities of biological creatures is their capacity to adapt their behavior to different contexts and environments (i.e., cognitive flexibility) (Mante, Sussillo, Shenoy, & Newsome, 2013; Dajani & Uddin, 2015) through learning. People can learn to call on various responses depending on the situation—for example, independently move the right and left hands when playing an instrument and speak several different languages. Such multitasking abilities are particularly crucial in communication with several people, who each demand subtly different forms of interaction (Taborsky & Oliveira, 2012; Parkinson & Wheatley, 2015). In this kind of situation, one needs to infer who has generated a heard voice—and infer that person's mental state—to respond in an appropriate manner. This is a requirement for exhibiting social intelligence—which usually indicates the ability of organisms to correctly recognize oneself and others—and behave adequately in the social environment with several conspecifics. This is an important challenge in understanding a key aspect of social intelligence. Experimental studies of primates have shown that the volumes of certain brain structures (e.g., the hippocampus) are correlated with the performance of cognitive and social tasks (Reader & Laland, 2002; Shultz & Dunbar, 2010) and that the ability to infer another's intentions increases with brain volume (Devaine et al., 2017). This speaks to a putative strategy for making inferences about several different conspecifics with a plurality of internal models, each associated with a particular of the community or econiche.

This ability of biological creatures contrasts with current notions of artificial general intelligence. The development of a synthetic system as flexible as the biological brain remains a challenge (LeCun, Bengio, & Hinton, 2015; Hassabis, Kumaran, Summerfield, & Botvinick, 2017). Here, we tried to understand how the brain might entertain distinct generative models in a context-sensitive setting. To do this, we focus on a social task, communication through birdsong, in which the conversational partner may change. This induces the dual task of inferring the identity of a conspecific and learning about that conspecific at the same time. Crucially, this learning should be specific to each partner.

To address this problem, we appeal to generalized Bayesian filtering, a corollary of the free-energy principle (Friston, 2008, 2010). We illustrate the behavior of the proposed scheme using artificial birdsongs and natural zebra finch songs. We consider a synthetic (student), whose generative model is based on a physiologically plausible model of birdsong production, and present the student bird with a song generated by one of several (teacher) conspecifics. During the exchange the student bird performs Bayesian model selection (Schwarz, 1978) to decide which teacher generated the heard song. Having accumulated sensory evidence under all

hypotheses or models, the parameters of the generative models are updated in proportion to the evidence for each competing model.

We show that over successive interactions, our student is able to learn the individual characteristics of multiple teachers and recognize them with increasing confidence Finally, possible neurobiological implementations of the proposed scheme are discussed. Despite our emphasis on birdsong, our interest (and expertise) is not in the theoretical neurobiology of songbirds. We use birdsong as a vehicle to introduce a computational perspective on perceptual categorization and learning in communication (of any sort) that inherits from Bayesian model selection. We hope the scheme we showcase may be useful in areas like voice recognition and in other domains of social exchange.

1.1 Concept of Modeling. In formulating the generative model, we have to contend with a mixture of random variables in continuous time (i.e., latent states of each singing bird) and categorical variables (i.e., the identity of the bird) that constitute a perceptual categorization problem. In short, the listening bird (i.e., student) has to make inferences in terms of beliefs over both continuous and discrete random variables in order to recognize who is singing and *what* they are singing. In a general setting, this would call on mixed generative models with a mixture of continuous and discrete states of the sort considered in Roweis and Ghahramani (1999) and, more recently, Friston, Parr, and de Vries (2017). A complementary way of combining categorical and continuous latent states is to work within a continuous generative model that includes switching variables that have a discrete (i.e., categorical) probability distribution, with an accompanying conjugate prior such as the Dirichlet distribution. The most common example of this would be a gaussian mixture model (see Roweis & Ghahramani, 1999, for details).

Heuristically, this means the generative model can be constructed in one of two ways. We can select a singing bird to generate a song, leading to a hierarchical model with a categorical latent variable at the top and a continuous model generating outcomes. Alternatively, we could generate continuous outcomes from all possible birds and then select one to constitute the actual stimulus. In the second (switching variable) case, the categorical variable plays the role of a switch, basically switching from one possible sensory "channel" to another.

In terms of model inversion and belief propagation, both generative models are isomorphic and lead to the same update equations via minimization of variational free energy. However, the way in which the generative models play out in terms of requisite message passing can have different forms. We could use a generative model with a single bird and try to infer which bird was singing (and, implicitly, the parameters of its generative process). Alternatively, the student may entertain all possible teachers "in mind" and then select the best hypothesis or explanation for

the sensory input. This would correspond to the second form of generative model, in which the dynamics are conditioned on the categorical variable (i.e., a student bird predicts songs under all possible hypotheses) and the best explanation is then selected. In this sense, the expectation about the identity of the singing bird acquires two complementary interpretations. In the first formulation, it is the posterior expectation about the bird that has been selected to generate the song. In the second interpretation, it becomes an expectation about the switching variable. This means the student (i.e., listening bird) effectively composes a Bayesian model average over all hypotheses (i.e., singing birds) entertained in providing posterior predictions of the song.

We can appeal to both forms when interpreting the results that follow. However, the second interpretation has some interesting aspects from a cognitive neuroscience perspective. In essence, the gating or selection of top-down predictions complements the gating or selection of ascending prediction errors usually associated with attention (Luck, Woodman, & Vogel, 2000; Green & Bavelier, 2003; Awh, Belopolsky, & Theeuwes, 2012). In other words, selecting (switching to) the best explanation from available hypotheses, when predicting sensory input, becomes a covert form of (mental) action. Examples of such attentional switching can be found in bistable visual illusions (Eagleman, 2001). This is in the sense that descending predictions are contextualized and selected on the basis of higher-order beliefs (i.e., expectations) about the most plausible hypothesis or context in play. The unique aspect of this gating rests on the fact that there are a discrete (categorical) number of competing hypotheses that are mutually exclusive. This is reminiscent of equivalent architectures in motor control (e.g., the MOSAIC architecture) and related mixture of experts (Roweis & Ghahramani, 1999; Lee, Lewicki, & Sejnowski, 2000; Haruno, Wolpert, & Kawato, 2003). In our case, a simple perceptual categorization paradigm mandates a selection among different possible categories and enforces a form of mental action through optimization of an implicit switching variable.

In what follows, we present the results of perceptual learning and inference using this form of model selection or structure learning, predicated on an ensemble or repertoire of generative models (using synthetic birds and real birdsongs). Using this setup, we show that Bayesian model averaging provides a plausible account of how multiple hypotheses can be combined to predict the sensorium, while Bayesian model selection enables perceptual categorization and selective learning. Crucially, all of these unsupervised processes conform to the same normative principle: the minimization of (the path integral of) a variational free energy bound on model evidence.

2 Results

2.1 Multiple Generative Models and Attentional Switching. Organisms continuously infer the causes of their sensations (unconscious

inference and the Bayesian brain hypothesis: Helmholtz, 1925; Knill & Pouget, 2004) and thereby predict what will happen in the immediate future (e.g., predictive coding) (Rao & Ballard, 1999; Friston, 2005). This sort of perceptual inference rests on an internal generative model that expresses beliefs about how sensory inputs are generated, where perceptual inference is formulated as the minimization of surprise or prediction errors. These models typically assume that sensations are generated by latent or hidden (unobservable) causes in the external world. Such causes may themselves be generated by other causes in a hierarchical manner. In the setting of continuous state-space models, hierarchical Bayesian filtering can be used to perform inference under a hierarchical generative model (Friston, 2008; Friston, Trujillo-Barreto, & Daunizeau, 2008). This filtering uses variational message passing to furnish approximate posterior probability (recognition) densities over the hidden states. In what follows, we describe the process-generating sensory inputs. We assume that the same generative structure is used by the brain as an internal generative model; however the brain needs to learn underlying model parameters to infer the values of hidden states (Dayan, Hinton, Neal, & Zemel, 1995; Friston, Kilner, & Harrison, 2006; George & Hawkins, 2009). A detailed description of the generative models used in this study is provided in section 4.

Let us consider a generative model of birdsong. In brief, this model is a deep generative model with two levels, both levels based on attractor dynamics in the form of neural circuits (see section 4; see also Kiebel, Daunizeau, & Friston, 2008, and Friston & Kiebel, 2009, for details). The goal of an agent (student bird) is to learn about and categorize several different birdsongs, and hence reproduce particular songs depending on the currently heard song. Crucially, the state of a (slow) higher attractor, associated with neuronal dynamics in the high vocal center (HVC) in the songbird brain, provides a control parameter for a (fast) attractor at a lower level in the auditory hierarchy. The hidden or latent states of the lower attractor, associated with the robust nucleus of the archistriatum (RA), then drive fluctuations in the amplitude and frequency of birdsong. (For related songbird studies, see Laje & Mindlin, 2002; Long & Fee, 2008; Amador, Perl, Mindlin, & Margoliash, 2013; and Calabrese & Woolley, 2015.)

In our case, we are interested in multiple models (i.e., multiple teachers), each specified by m_i with $i = 1, 2, \ldots \in M$ (see Figure 1, right). Each m_i indicates a specific model structure including certain functional forms and dimensions of latent variables and parameters. These models describe how sensory input (i.e., birdsong) s is generated by a set of latent variables $u^{(i)}$ that include hidden states $x^{(i)}$ and hidden causes $v^{(i)}$. Here, hidden states $x^{(i)}$ are variables whose dynamics are determined by differential equation (as described below), while hidden causes $v^{(i)}$ are variables generated at a higher level, with a probability distribution $p\left(v^{(i)}|m_i\right)$. In other words, hidden states are linked via dynamics within a hierarchical level, while hidden causes link successive hierarchical levels. Note that the bracketed
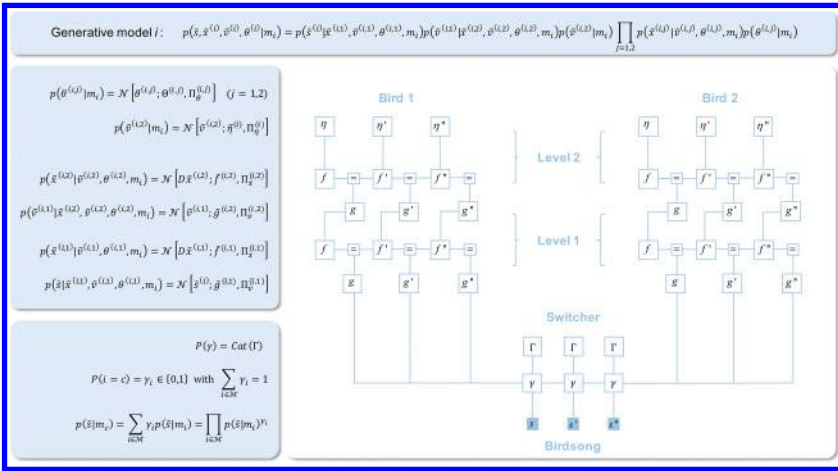
Figure 1: This figure illustrates how random (stochastic) differential equations of motion (see equation 2.1) can be interpreted as a probabilistic generative model. This probabilistic model comprises a joint distribution over latent variables and observable sensory input (upper panel) that can be factorized into the marginal distributions shown in the middle left panel. The large lower right panel depicts two generative models in the form of a normal (Forney) factor graph (Friston, Parr et al., 2017; Forney, 2001; Dauwels, 2007). This graphical form shows that the sensory input (birdsong) may be generated by one of two teacher birds, each represented by its own hierarchical generative model. Here, η at the top of the graph indicates the (prior) expectations of hidden causes. A switcher placed in the center determines which bird generates the sensory input as described in the bottom left panel. The bottom left panel shows that the switcher state γ_i corresponds to the probability of model i being selected, where only γ_c takes a value of one (i.e., m_c is the present model), while the remaining γ_i with $i \neq c$ are zero. Importantly, regardless of the switcher state, all models generate dynamics; for example, $p\left(\tilde{v}^{(i,2)}|m_i\right)$ indicates the probability of the hidden (generalized motion of) cause $\tilde{v}^{(i,2)}$, under the ith model structure m_i, while the selected model is denoted by m_c. The task of our synthetic (student) bird that hears the sensory input (birdsong) is to infer which (teacher) bird generated the song (i.e., to infer γ_i). Having done so, the parameters $\theta^{(i)}$ associated with bird i are updated in proportion to the evidence that bird i was, in fact, singing. (See also section 4 for details.)

superscript (i) indicates they belong to model i (or bird i). These variables are associated with trajectories that are specified in generalized coordinates of motion: $\tilde{s} \equiv (s, s', s'', \ldots)$, where dashes denote time derivatives. The processes that generate birdsong from these latent variables are parameterized by a set of parameters $\theta^{(i)}$. We can represent the generation of birdsong

under model i by the following stochastic differential equations:

$$\tilde{s} = \tilde{g}^{(i,1)}\left(\tilde{x}^{(i,1)}, \tilde{v}^{(i,1)}, \theta^{(i,1)}\right) + \tilde{\omega}^{(i,1)},$$

$$D\tilde{x}^{(i,1)} = \tilde{f}^{(i,1)}\left(\tilde{x}^{(i,1)}, \tilde{v}^{(i,1)}, \theta^{(i,1)}\right) + \tilde{z}^{(i,1)},$$

$$\tilde{v}^{(i,1)} = \tilde{g}^{(i,2)}\left(\tilde{x}^{(i,2)}, \tilde{v}^{(i,2)}, \theta^{(i,2)}\right) + \tilde{\omega}^{(i,2)},$$

$$D\tilde{x}^{(i,2)} = \tilde{f}^{(i,2)}\left(\tilde{x}^{(i,2)}, \tilde{v}^{(i,2)}, \theta^{(i,2)}\right) + \tilde{z}^{(i,2)}, \tag{2.1}$$

In the above, $\tilde{\omega}^{(i,j)}$ and $\tilde{z}^{(i,j)}$ represent random fluctuations. They follow gaussian distributions with mean zero and precision (inverse covariance) of $\Pi_v^{(i,j)}$ and $\Pi_x^{(i,j)}$, respectively. D is a block matrix operator that implements $D\tilde{x} = \tilde{x}' \equiv (x', x'', ...)^T$. The superscript (ij) indicates the jth level of model i (i.e. latent variables at level 2 generate latent variables at level 1 that generate sensory input). The model specified by these equations can be interpreted in terms of a set of probability distributions (see Figure 1 middle left), and their product provides the ith generative model (see Figure 1 top).

In our task design, although every generative model is running simultaneously, only the signal generated by a specific bird is selected as the sensory input (i.e., a teacher song) that the student can actually hear—as an analogy to social communication with several distinct conspecifics. This selection is controlled by a switcher (see also Figure 1 right). Suppose the currently selected model is indexed by c. We represent the switcher state by a set of binary variables $P(i = c) = \gamma_i \in \{0, 1\}$ where only $\gamma_c = 1$, while the remaining variables are zero to ensure $\sum_{i \in \mathcal{M}} \gamma_i = 1$. Note that γ_i indicates the probability of model i being selected but takes only either 0 or 1 by design. When this switching process is used, the probability of sensory input is $p(\tilde{s}|m_c) = E_{P(i=c)}\left[p(\tilde{s}|m_i)\right] = \sum_{i \in \mathcal{M}} \gamma_i p(\tilde{s}|m_i)$, where $p(\tilde{s}|m_i)$ is the conditional probability of the sensory input when model i is selected (see Figure 1 bottom left). We suppose that the switcher γ is sampled from a categorical prior distribution $P(\gamma) = Cat(\Gamma)$ for each epoch of singing. In sum all models (m_i) are generating fluctuations in hidden states, while only the output from m_c is selected as the sensory input.

2.2 Update Rules for Inference, Model Selection, and Learning. The inversion of a generative model corresponds to inferring the unknown variables and parameters, which we will treat as perceptual inference and learning respectively. Formally, in variational Bayes, this rests on optimizing an approximate posterior belief over unknown quantities by minimizing the variational free energy (and its path integral) under each model. This comprises three steps, as shown at the top of Figure 2: (1) in the inference step, latent variables under all models are updated over an epoch of birdsong; (2) in the model selection step, a softmax function of variational free action,
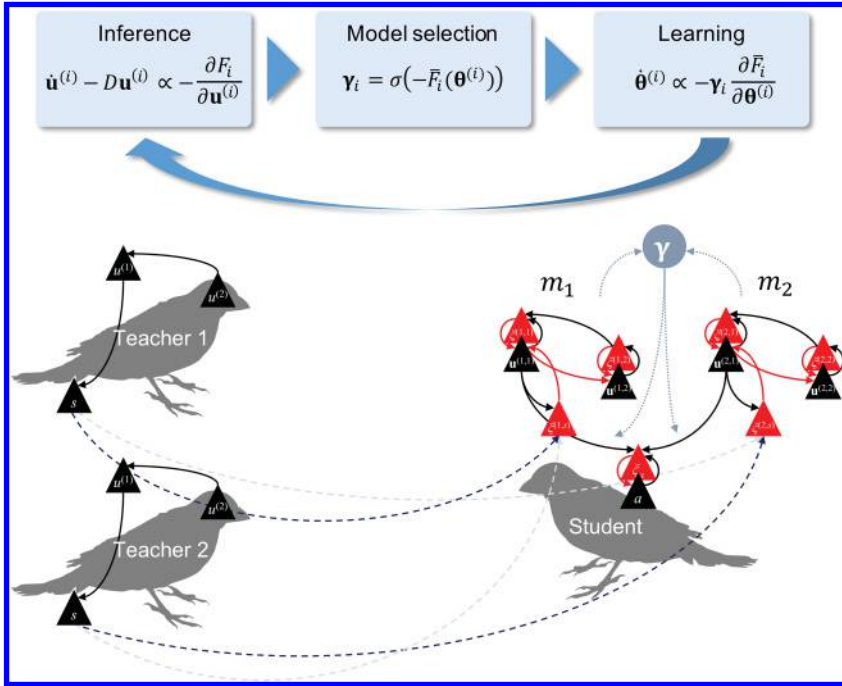
Figure 2: Schematics illustrating the variational update scheme (top): the models that our synthetic student (right) uses to make inferences about the songs generated by two teachers (left). A flowchart at the top summarizes the inference, model selection, and learning processes that the student must implement. Here, $\sigma\,(\cdot)$ is a softmax (normalized exponential) function that converts the conditional free actions to a model plausibility γ_i, and $\xi^{(i,s)}\xi^{(i,j)}\xi_a$ are error-encoding units that encode between actual and expected error in sensation, hidden states, and action, respectively (see Friston, 2008, for details). Our learning process is weighted by the model plausibility, ensuring that the model most likely to have generated the heard song updates its parameters during learning. See section 4 for further details.

under each model, gives the model posterior (i.e., model evidence); and (3) in the learning step, this posterior plays the role of an adaptive learning rate when updating model parameters using a descent on variational free action. This ensures that only models that are likely to be generating the birdsong (sensory data) are updated, while the remaining models retain their current parameters. In what follows, we derive the associated update rules to illustrate their general form (section 4 for details).

The latent variables and parameters under model i are given by $u^{(i)} \equiv \left(\tilde{x}^{(i,1)}, \tilde{x}^{(i,2)}, \tilde{v}^{(i,1)}, \tilde{v}^{(i,2)}\right)$ and $\theta^{(i)} \equiv \left(\theta^{(i,1)}, \theta^{(i,2)}\right)$, respectively. The

internal energy $U_i\left(\tilde{s}, u^{(i)}, \theta^{(i)}\right) \equiv -\log p\left(\tilde{s}, u^{(i)}|\theta^{(i)}, m_i\right)$ quantifies the amount of (squared) prediction error induced by sensory data for a given generative model i, that is, the likelihood of $(\tilde{s}, u^{(i)})$ when $\theta^{(i)}$ and m_i are given. Using this, the conditional free energy of model i is given by

$$F_i(t) \equiv \mathrm{E}_{q(u^{(i)})q(\theta^{(i)})}\left[U_i\left(\tilde{s}, u^{(i)}, \theta^{(i)}\right) + \log q\left(u^{(i)}\right)\right]$$

$$\approx U_i\left(\tilde{s}, \mathbf{u}^{(i)}, \boldsymbol{\theta}^{(i)}\right) + \text{const.} \tag{2.2}$$

Here, $q\left(u^{(i)}\right)$ and $q\left(\theta^{(i)}\right)$ are approximate posterior (i.e., recognition) densities over the latent variables and parameters of each model. The expression $\mathrm{E}_{q(u^{(i)})q(\theta^{(i)})}[\cdot]$ denotes the expectation over these posterior beliefs, and bold symbols $\mathbf{u}^{(i)}$ and $\boldsymbol{\theta}^{(i)}$ denote their posterior expectations (i.e., the means of $q\left(u^{(i)}\right)$ and $q\left(\theta^{(i)}\right)$), respectively. Thus, $\mathbf{u}^{(i)}$ and $\boldsymbol{\theta}^{(i)}$ are maximum a posteriori estimates of the latent variables and parameters under model i. If we ignore the second-order derivative of U_i, we can express F_i by simply substituting $\mathbf{u}^{(i)}$ and $\boldsymbol{\theta}^{(i)}$ into U_i up to a constant term. In neurobiological process theories, $\mathbf{u}^{(i)}$ and $\boldsymbol{\theta}^{(i)}$ are usually associated with neural activities and synaptic strengths, respectively (Bastos et al., 2012; Friston, FitzGerald, Rigoli, Schwartenbeck, & Pezzulo, 2017).

Inference optimizes the approximate posterior beliefs (expectations) about the latent variables. This can be expressed as a gradient flow in generalized coordinates of motion (noting that the solution satisfies the variational principle of least action):

$$\text{Inference:} \quad \dot{\mathbf{u}}^{(i)} - D\mathbf{u}^{(i)} \propto -\frac{\partial}{\partial \mathbf{u}^{(i)}}F_i(t). \tag{2.3}$$

Special cases of this Bayesian filtering reduce to Kalman filtering. The implicit optimization of $\mathbf{u}^{(i)}$ allows for inference to take place under every model. In addition, our agent needs to infer which model is currently generating its sensory input. This involves minimization of the free action (denoted by a bar) over models, given by

$$\bar{F} \equiv \mathrm{E}_{Q(i=c)}\left[\bar{F}_i\left(\boldsymbol{\theta}^{(i)}\right)\right] + D\left[Q(\gamma)||P(\gamma)\right] + \sum_{i \in \mathcal{M}} \mathcal{D}\left[q\left(\theta^{(i)}\right)||p\left(\theta^{(i)}|m_i\right)\right]. \tag{2.4}$$

Note that the first term is the weighted sum of the path integral of the conditional free energies, where $\bar{F}_i(\boldsymbol{\theta}^{(i)}) \equiv \int_0^T F_i(t)\,dt$ is the conditional free action of model i. In this expression, $Q(i=c) = \gamma_i \in [0, 1]$ with $\sum_{i \in \mathcal{M}} \gamma_i = 1$ denotes the posterior expectation about model i being selected, which is equivalent to the posterior belief about the switcher state $Q(\gamma) = \mathrm{Cat}(\gamma)$. The

second and third terms are complexity terms relating to the switcher and parameters, expressed by Kullback-Leibler divergence (Kullback & Leibler, 1951). When the prior distribution of the switcher state $P(\gamma)$ is the same for each model (i.e., all birds are equally likely), we obtain the posterior expectation of the switcher state γ_i that minimizes the total free action as

$$\text{Model selection:} \quad \gamma_i = \sigma\left(-\bar{F}_i\left(\theta^{(i)}\right)\right). \tag{2.5}$$

This means that the posterior expectation (i.e., evidence) that model i generated the song (denoted by γ_i) can be computed by taking a softmax $\sigma(\cdot)$ (normalized exponential) of the conditional free actions for each model, analogous to a post hoc Bayesian model selection (Friston & Penny, 2011) and a discrete categorical model (Friston, FitzGerald et al., 2017). We also refer to γ_i as the model plausibility since this quantifies how likely model i is to have generated the current sensory input.

Finally, learning entails updating posterior expectations about the parameters $\theta^{(i)}$ to minimize the total free action. Taking the gradient of the total free action with respect to the parameters furnishes the learning update rule. When the prior density of parameters $p\left(\theta^{(i)}|m_i\right)$ is flat for every model (i.e., no prior knowledge about parameters), this optimization is given by the minimization of the conditional free action weighed by γ_i:

$$\text{Learning:} \quad \dot{\theta}^{(i)} \propto -\gamma_i \frac{\partial}{\partial \theta^{(i)}} \bar{F}_i\left(\theta^{(i)}\right). \tag{2.6}$$

The novel aspect of this update rule is the weighting of its learning rate by the model evidence or plausibility. This means that only plausible models will change their parameters, which enables the learning of several different generative models in a (soft) winner-takes-all manner. (Detailed derivations of the above equations are in section 4.)

The posterior distribution of the switch $Q(\gamma)$ can be considered an attentional filter (Luck et al., 2000; Green & Bavelier, 2003; Awh et al., 2012). According to this view, an attended generative model and its associated posterior beliefs correspond to the marginal distributions over models and posteriors (i.e., because the attended model is more plausible than all others, the posteriors conditioned on this model will approximate those obtained through a Bayesian model average over all models). Let $u^!$ and $\theta^!$ be the marginal beliefs over latent variables and parameters, respectively. These may be thought of as Bayesian model averages over each of the internal models. When each model has the same structure and dimensions, these marginals are given by $p\left(\tilde{s}, u^! , \theta^!\right) \equiv \sum_{i \in \mathcal{M}} \gamma_i p\left(\tilde{s}, u^{(i)} = u^!, \theta^{(i)} = \theta^!|m_i\right), q\left(u^!\right) \equiv \sum_{i \in \mathcal{M}} \gamma_i q\left(u^{(i)} = u^!\right)$, and $q\left(\theta^!\right) \equiv \sum_{i \in \mathcal{M}} \gamma_i q\left(\theta^{(i)} = \theta^!\right)$. Thus, our model is formally analogous to a gaussian-mixture-model version of a Bayesian filter. On a more anthropomorphic

note, the marginal beliefs over latent variables $u^!$ and parameters $\theta^!$ are fictive (i.e., they do not exist in the external real world). One could imagine that they underwrite some conscious inference, with several competing generative models (i.e., hypotheses) running at a subpersonal or unconscious level in the brain.

Interestingly the above formulation can be applied to generative models that have different structures and dimensions because there is no direct interaction between generative models and the switcher receives only the output from each generative model. This property may be particularly pertinent for recognizing conspecifics, since conspecifics may not be best modeled using the same generative model structure.

2.3 Demonstrations of Multiple Internal Models Using Artificial and Natural Birdsongs. A birdsong has a hierarchical structure that enables the expression of complicated narratives using a finite set of notes (Suzuki, Wheatcroft, & Griesser, 2016). Young songbirds are known to learn such a song by mimicking adult birds' song (Tchernichovski, Mitra, Lints, & Nottebohm, 2001; Woolley, 2012; Lipkind et al., 2013; Yanagihara & Yazaki-Sugiyama, 2016; Lipkind et al., 2017). Previous studies have developed a songbird model that infers the dynamics of another's song based on a deep (two-layer) generative model (Kiebel et al., 2008; Friston & Kiebel, 2009). Perceptual inference requires an internal model of how the song was generated. However, in a social situation, several birds may produce different songs generated by different brain states (or generative models). In the simulations that follow, we consider a case where two birds (denoted by teacher 1, 2) sang two different songs in turn, as illustrated in Figure 2 (left). A song $s = (s_1, s_2)^T$ is given by a 4 s sequence of a two-dimensional vector, where s_2 and s_1 represent the mode of sound frequency and its power, respectively—analogous to a physiological model of birdsong vocalizations (Laje, Gardner, & Mindlin, 2002; Perl, Arneodo, Amador, Goller, & Mindlin, 2011). Here, we supposed that the generative model had two layers of three-neuron circuits (or circuits comprising three neural populations) for birdsong generation, the so-called Laje-Mindlin style model (Laje & Mindlin, 2002). In preliminary simulations, we confirmed that when a student with a single generative model heard their songs, it was unable to learn either teacher 1 or 2's song (see Figure 6 in appendix A). This is because a single generative model cannot generate two songs. Thus, the student tried to learn a spurious intermediate model of the two songs and failed to learn either.

This limitation can be overcome using a repertoire of generative models (see section 4 for details). We found that a student with two generative models (m_1, m_2; see Figure 2, right) can solve this unsupervised learning problem efficiently. We trained the model by providing two (alternating) teacher songs by updating an unknown parameter of both generative models.
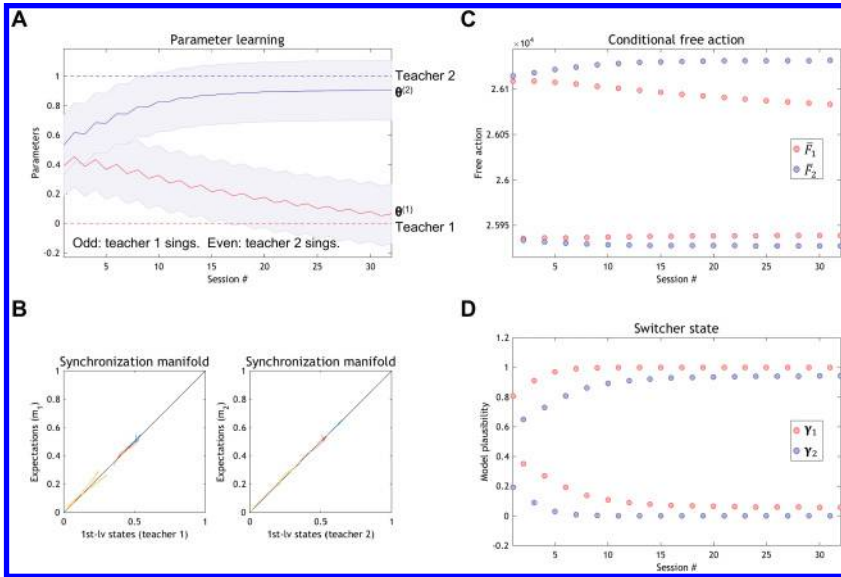
Figure 3: Simulation results when learning two birdsongs using multiple generative models. Teacher bird 1 generated a song in odd sessions, and teacher bird 2 sang in even sessions. At the beginning of each session, the initial latent variables of both teachers were reset to their initial values to ensure they generated quasi-periodic dynamics. The parameters of both teachers were fixed over sessions. A student bird was equipped with two generative models (m_1, m_2). (A) Trajectories of the posterior of a parameter that was optimized. The parameters for m_1 and m_2 (red and blue curves, respectively) were initialized from the middle point (≈ 0.5) and updated according to the variational scheme in the main text. After training, m_1 and m_2's parameter approximated the true parameter value of teacher 1 ($= 0$; red dashed line) and 2 ($= 1$; blue), respectively, reflecting veridical learning. Shaded areas indicate the standard deviation of the posterior density. (B) Comparisons between true (teacher) hidden states and their posterior expectations inferred by the student (left: teacher 1 versus m_1, right: teacher 2 versus m_2). (C) Trajectories of conditional free actions for m_1 (red) and m_2 (blue). When teacher 1 sang (odd sessions), $\bar{F}_1$ was lower than $\bar{F}_2$, and vice versa. A free action difference of about three corresponds to strong evidence for the presence of a song—that is, a log odds ratio of 3 ($= 20$ to 1). (D) Trajectories of model plausibility used for parameter updates. Simulation results with different experimental setup are provided in appendix A.

Posterior densities over parameters (i.e., synaptic strengths) were updated over learning and successfully converged to the true values used in the simulation (see Figure 3A). As a result, the student was able to make perceptual inferences about latent states generating both songs (see Figure 3B). In each

session of training, the free action (i.e., average free energy) was computed for both internal models. The trajectories of free action evince a process of specialization, where each model becomes an expert for one of two songs (see Figure 3C). At the beginning of each exposure, the probability of each model was around 0.5, which led to parameter updates in both models (see Figure 3D). Following learning, the difference in model plausibility became significantly larger—and only the most likely model updated its parameter following the appropriate song. In Figure 3, the hidden states of teachers were reset at the beginning of each session, which made the song sequence periodic and easy to learn. When the hidden states of teachers were not reset, the song sequence became chaotic and was more difficult to learn. However, even in this chaotic case, our model successfully learned from two distinct teachers (see Figure 7 in appendix A).

We next provided six distinct natural zebra finch songs to our model to see if it could learn and recognize six different teachers (see Figure 4; see also section 4 for details). Here we assumed a generative model with realistic song generation capacity comprising two layers of four-neuron circuits (or circuits comprising four neural populations), based on the Laje-Mindlin-style model (see section 4). The posteriors of the parameters of the student's internal models were randomized. However, for simplicity, the posteriors over hidden states at time $t = 0$, and the time constants of the differential equations, were optimized a priori (i.e., initialized) to be consistent with one of the six teacher songs. Before training, we tested the responses of the student to the teacher songs as a reference (movie 1; see appendix B).

A student bird with six internal models inferred latent states (with a small update rate) and calculated the accompanying free energy and model evidence (see Figures 4A and 4B). After exposure, the student generated a song to predict (or imitate) the current teacher song by running the generative models in a forward or active mode. In this mode, the bird reproduces its predicted sensory input based on a Bayesian model average of the dynamics generating a particular song. This Bayesian model average is the mixture of model-specific predictions weighted by model evidence or plausibility. However, prior to learning, the student could not reproduce the teacher song because it has not yet learned the teacher's parameters and could not categorize the teachers. During training, we randomly provided one of the six teacher songs for 60 sessions (movie 2; see appendix B). The student listened to the song and evaluated model plausibility for each of its six internal models. It then learned (924-dimensional) unknown model parameters, with a learning rate determined by model plausibility, to ensure only plausible models were updated. These parameters controlled a nonlinear (polynomial) mapping from latent states expressing the dynamics of the deep generative models to fluctuations in amplitude and peak frequency of the sensory input.

We found that the student's internal models became progressively specialized for one of six teacher songs (movie 3; see appendix B). After
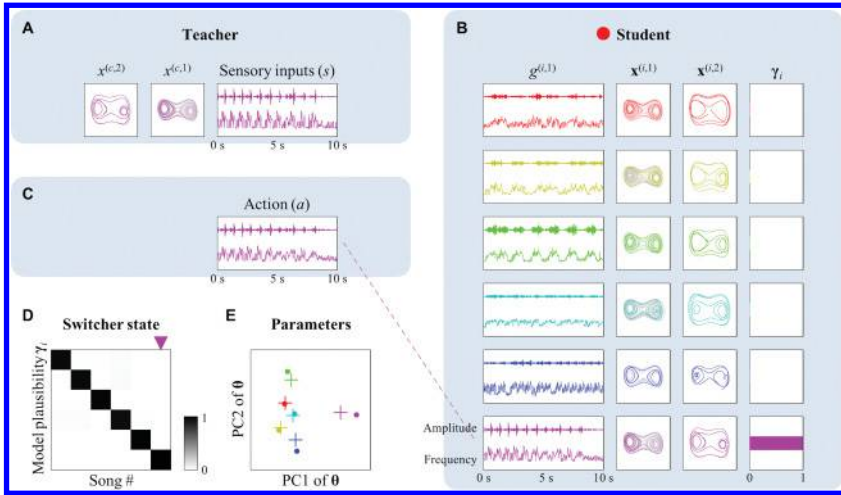
Figure 4: A demonstration of learning six natural zebra finch songs using the multiple generative models. The dynamics of teacher and student states before, during, and after training are provided in appendix B. This figure shows a snapshot of a movie after training. (A) A teacher song (right) and underlying dynamics of hidden states and causes (left and middle; arbitrary scale; they were estimated from the sensory data). Six real zebra finch songs were processed and used as teacher songs (illustrated in six colors). A song is given by a 10s sequence of s, illustrated by the mode of sound frequency and its amplitude (right; arbitrary scale; amplitude is plotted in both positive and negative sides). The currently selected song is indexed by c. (B) Six internal generative models in a student listening to a teacher, making inferences about latent states (centre), model plausibility (right), and the predicted trajectories of sensory input $g^{(i,1)}$ (left). These constitute the expected song sequences (output) under each model. The color of trajectories indicates the song for which each model is specialized. (C) To evaluate prediction capability, a student generates song (action) in the absence of sensory input. Action is given by the average of predictions of the six models weighted by model plausibility, the Bayesian model average. (D) Posterior expectations of the switcher state (or model plausibility) were updated in each model selection step (see equation 2.5). They were initialized from a uniform distribution and converged to a definitive identity matrix, suggesting that each model became specialized for a specific teacher song. (E) Posterior expectations of the parameters of six internal models (circles) and optimal parameters for six teacher songs (plus marks) plotted in a subspace of the first and second principal components (PC1, PC2) of parameter space. During learning, only internal models with high model plausibility enjoy parameter updates (see equation 2.6). The initial parameter values were adjusted in the absence of model selection to ensure their initial values generated an averaged song (see section 4 for details).

learning, only the most plausible model (with veridical parameters) contributed to the Bayesian model average, so that the student could reproduce the teacher songs in a remarkably accurate way (see Figure 4C). These results are particularly pleasing because they also suggest that real songbirds (zebra finches) learn and generate songs (in their RA and HVC) using dynamics with the form we have assumed. Indeed, to compare inferred and true hidden states (and parameters), the real zebra finch songs were learned separately and regenerated under the appropriate model to provide stimuli. Learning success was further confirmed by a specialization of each model for a specific teacher (see Figure 4D) and a convergence of posterior parameter expectations, under each model, to the teacher-specific values (see Figure 4E). These results suggest that the proposed scheme works robustly, even with natural data and a large number of songs.

Finally, we illustrate how inference is affected by either the absence of attentional switching or by a discrepancy between the number of internal models and teacher songs presented (see Figure 5). A standard Bayesian filter, lacking attentional switching, failed to find optimal internal models—to track six teacher songs separately—even when equipped with six internal models. In Figures 5A and 5B, this is evidenced by the absence of free-energy reduction (black line). Conversely, the current scheme, with attentional switching, was able to reduce free energy (red line). This is due to the suppression of the learning rate in implausible models during model inversion. The resulting difference is especially evident when the generative model comprises four-neuron circuits. The mixture model learned six different songs with a high degree of accuracy, as shown in Figure 4, thereby reducing the free energy substantially. When there were equal numbers of internal models and teachers, only one of six internal models was plausible for each session (see Figure 5C). When the number of internal models was greater than that of teachers, several internal models came to represent a teacher song, while the superfluous models were never considered plausible (see Figure 5D), indicating the continued success of the agent in categorizing and learning multiple songs. However, the confidence in these categorizations diminished relative to the correct model. Conversely, when the number of internal models was fewer than that of teachers, each internal model came to represent a mixture of teacher songs, thereby failing to recognize distinct teacher songs (see Figure 5E). Collectively, these findings highlight the potential utility of equipping an adequate number of generative models with attentional selectivity for learning (and inverting) context-sensitive models of the social world.

3 Discussion

The brain may use multiple generative models and select the most plausible explanation for any given context. Findings from comparative neuroanatomy (Reader & Laland, 2002; Shultz & Dunbar, 2010; Devaine et al.,
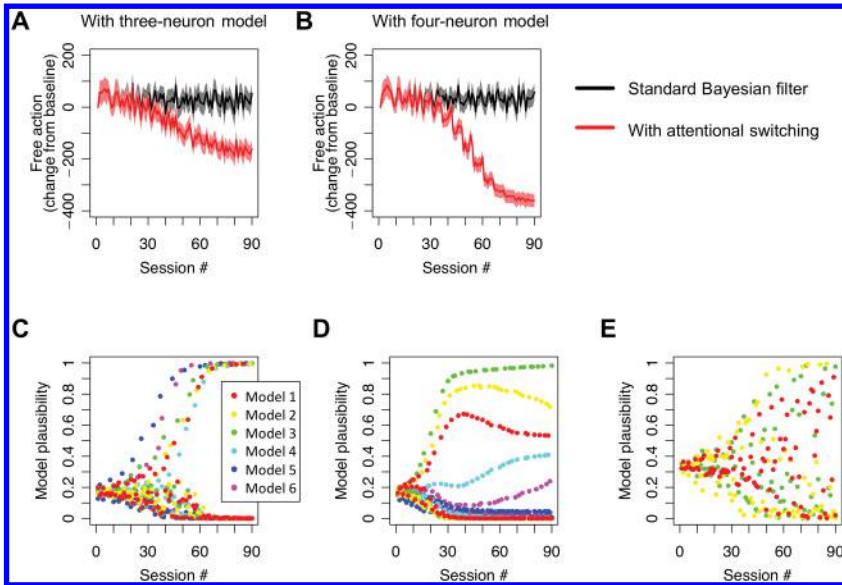
Figure 5: Comparison of learning with multiple internal models in the presence or absence of attentional switching (A, B) and when the number of internal models matches or differs from the teacher songs (C, D, E). (A) This figure illustrates the mean trajectories (lines) of averaged free energy, or free action, in each session and their standard errors (shaded areas). In all cases, a student bird has six internal models. Simulations were conducted 20 times, using different random initial states and song orders. The generative model had two layers of three-neuron circuits for generating birdsong. (B) The generative model had two layers of four-neuron circuits. (C) Trajectories of model plausibility for six internal models. The agent heard one of six teacher songs at random. (D) The same information as panel is displayed, but in the case where the agent received one of three teacher songs. (E) Trajectories of model plausibility for three internal models. In this case, the agent had only three internal models, while hearing one of six teacher songs.

2017) suggest that as the brain becomes larger, it can entertain more hypotheses, or internal models, about how its sensations were caused. This strategy can be used to learn and recognize particular conspecifics in a communication or social setting. A key question here is how the brain separately establishes distinct generative models before it recognizes which model is fit for purpose, and vice versa. In this study, we introduce a novel learning scheme for updating the parameters of the multiple internal models that are themselves being used to filter continuous data. First, several alternative generative models run in parallel to explain the sensory input in terms of inferred latent variables; this enables the free action (under each model) and

associated model plausibility to be evaluated; finally, the parameters of each model are updated with a learning rate that is proportional to the model plausibility. This ensures that only models with high model plausibility or evidence are informed by sensory experience. The proposed scheme allows an agent to establish and maintain several different generative models (or hypotheses) and to perform an adaptive online Bayesian model selection (i.e., switching) of generative models depending on the provided input.

The definition of *social intelligence* varies greatly. The term could be applied broadly to species able to engage in coordinated behaviors with other conspecifics (e.g., swarm behaviors or shoals of fish; Mann & Garnett, 2015). For an account of the sort of inferences required for a creature to "know its place" in this sort of society, Friston, Levin, Sengupta, and Pezzula (2015) illustrate how this can be achieved in the absence of inferences about other individuals in an ensemble. The sort of intelligence we are interested in here is of a more sophisticated sort: the capacity to recognize oneself and others. We are interested in creatures that interact with their conspecifics at an individual level and can tailor their behavior to whomever they interact with. This requires not just (a minimal) theory of mind but a theory of multiple minds and is closer to the sorts of social intelligence thought to be impaired in conditions like autism (Happé & Frith, 1995).

The ideas presented here address a key challenge for social systems: that of disambiguating between and learning about other conspecifics or members of a society. Our hope is that this takes us a step closer to a formal theory of social intelligence. A complete formal theory would entail computational approaches to solving other aspects of social behavior, including those addressing behavioral economic and trust games (Moutoussis, Trujillo-Barreto, El-Deredy, Dolan, & Friston, 2014), and approaches to understanding the optimal depth of recursive sophistication for social interactions (Devaine, Hollard, & Daunizeau, 2014).

The Bayesian filtering or (sensory) evidence accumulation simulated in this letter offers proof of concept that biologically plausible schemes can be used to recognize the source of dynamically rich sensory streams. The simulations show how neuronal-like message passing can solve two key problems: (1) abstracting or deconvolving a time-invariant representation of how fluctuating sensations are generated and (2) disambiguating among alternative sources. The particular message passing used in this study and in a number of previous publications (Friston, Adams, Perrinet, & Breakspear, 2012; Friston & Frith, 2015b; Friston & Herreros, 2016) can be regarded as a generalization of predictive coding that has growing empirical support as a scheme that the brain might use (Kok, Rahnev, Jehee, Lau, & de Lange, 2012; Brodski-Guerniero et al., 2017; Heilbron & Chait, 2017). A review of the evidence for the basic architecture and ideas can be found in several papers (Bastos et al., 2012; Adams, Shipp, & Friston, 2013; Shipp, 2016). A more technical treatment based on message passing on factor graphs can

be found in other publications (Friston, Parr et al., 2017). This letter pursues the biological plausibility of belief updating and, in particular, shows the formal similarities between neuronal message passing required under generative models of both discrete and continuous state spaces.

One might ask if there are alternative schemes that could perform equally well—for example, classification schemes from machine learning (LeCun et al., 2015). This is a potentially important question that would speak to different computational architectures and neurophysiological implementation. However, current machine learning approaches would probably converge on the Bayesian filtering scheme under the deep temporal models used above. This follows from the fact that high-end machine learning schemes use exactly the same (variational free energy) objective function used in Bayesian filtering (and generalized predictive coding). We have in mind here variational autoencoders based on a deep bottleneck architecture, for example (Suh, Chae, Kang, & Choi, 2016). Our model is a mixture model for generalized Bayesian filtering. In this sense, our model can be viewed as a time-domain extension of an autoencoder mixture model (Aljundi, Chakravarty, & Tuytelaars, 2017). A theoretical comparison also supports a close link between learning mechanisms in predictive coding and backpropagation (Whittington & Bogacz, 2017). At the present time, most variational autoencoders do not deal with time-varying data. The implication is that extending current deep learning and variational inference in machine learning to solve the inference problem in a dynamic setting will produce the same scheme as the one used in our simulations.

From a technical perspective, one of the key contributions to the literature of this study is the evaluation of the evidence for competing hypotheses about the sources of sensory input. This evaluation can be cast in terms of Bayesian model selection. From a psychological perspective, this sheds light on our capacity for perceptual categorization, where the underlying selective processes may be cast in terms of attentional selection (Deubel & Schneider, 1996; Itti & Koch, 2001; Bosman et al., 2012). From the perspective of optimal control theory and machine learning, this has clear homologues with selection from mixtures of experts in motor control (Tani & Nolfi, 1999; Haruno et al., 2003) and, indeed, any selective process that involves mutual or lateral inhibition leading to a winner-takes-alll-ike behavior (Zelinsky & Bisley, 2015).

Recent machine learning studies show that task-specific synaptic consolidation can protect the network from forgetting previously learned associations while learning new associations (Kirkpatrick et al., 2017; Zenke, Poole, & Ganguli, 2017). We have focused on disambiguating between, and learning with, multiple internal models. Therefore, we do not consider the explicit protection of previously learned associations. However, a combination of attentional switching and synaptic consolidation would be a potentially interesting extension. We would like to address this issue in the future work.

When an agent encounters an environment that generates data in several possible ways, it can model the environment as either a single generative model (with distinct contextual levels) or multiple generative models. In contrast, when an agent encounters an environment with multiple conspecifics, explaining data with a single model is not straightforward because there are multiple sources of sensory data, calling for a mixture of generative models. In this sense, our model is particularly useful when an agent encounters several different agents in a social context.

Neurobiologically, our learning update rule might be implemented by associative (Hebbian) plasticity modulated by a third factor, a concept that has recently received attention (Pawlak, Wickens, Kirkwood, & Kerr, 2010; Frémaux & Gerstner, 2016; Kuśmierz, Isomura, & Toyoizumi, 2017). While Hebbian plasticity occurs depending on the spike timings of pre- and postsynaptic neurons (Hebb, 1949; Bliss & Lømo, 1973; Markram, Lübke, Frotscher, Sakmann, 1997; Bi & Poo, 1998; Froemke & Dan, 2002; Malenka & Bear, 2004; Feldman, 2012), recent studies have reported that various neuromodulators (Reynolds, Hyland, & Wickens, 2001; Seol et al., 2007; Zhang, Lau, & Bi, 2009; Salgado, Köhr, & Treviño, 2012; Yagishita et al., 2014; Johansen et al., 2014), GABAergic inputs (Paille et al., 2013; Hayama et al., 2013), and glial factors (Ben Achour & Pascual, 2010) can modulate Hebbian plasticity in various ways. Our learning update rule consists of the product of the (conditional) free action gradient providing a Hebbian-like term (see Friston, 2008 for details) and the posterior belief of the switcher state, in which the latter might be implemented by such additional neurobiological factors.

Previous studies have modeled communication between agents in analogy with the mirror neuron system (Kilner, Friston, & Frith, 2007; Friston, Mattout, & Kilner, 2011; Friston & Frith, 2015a, 2015b). These simulations involve two birds that make inferences about each other, converging onto the same internal state and generating the same song. This has been used as a model of hermeneutics, cast in terms of generalized synchrony. Heuristically, both birds come to sing from the same "hymn sheet" and thereby come to "know each other" through knowing themselves. Such a synchronous exchange minimizes the joint free energy of both birds because both birds become mutually predictable. This might be related to an experimental observation that a birdsong can propagate emotional information to another bird and have influence on its behavior (Schwing, Nelson, Wein, & Parsons, 2017). This setup can be generalized when more than two birds are singing the same song. However, when several conspecifics generate different songs (or speak different languages), an agent with a single generative model is no longer fit for purpose. We addressed this limitation by equipping synthetic birds with alternative attractors or hypotheses. Interestingly, it has been reported that some birdsongs display such a learning flexibility—for example, white-crowned sparrows learn multiple songs during the vocal development stage and later switch between learned songs

(Hough, Nelson, & Volman, 2000). In our case, generative models are explicitly decomposed and their learning rates are tuned by an attentional switch, allowing an agent to optimize a specific model for each context. In other words, inference about a particular correspondent's "state of mind" can be modeled by synchronous dynamics during conversation, where one of the listener's internal models converges to an attractor representing the speaker's. In this view, empathic capacity may be quantified by how many attractors the agent can deploy and how well it can optimize each attractor. In future work, it will be interesting to consider the relationship between this conceptual model and recent experimental work that suggests that the human brain uses dissociable activity patterns to separately represent self and other (Ereira, Dolan, & Kurth-Nelson, 2018).

In terms of relating the dynamics of learning and inference to empirical observations, the ability to simulate learning and inference, implicit in our simulations, raises the possibility of using empirical data to constrain the scheme's parameters. In other words, there is, in principle, an opportunity to use the simulations of the birdsong recognition above as an observation model to explain the empirical time course of perceptual categorization, learning, and their neuronal correlates. For example, the time course of learning in Figures 3 to 5 suggests that a unit of time (the number of sessions) would correspond to a range from a few hours to a day, given the results reported in Tchernichovski et al. (2001). In our simulation, the first few sessions exhibited a small free-energy reduction because model plausibility was almost uniform and the associated learning rates were similarly small. After a difference in model evidence emerged, the rate of free-energy reduction reached a peak and then gradually returned to zero. This might correspond to the learning process of songbirds that learn the prototypes of songs early and later learn the details of respective songs (Tchernichovski et al., 2001). In the songbird brain, a population of neurons in the auditory association areas exhibits an experience-dependent selective response to one of several learned songs (Gentner & Margoliash, 2003), suggesting that neurons encode posterior expectations of individual songs based on experience. Moreover, the HVC plays an important role not only in song production but also in the formation of associations between a song and a conspecific that emits that song (Gentner, Hulse, Bentley, & Ball, 2000). Again, this is consistent with our generative model that is used for both generation and recognition. The existence of neurons with a teacher-specific activity has been reported in the higher-level auditory cortex of the songbird, the caudomedial nidopallium (NCM; Yanagihara & Yazaki-Sugiyama, 2016). In our model, the switcher exhibits teacher-specific activity by accumulating model evidence. One can imagine that neurons in the NCM might encode the posterior belief about the switcher state. In addition, the accuracy of song recognition might be bounded by the memory capacity of the neural circuit encoding songs (Gentner, 2004), which is similar to the memory capacity of our model, as determined by the number of

generative models that can be supported by the neuronal infrastructure. In short, these empirical observations support the neurobiological plausibility of our model and speak to the empirical tests.

While we have randomly interspersed the order of presentation from each teacher, it would be interesting to examine the influence of more systematic changes in presentation order. In a social neuroscience context, this could be important in understanding things like multiple language acquisition in children who speak to different family members in different languages. Specifically, there appear to be differences in bilingualism when languages are learned simultaneously (i.e., interspersed like the model here) or successively (Klein, Mok, Chen, & Watkins, 2014). This might imply different mechanisms for the latter compared to the former and require an extension of the generative model employed here.

In terms of the questions and challenges for empirical neuroscience, the picture that emerges from the current solution raises the following considerations: the computational anatomy in Figure 1 communicates with a deep (temporal) architecture in which there are multiple, competing attractor networks in the brain. These effectively compete to explain the sensory data, and their ability to do so determines the rate of perceptual learning. In turn, this means that one would predict distinct autonomous dynamics corresponding to competing hypotheses about the current dynamical form of sensory input. This means that there should be neuronal correlates of distinct pattern generators that are engaged contemporaneously during perceptual synthesis. Second, it suggests a convergence of descending projections to lower (e.g., primary) sensory systems. This brings an interesting and complementary perspective on the divergent neuroanatomy of descending backward connections in cortical hierarchies in the brain (Zeki & Shipp, 1988; Angelucci & Bressloff, 2006). We mean this in the sense that usually one interprets the asymmetry between convergent and divergent zones in terms of things like extra classical receptive field effects, particularly in the visual cortex (Angelucci & Bressloff, 2006). A complementary perspective is that the divergence of descending efferents can also be looked upon as a convergence of descending afferents. This is precisely the architecture described in Figure 1. More interesting, the factor graph representation of neuronal architectures speaks to a selective (Bayesian model selection) modulation of the messages converging on any given lower level. Physiologically, this means that there must be a neuromodulatory mechanism in play that can handle multiple convergent inputs to a postsynaptic neuron or population. In effect, this implies a winner-takes-all-like mechanism at the level of synaptic efficacy, as opposed to synaptic activity. One could speculate about the neurotransmitter basis of this selection process—for example, appealing to the neuromodulatory effects of neurotransmitter systems targeting cholinergic and 5HT receptors (Everitt & Robbins, 1997; Collerton, Perry, & McKeith, 2005; Vossel, Bauer, Bauer, & Mathys, 2014; Hedrick & Waters, 2015; Doya, 2002; Yu & Dayan, 2005; Dayan, 2012)—on

inhibitory interneurons in superficial layers and deep pyramidal cells in deep layers. Many of these speculations have often been rehearsed in relation to the deployment of attention in the context of predictive coding and usually implicate synchronous gain mechanisms via the action of inhibitory interneurons and the current connections with pyramidal cells (Fries, 2005; Womelsdorf & Fries, 2006; Saalmann & Kastner, 2009; Feldman & Friston, 2010; Buschman & Kastner, 2015). Finally, one key message of this theoretical work is that the rate of perceptual learning is determined by the evidence for competing models of sensory input. In principle, this predicts that the rate of sensory learning under ambiguity should be sensitive to the relative probability ascribed to different explanations for sensory input. In turn, this relates to psychophysical studies of perceptual learning under ambiguity with, for example, ambiguous figures or other forms of multistable perception (Tani & Nolfi, 1999; Wurtz, 2008; Hohwy, Paton, & Palmer, 2016).

In summary, we have introduced a novel learning scheme that integrates Bayesian filtering and model selection to learn and deploy multiple generative models. We assumed that a switching variable selects a particular model to generate current sensory input (like switching to a particular radio channel from a repertoire of radio programs), while many alternative generative models are running in the background. To deal with the problem of context-sensitive learning, the proposed scheme calculates the model plausibility (i.e., model evidence) of each generative model based on conditional free actions and updates parameters only in models with a convincing degree of evidence. Our synthetic agents were able to both learn and recognize different artificial and natural birdsongs. These results highlight the potential utility of equipping agents with multiple generative models to make inferences in context-sensitive environments.

4 Methods

The proposed variational update scheme is described in section 2. Further details are provided in this section.

4.1 Generative Model. Formally, the multiple generative models are defined as the following. Hierarchical Bayesian filtering supposes a model that consists of latent variables u (a set of hidden states x and hidden causes v) and parameters θ and infers and learns their approximate probability (recognition) densities. To extend this for a multiple-model version, we express the ith generative model consisting of two layers as m_i with $i \in \mathcal{M} \equiv \{1, 2, 3, \ldots\}$. This m_i indicates a specific model structure including certain forms of functions and dimensions of latent variables and parameters. Let $\tilde{s}$ be sensory inputs (i.e., teacher song) generated by m_i, and $\tilde{x}^{(i,j)}$, $\tilde{v}^{(i,j)}$, and $\theta^{(i,j)}$ be hidden states, hidden causes, and parameters in the jth layer ($j = 1,2$) of m_i, respectively. The tilde over a symbol denotes a set of a variable and its time derivatives $\tilde{s} \equiv (s, s', s'', \ldots)$. Throughout this letter,

i indices the model while j indices the level of layers. The ith generative model is given by equation 2.1. The corresponding probabilities are defined as gaussian distributions and written as

$$\tilde{s} = \tilde{g}^{(i,1)} + \tilde{\omega}^{(i,1)} \iff p\left(\tilde{s}|\tilde{x}^{(i,1)}, \tilde{v}^{(i,1)}, \theta^{(i,1)}, m_i\right) \equiv \mathcal{N}\left[\tilde{s}; \tilde{g}^{(i,1)}, \Pi_s^{(i)}\right],$$

$$D\tilde{x}^{(i,1)} = \tilde{f}^{(i,1)} + \tilde{z}^{(i,1)} \iff p\left(\tilde{x}^{(i,1)}|\tilde{v}^{(i,1)}, \theta^{(i,1)}, m_i\right)$$

$$\equiv \mathcal{N}\left[D\tilde{x}^{(i,1)}; \tilde{f}^{(i,1)}, \Pi_x^{(i,1)}\right],$$

$$\tilde{v}^{(i,1)} = \tilde{g}^{(i,2)} + \tilde{\omega}^{(i,2)} \iff p\left(\tilde{v}^{(i,1)}|\tilde{x}^{(i,2)}, \tilde{v}^{(i,2)}, \theta^{(i,2)}, m_i\right)$$

$$\equiv \mathcal{N}\left[\tilde{v}^{(i,1)}; \tilde{g}^{(i,2)}, \Pi_v^{(i,1)}\right],$$

$$D\tilde{x}^{(i,2)} = \tilde{f}^{(i,2)} + \tilde{z}^{(i,2)} \iff p\left(\tilde{x}^{(i,2)}|\tilde{v}^{(i,2)}, \theta^{(i,2)}, m_i\right)$$

$$\equiv \mathcal{N}\left[D\tilde{x}^{(i,2)}; \tilde{f}^{(i,2)}, \Pi_x^{(i,2)}\right],$$

$$p\left(\tilde{v}^{(i,2)}|m_i\right) \equiv \mathcal{N}\left[\tilde{v}^{(i,2)}; \tilde{\eta}^{(i)}, \Pi_v^{(i,2)}\right],$$

$$p\left(\theta^{(i,j)}|m_i\right) \equiv \mathcal{N}\left[\theta^{(i,j)}; \Theta^{(i,j)}, \Pi_\theta^{(i,j)}\right],$$

$$(j = 1, 2). \tag{4.1}$$

Note that $D\tilde{x}^{(i,j)}$ is the derivative of $\tilde{x}^{(i,j)}$ with respect to time; $\tilde{\omega}^{(i,1)} \sim N[\tilde{\omega}^{(i,1)}; 0, \Pi_s^{(i)}]$, $\tilde{\omega}^{(i,2)} \sim N[\tilde{\omega}^{(i,2)}; 0, \Pi_v^{(i,1)}]$ and $\tilde{z}^{(i,j)} \sim N[\tilde{z}^{(i,j)}; 0, \Pi_x^{(i,j)}]$ are background gaussian noises; $\tilde{g}^{(i,j)} \equiv \tilde{g}^{(i,j)}\left(\tilde{x}^{(i,j)}, \tilde{v}^{(i,j)}, \theta^{(i,j)}\right)$ and $\tilde{f}^{(i,j)} \equiv \tilde{f}^{(i,j)}\left(\tilde{x}^{(i,j)}, \tilde{v}^{(i,j)}, \theta^{(i,j)}\right)$ are arbitrary functions of $\tilde{x}^{(i,j)}$ and $\tilde{v}^{(i,j)}$ parameterized by $\theta^{(i,j)}$; $\Pi_s^{(i)}$, $\Pi_x^{(i,j)}$, and $\Pi_v^{(i,j)}$ are precision matrices; and $p\left(\tilde{v}^{(i,2)}|m_i\right)$ and $p\left(\theta^{(i,j)}|m_i\right)$ are gaussian priors parameterized by the mean and the precision matrix. To simplify notation, we define $u^{(i,j)} \equiv \left(\tilde{x}^{(i,j)}, \tilde{v}^{(i,j)}\right)$ as latent variables, $u^{(i)} \equiv \left(u^{(i,1)}, u^{(i,2)}\right)$ as a set of latent variables in all layers of m_i, and $\theta^{(i)} \equiv \left(\theta^{(i,1)}, \theta^{(i,2)}\right)$ as a set of parameters in all layers of m_i. By multiplying all equations on the right-hand side of equation 4.1, the ith generative model is expressed as

$$p\left(\tilde{s}, u^{(i)}, \theta^{(i)}|m_i\right) = p\left(\tilde{s}|\tilde{x}^{(i,1)}, \tilde{v}^{(i,1)}, \theta^{(i,1)}, m_i\right) p\left(\tilde{v}^{(i,1)}|\tilde{x}^{(i,2)}, \tilde{v}^{(i,2)}, \theta^{(i,2)}, m_i\right)$$

$$\cdot p\left(\tilde{v}^{(i,2)}|m_i\right) \prod_{j=1,2} p\left(\tilde{x}^{(i,j)}|\tilde{v}^{(i,j)}, \theta^{(i,j)}, m_i\right) p\left(\theta^{(i,j)}|m_i\right),$$

$$\tag{4.2}$$

as shown in the top panel in Figure 1.

4.1.1 Sensory Inputs. The sensory input that an agent (i.e., a student bird) actually receives is selected by one of the models in $M \equiv \{1, 2, \ldots\}$, where we index the currently selected model by c. The sensory input is expressed by the sum of the product of the conditional probability of $\tilde{s}$ under each model and the probability of each model being selected:

$$p\left(\tilde{s}|m_c\right) = \mathrm{E}_{P(i=c)}\left[p\left(\tilde{s}|m_i\right)\right] = \sum_{i \in M} P\left(i = c\right) p\left(\tilde{s}|m_i\right)$$

$$= \sum_{i \in M} \gamma_i p\left(\tilde{s}|m_i\right) = \sum_{i \in M} p\left(\tilde{s}|m_i\right)^{\gamma_i}. \tag{4.3}$$

Note that the sufficient statistics $\gamma_i \equiv P\left(i = c\right) \in \{0, 1\}$ satisfies $\sum_{i \in M} \gamma_i = 1$ by design, where only γ_c takes a value of one (and the others zero). In this setting, γ_i plays a role of a switcher that switches which model has generated the current sensory input, while all models are running in the background. The probability of $\gamma = (\gamma_1, \gamma_2, \ldots)$ follows a categorical prior distribution $P\left(\gamma\right) = \mathrm{Cat}\left(\Gamma\right)$.

Interestingly, this definition of the multiple generative models and the switcher is slightly different from supposing a large generative model with switcher-dependent parameters. This idea is rather an assumption that an agent has a set of hypotheses (generative models) about sensory inputs from which to select in a given context. Each generative model is running independently from the others. They interact only via sensory inputs through the model switching scheme, but this interaction does not change their latent variables or parameters. Because of this conditional independence, each model can have different forms and dimensions, although in this work, we assume models based on the same model structure and dimension with different latent variables and parameters.

4.1.2 Free Energy and Free Action. The negative log of $p\left(\tilde{s}|m_c\right)$ denotes the surprise (aka surprisal) associated with sensory inputs. Variational free energy is defined as an upper bound on surprise. In this work, we slightly modify the derivation of variational free energy to include the model selection procedure. First, we show that Bayesian model averaging of the conditional surprises provides an upper bound of surprise. Suppose $Q\left(i = c\right) \equiv \gamma_i \in [0, 1]$ with $\sum_{i \in M} \gamma_i = 1$ is the posterior expectation of the switcher state. This is equivalent to a categorical posterior distribution of γ given by $Q\left(\gamma\right) = \mathrm{Cat}\left(\gamma\right)$. Since model c is selected, the following inequality holds from the nonnegativity of Kullback-Leibler divergence (Kullback & Leibler, 1951):

$$\mathrm{E}_{p(\tilde{s}|m_c)}\left[\log p\left(\tilde{s}|m_c\right) - \sum_{i \in M} \gamma_i \log p\left(\tilde{s}|m_i\right)\right]$$

$$= \sum_{i \in \mathcal{M}} \gamma_i E_{p(\tilde{s}|m_c)} \left[\log p\left(\tilde{s}|m_c\right) - \log p\left(\tilde{s}|m_i\right) \right]$$

$$= \sum_{i \in \mathcal{M}} \gamma_i \mathcal{D}_{\mathrm{KL}} \left[p\left(\tilde{s}|m_c\right) ||p\left(\tilde{s}|m_i\right) \right] \geq 0, \tag{4.4}$$

where $\log p\left(\tilde{s}|m_i\right)$ is a conditional surprise under model i and $\mathcal{D}_{\mathrm{KL}}\left[\cdot||\cdot\right]$ denotes the Kullback-Leibler divergence between two distributions. The expectation over sensory inputs $E_{p(\tilde{s}|m_c)}\left[\cdot\right]$ can be approximated using the time average of surprise. From equation 4.4, we have

$$E_{p(\tilde{s}|m_c)Q(i=c)} \left[-\log p\left(\tilde{s}|m_i\right) \right] = E_{p(\tilde{s}|m_c)} \left[-\sum_{i \in \mathcal{M}} \gamma_i \log p\left(\tilde{s}|m_i\right) \right]$$

$$\geq E_{p(\tilde{s}|m_c)} \left[-\log p\left(\tilde{s}|m_c\right) \right] \tag{4.5}$$

and

$$E_{Q(i=c)} \left[-\int_0^T \log p\left(\tilde{s}|m_i\right)dt \right] = -\sum_{i \in \mathcal{M}} \gamma_i \int_0^T \log p\left(\tilde{s}|m_i\right)dt$$

$$\geq -\int_0^T \log p\left(\tilde{s}|m_c\right)dt, \tag{4.6}$$

where T is the measurement time within a session. Finally, we define total free action (the path integral of free energy) as an upper bound of $E_{Q(i=c)}[-\int_0^T \log p(\tilde{s}|m_i)dt]$:

$$\bar{F} \equiv E_{Q(i=c)} \left[\int_0^T \left(-\log p\left(\tilde{s}|m_i\right) + E_{q(\theta^{(i)})} \left[\mathcal{D}_{\mathrm{KL}} \left[q\left(u^{(i)}\right) ||p\left(u^{(i)}|\tilde{s}, \theta^{(i)}, m_i\right) \right] \right] \right) dt \right]$$

$$+ \mathcal{D}_{\mathrm{KL}} \left[Q\left(\gamma\right) ||P\left(\gamma\right) \right] + \sum_{i \in \mathcal{M}} \mathcal{D}_{\mathrm{KL}} \left[q\left(\theta^{(i)}\right) ||p\left(\theta^{(i)}|m_i\right) \right]$$

$$= E_{Q(i=c)} \left[\bar{F}_i + \log \gamma_i - \log \Gamma_i \right] + \sum_{i \in \mathcal{M}} E_{q(\theta^{(i)})} \left[\log q\left(\theta^{(i)}\right) - \log p\left(\theta^{(i)}|m_i\right) \right]$$

$$\geq E_{Q(i=c)} \left[-\int_0^T \log p\left(\tilde{s}|m_i\right)dt \right]. \tag{4.7}$$

Note that $q\left(u^{(i)}\right)$ and $q\left(\theta^{(i)}\right)$ are the posterior densities over latent variables and parameters under model i, respectively. In this expression, the total free energy is defined as the weighed sum of conditional free actions $\bar{F}_1, \bar{F}_2, \ldots$ plus the Kullback-Leibler divergence (i.e., complexity) associated with the

switcher state and parameters. The free action is defined by

$$\bar{F}_i \equiv \int_0^T \mathrm{E}_{q\left(u^{(i)}\right)q\left(\theta^{(i)}\right)} \left[-\log p \left(\tilde{s}, u^{(i)} | \theta^{(i)}, m_i \right) + \log q \left(u^{(i)} \right) \right] dt$$

$$= \int_0^T F_i \left(t \right) dt, \tag{4.8}$$

where $F_i \left(t \right) \equiv \mathrm{E}_{q\left(u^{(i)}\right)q\left(\theta^{(i)}\right)} \left[-\log p \left(\tilde{s}, u^{(i)} | \theta^{(i)}, m_i \right) + \log q \left(u^{(i)} \right) \right]$ is free energy given model i. The first term of $F_i \left(t \right)$ is the negative log of the generative model (see equation 4.2) divided by $p \left(\theta^{(i)} | m_i \right)$ and is referred to as internal energy under model i: $U_i \left(\tilde{s}, u^{(i)}, \theta^{(i)} \right) \equiv -\log p \left(\tilde{s}, u^{(i)} | \theta^{(i)}, m_i \right)$. Note that $F \left(t \right) \equiv \mathrm{E}_{Q(i=c)} \left[F_i \left(t \right) \right] = \sum_{i \in \mathcal{M}} \gamma_i F_i \left(t \right)$ denotes total free energy.

4.1.3 Posteriors. From the Laplace assumption, the posterior density of latent variables $u^{(i)}$ is approximated as a gaussian distribution $q \left(u^{(i)} \right) = N[u^{(i)}; \mathbf{u}^{(i)}, P_u^{(i)}]$ with an expectation (or mode) vector $\mathbf{u}^{(i)}$ and a precision matrix $P_u^{(i)}$. The posterior density of parameters $\theta^{(i)}$ is approximated as a gaussian distribution $q \left(\theta^{(i)} \right) = N[\theta^{(i)}; \boldsymbol{\theta}^{(i)}, P_\theta^{(i)}]$ with an expectation vector $\boldsymbol{\theta}^{(i)}$ and a precision matrix $P_\theta^{(i)}$. As described above, the posterior distribution of the switcher state (i.e., model plausibility) has been defined as a categorical distribution $Q \left(i = c \right) = \gamma_i$ with $\sum_{i \in \mathcal{M}} \gamma_i = 1$, which is equivalent to $Q \left(\gamma \right) = \mathrm{Cat} \left(\gamma \right)$.

4.2 Variational Update Rules. Updates of the posteriors of the latent variables, the switcher state, and the parameters are conducted in the inference, model selection, and learning steps, respectively. In the simulation, these three steps are repeated in order for each session. In what follows, we formally derive update rules from the minimization of free energy or free action.

4.2.1 Inference (Neural Activity). The optimal $q \left(u^{(i)} \right)$ is obtained by solving the variation of F_i, $\delta F_i = \int \{\mathrm{E}_{q\left(\theta^{(i)}\right)}[U_i \left(\tilde{s}, u^{(i)}, \theta^{(i)} \right)] + \log q \left(u^{(i)} \right) + 1\} \delta q \left(u^{(i)} \right) du^{(i)}$. To satisfy $\delta F_i = 0$, the posterior density should be $q \left(u^{(i)} \right) \propto \exp[-\mathrm{E}_{q\left(\theta^{(i)}\right)}[U_i \left(\tilde{s}, u^{(i)}, \theta^{(i)} \right)]] \approx \exp[-U_i(\tilde{s}, u^{(i)}, \boldsymbol{\theta}^{(i)})]$, where $U_i(\tilde{s}, u^{(i)}, \boldsymbol{\theta}^{(i)})$ is the first-order approximation of the variational energy for latent variables. When $\mathbf{u}^{(i)}$ has been optimized, the path of the mode $\dot{\mathbf{u}}^{(i)}$ should be equal to the mode of the path $Du^{(i)}$, $\dot{\mathbf{u}}^{(i)} = D\mathbf{u}^{(i)}$, in addition to minimizing $U_i(\tilde{s}, u^{(i)}, \boldsymbol{\theta}^{(i)})$ (see Friston, 2008, and Friston et al., 2008, for details). Thus, the gradient descent rule to minimize $F_i \left(t \right)$ with respect to $\mathbf{u}^{(i)}$ is given by

$$\dot{\mathbf{u}}^{(i)} - D\mathbf{u}^{(i)} \propto - \left. \frac{\partial}{\partial u^{(i)}} U_i \left(\tilde{s}, u^{(i)}, \theta^{(i)} \right) \right|_{u^{(i)} = \mathbf{u}^{(i)}}$$

$$= - \frac{\partial}{\partial \mathbf{u}^{(i)}} U_i \left(\tilde{s}, \mathbf{u}^{(i)}, \theta^{(i)} \right) \approx - \frac{\partial}{\partial \mathbf{u}^{(i)}} F_i (t). \tag{4.9}$$

Moreover, $P_u^{(i)}$ that minimizes $F_i(t)$ is given by the Hessian:

$$P_u^{(i)} = \left. \frac{\partial^2}{\partial \left(u^{(i)} \right)^2} U_i \left(\tilde{s}, u^{(i)}, \theta^{(i)} \right) \right|_{u^{(i)} = \mathbf{u}^{(i)}} \approx \frac{\partial^2}{\partial \left(\mathbf{u}^{(i)} \right)^2} F_i (t). \tag{4.10}$$

Hence, we find equation 2.3. Equation 4.9 is usually supposed to be the dynamics of state coding neurons.

4.2.2 Model Selection (Attentional Switch). This step performs online Bayesian model selection analogous to post hoc Bayesian model selection (Friston & Penny, 2011). Since the switcher state comprises discrete variables, this step is similar to a Markov decision process scheme (Friston, FitzGerald et al., 2017). From equation 4.7, the derivative of $\bar{F}$ with respect to γ_i is given by $\delta \bar{F} = \left(\bar{F}_i \left(\theta^{(i)} \right) + \log \gamma_i - \log \Gamma_i + 1 \right) \delta \gamma_i$. From $\delta \bar{F} = 0$, we find γ_i that minimizes $\bar{F}$ as

$$\gamma_i = \sigma \left(-\bar{F}_i \left(\theta^{(i)} \right) + \log \Gamma_i \right). \tag{4.11}$$

Here $\sigma (\cdot)$ is a softmax function defined by $\sigma (\cdot_i) \equiv \exp (\cdot_i) / \sum_{k \in M} \exp (\cdot_k)$. This γ_i expresses the model plausibility of model i. When $P(\gamma_i)$ is the flat (uniform) prior distribution, equation 4.11 becomes equation 2.5.

4.2.3 Learning (Synaptic Plasticity). Estimation of parameters is based on a conventional gradient descent approach. To satisfy $\delta \bar{F} = \int \{ \gamma_i \int_0^T E_{q(u^{(i)})} [U_i \left(\tilde{s}, u^{(i)}, \theta^{(i)} \right)] dt + \log q \left(\theta^{(i)} \right) - \log p \left(\theta^{(i)} | m_i \right) + 1 \} \delta q \left(\theta^{(i)} \right) d\theta^{(i)} = 0$, from $E_{q(u^{(i)})} [U_i \left(\tilde{s}, u^{(i)}, \theta^{(i)} \right)] \approx U_i \left(\tilde{s}, \mathbf{u}^{(i)}, \theta^{(i)} \right)$, the density should be $q \left(\theta^{(i)} \right) \propto \exp[-\gamma_i \int_0^T U_i \left(\tilde{s}, \mathbf{u}^{(i)}, \theta^{(i)} \right) dt + \log p \left(\theta^{(i)} | m_i \right)]$, where $\int_0^T U_i \left(\tilde{s}, \mathbf{u}^{(i)}, \theta^{(i)} \right) dt$ is the approximate variational action for parameters. The gradient descent rule to minimize $\bar{F}$ with respect to $\theta^{(i)}$ is given by

$$\dot{\theta}^{(i)} \propto - \left. \frac{\partial}{\partial \theta^{(i)}} \left\{ \gamma_i \int_0^T U_i \left(\tilde{s}, \mathbf{u}^{(i)}, \theta^{(i)} \right) dt - \log p \left(\theta^{(i)} | m_i \right) \right\} \right|_{\theta^{(i)} = \theta^{(i)}}$$

$$\approx - \frac{\partial}{\partial \theta^{(i)}} \left(\gamma_i \bar{F}_i \left(\theta^{(i)} \right) - \log p \left(\theta^{(i)} | m_i \right) \right). \tag{4.12}$$

Moreover, $P_\theta^{(i)}$ that minimizes $\bar{F}$ is given by the Hessian:

$$\dot{P}_\theta^{(i)} \propto -P_\theta^{(i)} + \frac{\partial^2}{\partial \left(\theta^{(i)}\right)^2} \left\{ \gamma_i \int_0^T U_i \left(\tilde{s}, \mathbf{u}^{(i)}, \theta^{(i)}\right) dt - \log p \left(\theta^{(i)} | m_i\right) \right\} \Bigg|_{\theta^{(i)} = \boldsymbol{\theta}^{(i)}}$$

$$\approx -P_\theta^{(i)} + \frac{\partial^2}{\partial \left(\boldsymbol{\theta}^{(i)}\right)^2} \left(\gamma_i \bar{F}_i \left(\boldsymbol{\theta}^{(i)}\right) - \log p \left(\boldsymbol{\theta}^{(i)} | m_i\right)\right). \tag{4.13}$$

When $p\left(\theta^{(i)} | m_i\right)$ is the flat (uniform) prior distribution, equation 4.12 becomes equation 2.6.

Accordingly, we obtain posterior beliefs of the latent variables, the switcher state, and the parameters that minimize free action. The difference in learning rate mediated by the model plausibility enables that only the parameters in the most plausible models are updated, while the parameters in the remaining models are maintained in a winner-takes-all manner, whereas the latent variables in all models are updated with a fixed update rate. Therefore, inference occurs for all generative models, while learning occurs only for the most plausible generative models. This mechanism enables the agent to make inferences and learning with several different generative models.

4.2.4 Action. Action a is generated to minimize the total free energy $F(t) = \sum_{i \in \mathcal{M}} \gamma_i F_i(t)$: $\dot{a} \propto -\partial F / \partial a$. In the absence of the external sensory input, action directly induces sensory input, that is, $\tilde{s} = a$. Suppose all internal models use the same precision matrix. In this special case, the optimal action is approximately solved as

$$\dot{a} \propto -\frac{\partial F}{\partial a} \approx \sum_{i \in \mathcal{M}} \gamma_i g^{(i,1)} \left(\mathbf{u}^{(i,1)}, \boldsymbol{\theta}^{(i,1)}\right). \tag{4.14}$$

4.3 Songbird Model. A generative model for birdsong generation is defined as a two-layer hierarchical generative model as mentioned in previous studies (Kiebel et al., 2008; Friston & Kiebel, 2009), in which each layer has three or four hidden states that express biological neural circuits for birdsong generation. For simulation purposes, several different teacher songs and the same number of internal models were used.

4.3.1 For Figure 3. Two teacher songs were generated from two generative models with different parameters. A student was supposed to have two internal models. A generative model with two layers was defined. Each layer has three hidden states $x_1^{(i,j)}, x_2^{(i,j)}, x_3^{(i,j)}$ that recapitulate the Laje-Mindlin-style three-neuron circuit model for birdsong production (Laje &

Mindlin, 2002). Layer 1 has one hidden cause $v_1^{(i,1)}$ and one parameter $\theta_1^{(i,1)}$, while layer 2 has no hidden cause or parameter. Several functions for the generative model were defined as follows:

$$g^{(i,1)} \equiv \begin{pmatrix} x_1^{(i,1)} \\ x_3^{(i,1)} \end{pmatrix},$$

$$f^{(i,1)} \equiv \frac{1}{\tau_1} \left[-\underbrace{\begin{pmatrix} x_1^{(i,1)} \\ x_2^{(i,1)} \\ 4x_3^{(i,1)} \end{pmatrix}}_{\text{intrinsic dynamics}} + \underbrace{\begin{pmatrix} \text{sig}\left(10x_1^{(i,1)} - 10x_2^{(i,1)}\right) \\ \text{sig}\left(8.5x_1^{(i,1)} + 2x_2^{(i,1)} + 2x_3^{(i,1)} - 5.5 - \theta_1^{(i,1)} + 2.7x_v^{(i,1)}\right) \\ 4\text{sig}\left(-20x_2^{(i,1)} + 4x_3^{(i,1)} + 6\right) \end{pmatrix}}_{\text{synaptic input}} \right],$$

$$g^{(i,2)} \equiv \left(x_3^{(i,2)}\right),$$

$$f^{(i,2)} \equiv \frac{1}{\tau_2} \left[-\begin{pmatrix} x_1^{(i,2)} \\ x_2^{(i,2)} \\ 4x_3^{(i,2)} \end{pmatrix} + \begin{pmatrix} \text{sig}\left(10x_1^{(i,2)} - 10x_2^{(i,2)}\right) \\ \text{sig}\left(8.5x_1^{(i,2)} + 2x_2^{(i,2)} + 2x_3^{(i,2)} - 5.5\right) \\ 4\text{sig}\left(-20x_2^{(i,2)} + 4x_3^{(i,2)} + 6\right) \end{pmatrix} \right].$$

$$(4.15)$$

Here, $\text{sig}(x) \equiv 1/(1 + e^{-x})$ is the sigmoid function. Neurobiologically, $x_1^{(i,j)}$ and $x_3^{(i,j)}$ correspond to excitatory neurons, while $x_2^{(i,j)}$ corresponds to an inhibitory neuron. Different teacher songs use different parameter $\theta_1^{(i,1)}$ (0 for teacher 1 and 1 for teacher 2). This parameter was learned by a student without supervision. Training was repeated for 32 sessions. Each session was a 4 s sequence. Time resolution $dt = 1/64$ [s] and the time constants $\tau_1 = 32/3$ [s], $\tau_2 = 128/3$ [s] were used.

4.3.2 *For Figures 4 and 5.* Six natural zebra finch songs were used as teacher songs, and a student that has six internal models was supposed.

These models share the same structure while using different latent variables and parameters. To enable the model to accurately imitate natural zebra finch songs, we extended the original three-neuron Laje-Mindlin model to a four-neuron circuit by adding an inhibitory neuron (see Figure 8 in appendix A). This model was used in Figures 4 and 5B. This addition served to introduce a delay in the attractor, whereas in Figure 5A, the original three-neuron attractor was used for comparison. We also supposed a nonlinear mapping from the four-neuron attractor to the outputs (songs), which afforded the capability to generate (imitate) complex songs. Each layer has four hidden states: $x_1^{(i,j)}$ and $x_4^{(i,j)}$ correspond to excitatory neurons, whereas $x_2^{(i,j)}$ and $x_3^{(i,j)}$ correspond to inhibitory neurons. Layer 1 has two hidden causes $v_1^{(i,1)}$, $v_2^{(i,1)}$ and 924-dimensional parameters (a 2×462 matrix $\theta^{(i,1)}$), while layer 2 has no hidden cause or parameter. As before, several functions for the generative model were defined as follows:

$$
g^{(i,1)} \equiv \begin{pmatrix} \theta_{1,1}^{(i,1)} & \theta_{1,2}^{(i,1)} & \cdots & \theta_{1,462}^{(i,1)} \\ \theta_{2,1}^{(i,1)} & \theta_{2,2}^{(i,1)} & \cdots & \theta_{2,462}^{(i,1)} \end{pmatrix} \begin{pmatrix} 1 \\ u^{(i,1)} \\ u^{(i,1)} \otimes u^{(i,1)} \\ u^{(i,1)} \otimes u^{(i,1)} \otimes u^{(i,1)} \\ u^{(i,1)} \otimes u^{(i,1)} \otimes u^{(i,1)} \otimes u^{(i,1)} \\ u^{(i,1)} \otimes u^{(i,1)} \otimes u^{(i,1)} \otimes u^{(i,1)} \otimes u^{(i,1)} \end{pmatrix},
$$

$$
f^{(i,j)} \equiv \frac{1}{\tau_j} \left[-\rho \underbrace{\begin{pmatrix} x_1^{(i,1)} \\ x_2^{(i,1)} \\ x_3^{(i,1)} \\ x_4^{(i,1)} \end{pmatrix}}_{\text{intrinsic dynamics}} + \underbrace{\begin{pmatrix} -x_2^{(i,j)} \\ \psi\left(x_1^{(i,j)} - \epsilon x_3^{(i,j)}\right) - \epsilon x_3^{(i,j)} \\ -\epsilon x_2^{(i,j)} + x_4^{(i,j)} \\ -\lambda x_3^{(i,j)} \end{pmatrix}}_{\text{synaptic input}} \right], \quad (j = 1, 2),
$$

$$
g^{(i,2)} \equiv \begin{pmatrix} x_1^{(i,2)} \\ x_4^{(i,2)} \end{pmatrix}. \tag{4.16}
$$

Here $u^{(i,1)} = \left(x_1^{(i,1)}, x_2^{(i,1)}, x_3^{(i,1)}, x_4^{(i,1)}, v_1^{(i,1)}, v_2^{(i,1)}\right)^T$ is a vector of latent variables $\otimes$ expresses an outer product (arranging all pair-wise products of elements of two vectors in vertical line except for duplicate terms), and $\psi(x) \equiv -x + x^3$ is a nonlinear activation function that exhibits bistable

neural dynamics. This modification was implemented to ensure that this circuit exhibited chaotic dynamics with separatrix crossing (Fukuda, Petrosky, & Konishi, 2016; Ogawa et al., 2016), which is caused by the bistable dynamics of an excitatory neuron ($x_1^{(i,j)}$). For simplification, the other three neurons were supposed to be linear functions instead of sigmoid functions. In this circuit model, two excitatory-inhibitory couplings ($x_1^{(i,j)} - x_2^{(i,j)}$ and $x_3^{(i,j)} - x_4^{(i,j)}$) exhibited oscillatory dynamics, whereas a weak mutual connection between two inhibitory neurons $x_2^{(i,j)}$ and $x_3^{(i,j)}$ made the attractor chaotic.

In the simulations, the leakage parameter $\rho = 0.1$ (to ensure stability), a small inhibitory synaptic weight $\epsilon = 0.2$ (which controlled the coupling strength between the first and second oscillators), and a large inhibitory synaptic weight $\lambda = 1.2$ (which determined the period of the second neural oscillator) were used. We supposed that the hidden causes $v^{(i,1)}$ converged to $x^{(i,2)}$ to simplify the simulation. The definition of $g^{(i,1)}$ was chosen to ensure that it could express a general quintic function by a linear product of a 2×462 matrix and a 462-dimensional vector. These parameters were learned by a student without supervision. When updating the posterior belief of hidden states, we smoothed the posterior trajectory by adding small amounts of components of the prior (i.e., a trajectory of an attractor without perturbation) to avoid a divergence of variables induced by a large perturbation. Training was repeated for 60 sessions. Each session had a 10 s sequence. Time resolution $dt = 10^{-3}$ [s] was used.

The following procedure was applied before training: (1) the student's initial states of $\mathbf{x}^{(i,1)}, \mathbf{x}^{(i,2)}$ at time $t = 0$ and the time constants τ_1, τ_2 [s] were optimized in relation to one of six teacher songs (these were in the range of $-1 \leq \mathbf{x}_1^{(i,j)}, \mathbf{x}_2^{(i,j)}, \mathbf{x}_3^{(i,j)} \mathbf{x}_4^{(i,j)} \leq 1, 1/60 \leq \tau_1 \leq 1/40$, and $1/6 \leq \tau_2 \leq 1/4$); (2) the posterior expectation of parameters $\theta^{(i,1)}$ was randomly generated; and (3) $\theta^{(i,1)}$ were modified by pretraining, in which each model randomly received one of six teacher songs, made an inference, and updated the parameters without model selection for 18 sessions to ensure that each internal model initially represented an averaged song. Then the response songs of a student were tested with different teacher songs (movie 1; see appendix B). For training, we randomly selected one of six teacher songs and provided it to a student (movie 2; see appendix B). Training was repeated for 60 sessions. After training, we tested the response songs again (see Figure 4 and movie 3; see appendix B).

4.4 Preprocessing for Natural Birdsong Data. The birdsong data used for Figures 4 and 5 and the supplementary movies were downloaded from http://ofer.sci.ccny.cuny.edu/song_database/zebra-finch-song -library-2015/view. This data set was recorded by the Tchernichovski

group (see Tchernichovski et al., 2001). We treated the data as follows. First, we acquired a spectrogram of the song by performing a Fourier transform with a 23.2 ms time window. As an analogy to a physiological model of vocal coda that generates a birdsong by sequences of power and tone (frequency) of the voice (Laje, 2002), we defined the leading frequency (s_2) and the amplitude (s_1) of a song by the mode of its frequency and the power of the mode frequency for each time step, respectively. They were normalized and introduced as sensory inputs $s = (s_1, s_2)^T$.

Appendix A: Supplementary Figures
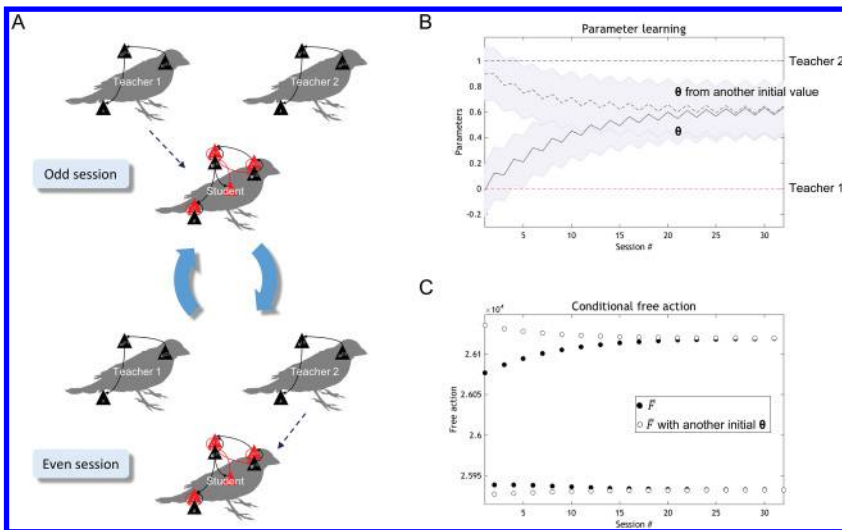
Figure 6: A schematic illustrating an experimental procedure (A) and simulation results of a synthetic bird with a single generative model (B,C). (A) Experimental procedure. Teacher bird 1 sings in odd sessions, while teacher bird 2 sings in even sessions. Our synthetic bird (student) listens to either song in turn. (B) Trajectory of the posterior expectation of a parameter (θ) of the student that employs a single generative model (black sold line). A black dashed line shows another trajectory where θ started from a different initial value. Shaded areas indicate the standard deviation. Red and blue dashed lines express the true parameter of teachers 1 and 2, respectively. The student tried to infer either parameter of teachers 1 or 2, but it failed to learn either parameter; even its posterior belief was initialized to the same value as either teacher 1 or 2's parameter. This is because the student inferred the intermediate value of teacher 1 and 2's parameters. (C) Transition of free action (filled circles). Open circles show the transition of free action with another initial θ.
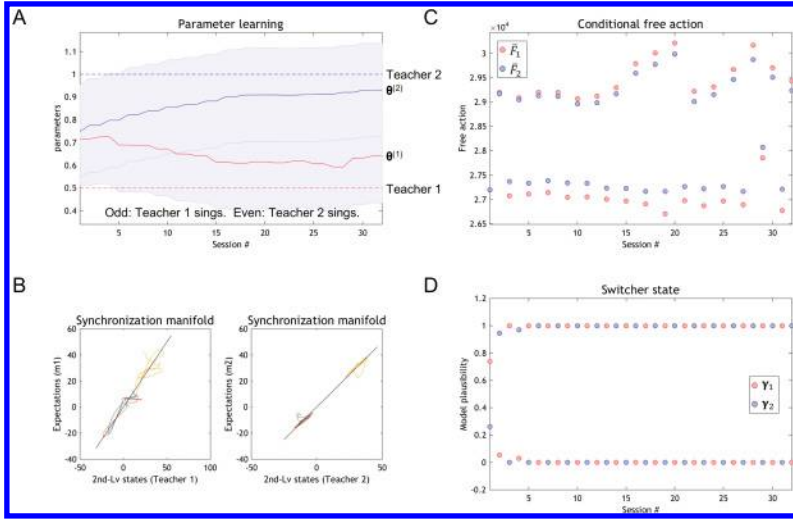
Figure 7: Simulation results when learning two birdsongs using multiple generative models. Simulation setup and layout are the same as Figure 3, but initial hidden states of teachers 1 and 2 were not reset for each session. This yielded the chaotic dynamics in their songs. Even in this case, a student bird employing the proposed scheme could learn from two distinct teachers. In this figure, to generate various song trajectories, a chaotic attractor considered in Kiebel et al. (2008) and Friston and Kiebel (2009) was used as the generative model instead of the Laje-Mindlin style model.
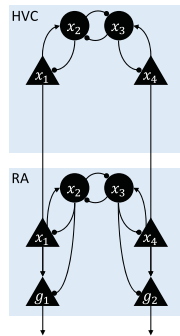

Figure 8: A birdsong generative model with two-layer four neuron circuits. This model is defined by extending the Laje-Mindlin style model to facilitate the song generation capability. The lower layer (level 1) corresponds to RA, while the higher layer (level 2) corresponds to HVC. In each layer, x_1 and x_4 are associated with excitatory neurons, while x_2 and x_3 are associated with inhibitory neurons. Signals of x_1 and x_4 in the HVC are translated into v_1 and v_2 in the RA. The sensory input (i.e., song) is generated through nonlinear functions g_1 and g_2 which receive inputs from $x_1, x_2, x_3, x_4, v_1,$ and v_2 (see section 4 for details).

Appendix B: Supplementary Movies 1–3

Movies 1 to 3 are available online at https://www.mitpressjournals.org /doi/suppl/10.1162/neco_a_01239. These movies show the dynamics of teacher and student birds before, during, and after training. The details are described in the caption to Figure 4.

Acknowledgments

Author Contributions

K.J.F. conceptualized the free-energy principle; T.I. conceived and designed the method using the multiple internal models and performed the simulations. T.I., T.P., and K.J.F. wrote the paper.

Competing Interests

The authors declare that they have no competing interests.

References

Adams, R. A., Shipp, S., & Friston, K. J. (2013). Predictions not commands: Active inference in the motor system. *Brain Structure and Function*, *218*, 611–643. doi:10.1007/s00429-012-0475-5. PMID:23129312.

Aljundi, R., Chakravarty, P., & Tuytelaars, T. (2017). Expert gate: Lifelong learning with a network of experts. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition* (pp. 3366–3375). Piscataway, NJ: IEEE. https://arxiv.org/abs/1611.06194.

Amador, A., Perl, Y. S., Mindlin, G. B., & Margoliash, D. (2013). Elemental gesture dynamics are encoded by song premotor cortical neurons. *Nature*, *495*, 59–65. doi:10.1038/nature11967. PMID:23446354.

Angelucci, A., & Bressloff, P. C. (2006). Contribution of feedforward, lateral and feedback connections to the classical receptive field center and extra-classical receptive field surround of primate V1 neurons. *Progress in Brain Research*, *154*, 93–120. doi:10.1016/S0079-6123(06)54005-1. PMID:17010705.

Awh, E., Belopolsky, A. V., & Theeuwes, J. (2012). Top-down versus bottom-up attentional control: A failed theoretical dichotomy. *Trends in Cognitive Sciences*, *16*, 437–443. doi:10.1016/j.tics.2012.06.010. PMID:22795563.

Bastos, A. M., Usrey, W. M., Adams, R. A., Mangun, G. R., Fries, P., & Friston, K. J. (2012). Canonical microcircuits for predictive coding. *Neuron*, *76*, 695–711. doi:10.1016/j.neuron.2012.10.038. PMID:23177956.

Ben Achour, S., & Pascual, O. (2010). Glia: the many ways to modulate synaptic plasticity. *Neurochemistry International*, *57*, 440–445. doi:10.1016/j.neuint.2010.02.013. PMID:20193723.

Bi, G. Q., & Poo, M. M. (1998). Synaptic modifications in cultured hippocampal neurons: Dependence on spike timing, synaptic strength, and postsynaptic cell type. *Journal of Neuroscience*, *18*, 10464–10472. doi:10.1523/JNEUROSCI.18-24-10464.1998. PMID:9852584.

Bliss, T. V., & Lømo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *Journal of Physiology*, *232*, 331–356. doi:10.1113/jphysiol.1973.sp010273. PMID:4727084.

Bosman, C. A., Schoffelen, J. M., Brunet, N., Oostenfeld, Bastos, A. M., Womelsdorf, T., . . . Fries, P. (2012). Attentional stimulus selection through selective synchronization between monkey visual areas. *Neuron*, *75*, 875–888. doi:10.1016/j.neuron.2012.06.037. PMID:22958827.

Brodski-Guerniero, A., Paasch, G. F., Wollstadt, P., Özdemir, I., Lizier, J. T., & Wibral, M. (2017). Information-theoretic evidence for predictive coding in the face-processing system. *Journal of Neuroscience*, *37*, 8273–8283. doi:10.1523/JNEUROSCI.0614-17.2017. PMID:28751458.

Buschman, T. J., & Kastner, S. (2015). From behavior to neural dynamics: An integrated theory of attention. *Neuron*, *88*, 127–144. doi:10.1016/j.neuron.2015.09.017. PMID:26447577.

Calabrese, A., & Woolley, S. M. (2015). Coding principles of the canonical cortical microcircuit in the avian brain. *Proceedings of the National Academy of Sciences of the USA*, *112*, 3517–3522. doi:10.1073/pnas.1408545112. PMID:25691736.

Collerton, D., Perry, E., & McKeith, I. (2005). Why people see things that are not there: A novel perception and attention deficit model for recurrent complex visual hallucinations. *Behavioral and Brain Sciences*, *28*, 737–757. doi:10.1017/S0140525X05000130. PMID:16372931.

Dajani, D. R., & Uddin, L. Q. (2015). Demystifying cognitive flexibility: Implications for clinical and developmental neuroscience. *Trends in Neurosciences*, *38*, 571–578. doi:10.1016/j.tins.2015.07.003. PMID:26343956.

Dauwels, J. (2007). On variational message passing on factor graphs. In *Proceedings of the 2007 IEEE International Symposium on Information Theory* (pp. 2546–2550). Piscataway, NJ: IEEE. doi:10.1109/ISIT.2007.4557602.

Dayan, P. (2012). Twenty-five lessons from computational neuromodulation. *Neuron*, *76*, 240–256. doi:10.1016/j.neuron.2012.09.027. PMID:23040818.

Dayan, P., Hinton, G. E., Neal, R. M., & Zemel, R. S. (1995). The Helmholtz machine. *Neural Computation*, *7*, 889–904. doi:10.1162/neco.1995.7.5.889. PMID:7584891.

Deubel, H., & Schneider, W. X. (1996). Saccade target selection and object recognition: Evidence for a common attentional mechanism. *Vision Research*, *36*, 1827–1837. doi:10.1016/0042-6989(95)00294-4. PMID:8759451.

Devaine, M., Hollard, G., & Daunizeau, J. (2014). Theory of mind: Did evolution fool us? *PLoS One*, *9*, e87619. doi:10.1371/journal.pone.0087619. PMID:24505296.

Devaine, M., San-Galli, A., Trapanese, C., Bardino, G., Hano, C., Saint Jeime, M., Daunizeau, J. (2017). Reading wild minds: A computational assay of theory of mind sophistication across seven primate species. *PLoS Computational Biology*, *13*, e1005833. doi:10.1371/journal.pcbi.1005833. PMID:29112973.

Doya, K. (2002). Metalearning and neuromodulation. *Neural Networks*, *15*, 495–506. doi:10.1016/S0893-6080(02)00044-8. PMID:12371507.

Eagleman, D. M. (2001). Visual illusions and neurobiology. *Nature Reviews Neuroscience*, *2*, 920–926. doi:10.1038/35104092. PMID:11733799.

Ereira, S., Dolan, R. J., & Kurth-Nelson, Z. (2018). Agent-specific learning signals for self–other distinction during mentalising. *PLoS Biology*, *16*, e2004752. doi:10.1371/journal.pbio.2004752. PMID:29689053.

Everitt, B. J., & Robbins, T. W. (1997). Central cholinergic systems and cognition. *Annual Review of Psychology*, *48*, 649–684. doi:10.1146/annurev.psych.48.1.649. PMID:9046571.

Feldman, D. E. (2012). The spike-timing dependence of plasticity. *Neuron*, *75*, 556–571. doi:10.1016/j.neuron.2012.08.001. PMID:22920249.

Feldman, H., & Friston, K. J. (2010). Attention, uncertainty, and free-energy. *Frontiers in Human Neuroscience*, *4*, 215. doi:10.3389/fnhum.2010.00215. PMID:21160551.

Forney, G. D. (2001). Codes on graphs: Normal realizations. *IEEE Transactions on Information Theory*, *47*, 520–548. doi:10.1109/18.910573.

Frémaux, N., & Gerstner, W. (2016). Neuromodulated spike-timing-dependent plasticity, and theory of three-factor learning rules. *Frontiers in Neural Circuits*, *9*, 85. doi:10.3389/fncir.2015.00085. PMID:26834568.

Fries, P. (2005). A mechanism for cognitive dynamics: Neuronal communication through neuronal coherence. *Trends in Cognitive Sciences*, *9*, 476–480. doi:10.1016/j.tics.2005.08.011. PMID:16150631.

Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society London B Biological Sciences*, *360*, 815–836. doi:10.1098/rstb.2005.1622. PMID:15937014.

Friston, K. (2008). Hierarchical models in the brain. *PLoS Computational Biology, 4*, e1000211. doi:10.1371/journal.pcbi.1000211. PMID:18989391.

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*, 127–138. doi:10.1038/nrn2787. PMID:20068583.

Friston, K., Adams, R. A., Perrinet, L., & Breakspear, M. (2012). Perceptions as hypotheses: Saccades as experiments. *Frontiers in Psychology*, *3*, 151. doi:10.3389/fpsyg.2012.00151. PMID:22654776.

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. (2017). Active inference: A process theory. *Neural Computation, 29*, 1–49. doi:10.1162/NECO_a_00912. PMID:27870614.

Friston, K. J, & Frith, C. D. (2015a). Active inference, communication and hermeneutics. *Cortex*, *68*, 129–143. doi:10.1016/j.cortex.2015.03.025. PMID:25957007.

Friston, K., & Frith, C. (2015b). A duet for one. *Consciousness and Cognition*, *36*, 390–405. doi:10.1016/j.concog.2014.12.003. PMID:25563935.

Friston, K., & Herreros, I. (2016). Active inference and learning in the lerebellum. *Neural Computation, 28*, 1812–1839. doi:10.1162/NECO_a_00863. PMID:27391681.

Friston, K., & Kiebel, S. (2009). Cortical circuits for perceptual inference. *Neural Networks*, *22*, 1093–1104. doi:10.1016/j.neunet.2009.07.023. PMID:19635656.

Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *Journal of Physiology Paris, 100*, 70–87. doi:10.1016/j.jphysparis.2006.10.001. PMID:17097864.

Friston, K., Levin, M., Sengupta, B., & Pezzulo, G. (2015). Knowing one's place: A free-energy approach to pattern regulation. *J. Royal Soc. Interface*, *12*, 20141383. doi:10.1098/rsif.2014.1383. PMID:25788538.

Friston, K., Mattout, J., & Kilner, J. (2011). Action understanding and active inference. *Biological Cybernetics, 104*, 137–160. doi:10.1007/s00422-011-0424-z. PMID:21327826.

Friston, K. J., Parr, T., & de Vries, B. D. (2017). The graphical brain: Belief propagation and active inference. *Network Neuroscience*, *1*, 381–414. doi:10.1162/NETN_a_00018. PMID:29417960.

Friston, K., & Penny, W. (2011). Post hoc Bayesian model selection. *NeuroImage*, *56*, 2089–2099. doi:10.1016/j.neuroimage.2011.03.062. PMID:21459150.

Friston, K. J., Trujillo-Barreto, N., & Daunizeau, J. (2008). DEM: A variational treatment of dynamic systems. *NeuroImage 41*, 849–885. doi:10.1016/j.neuroimage.2008.02.054. PMID:18434205.

Froemke, R. C., & Dan, Y. (2002). Spike-timing-dependent synaptic modification induced by natural spike trains. *Nature, 416*, 433–438. doi:10.1038/416433a. PMID:11919633.

Fukuda, H., Petrosky, T., & Konishi, T. (2016). Analytical investigation of an isomerization system using the resonance overlap criterion. *Progress of Theoretical and Experimental Physics*, *2016*, 093A01. doi:10.1093/ptep/ptw122.

Gentner, T. Q. (2004). Neural systems for individual song recognition in adult birds. *Annals of the New York Academy of Sciences*, *1016*, 282–302. doi:10.1196/annals.1298.008. PMID:15313781.

Gentner, T. Q., Hulse, S. H., Bentley, G. E., & Ball, G. F. (2000). Individual vocal recognition and the effect of partial lesions to HVc on discrimination, learning, and categorization of conspecific song in adult songbirds. *Journal of Neurobiology*, *42*, 117–133. doi:10.1002/(SICI)1097-4695(200001)42:1⟨117::AID-NEU11⟩3.0.CO;2-M. PMID:10623906.

Gentner, T. Q., & Margoliash, D. (2003). Neuronal populations and single cells representing learned auditory objects. *Nature*, *424*, 669–674. doi:10.1038/nature01731. PMID:12904792.

George, D., & Hawkins, J. (2009). Towards a mathematical theory of cortical micro-circuits. *PLoS Computational Biology*, *5*, e1000532. doi:10.1371/journal.pcbi.1000532. PMID:19816557.

Green, C. S., & Bavelier, D. (2003). Action video game modifies visual selective attention. *Nature*, *423*, 534–537. doi:10.1038/nature01647. PMID:12774121.

Happé, F., & Frith, U. (1995). Theory of mind in autism. In E. Schopler & G. B. Mesibov (Eds.), *Learning and cognition in autism*. Boston: Springer.

Haruno, M., Wolpert, D. M., & Kawato, M. (2003). Hierarchical MOSAIC for movement generation. *International Congress Series*, *1250*, 575–590. doi:10.1016/S0531-5131(03)00190-0.

Hassabis, D., Kumaran, D., Summerfield, C., & Botvinick, M. (2017). Neuroscience-inspired artificial intelligence. *Neuron*, *95*, 245–258. doi:10.1016/j.neuron.2017.06.011. PMID:28728020.

Hayama, T., Noguchi, J., Watanabe, S., Takahashi, N., Hayashi-Takagi, A., Ellis-Davies, G. C., . . . Kasai, H. (2013). GABA promotes the competitive selection of dendritic spines by controlling local CA2 signaling. *Nature Neuroscience*, *16*, 1409–1416. doi:10.1038/nn.3496. PMID:23974706.

Hebb, D. O. (1949). *The organization of behavior: A neuropsychological theory*. New York: Wiley.

Hedrick, T., & Waters, J. (2015). Acetylcholine excites neocortical pyramidal neurons via nicotinic receptors. *Journal of Neurophysiology*, *113*, 2195–2209. doi:10.1152/jn.00716.2014. PMID:25589590.

Heilbron, M., & Chait, M. (2017). Great expectations: Is there evidence for predictive coding in auditory cortex? *Neuroscience* pii:S0306-4522(17)30547-X. doi:10.1016/j.neuroscience.2017.07.061. PMID:28782642.

Helmholtz, H. (1925). *Treatise on physiological optics* (vol. *3*). Washington, DC: Optical Society of America.

Hohwy, J., Paton, B., & Palmer, C. (2016). Distrusting the present. *Phenomenology and the Cognitive Sciences*, *15*, 315–335. doi:10.1007/s11097-015-9439-6.

Hough, II, G. E., Nelson, D. A., & Volman, S. F. (2000). Re-expression of songs deleted during vocal development in white-crowned sparrows, Zonotrichia leucophrys. *Animal Behaviour*, *60*, 279–287. doi:10.1006/anbe.2000.1498. PMID:11007636.

Itti, L., & Koch, C. (2001). Computational modelling of visual attention. *Nature Reviews Neuroscience*, *2*, 194–203. doi:10.1038/35058500. PMID:11256080.

Johansen, J. P., Diaz-Matair, L., Hamanaka, H., Ozawa, T., Ycu, E., Koivumaa, J., . . . Le Doux, J. E. (2014). Hebbian and neuromodulatory mechanisms interact to trigger associative memory formation. *Proceedings of the National Academy of Sciences of the USA*, *111*, E5584–E5592. doi:10.1073/pnas.1421304111. PMID:25489081.

Kiebel, S. J., Daunizeau, J., & Friston, K. J. (2008). A hierarchy of time-scales and the brain. *PLoS Computational Biology*, *4*, e1000209. doi:10.1371/journal.pcbi.1000209. PMID:19008936.

Kilner, J. M., Friston, K. J., & Frith, C. D. (2007). Predictive coding: An account of the mirror neuron system. *Cognitive Processing*, *8*, 159–166. doi:10.1007/s10339-007-0170-2. PMID:17429704.

Kirkpatrick, J., Pascanu, R., Rabinowitz, N., Veness, J., Desjardins, G., Rusu, A. A., . . . Hadsell, R. (2017). Overcoming catastrophic forgetting in neural networks. *Proceedings of the National Academy of Sciences of the USA*, *114*, 3521–3526. doi:10.1073/pnas.1611835114. PMID:28292907.

Klein, D., Mok, K., Chen, J. K., & Watkins, K. E. (2014). Age of language learning shapes brain structure: A cortical thickness study of bilingual and monolingual individuals. *Brain Lang.*, *131*, 20–24. doi:10.1016/j.bandl.2013.05.014. PMID:23819901.

Knill, D. C., & Pouget, A. (2004). The Bayesian brain: The role of uncertainty in neural coding and computation. *Trends in Neurosciences*, *27*, 712–719. doi:10.1016/j.tins.2004.10.007. PMID:15541511.

Kok, P., Rahnev, D., Jehee, J. F. M., Lau, H. C., & de Lange, F. P. (2012). Attention reverses the effect of prediction in silencing sensory signals. *Cerebral Cortex*, *22*, 2197–2206. doi:10.1093/cercor/bhr310. PMID:22047964.

Kullback, S., & Leibler, R. A. (1951). On information and sufficiency. *Annals of Mathematical Statistics*, *22*, 79–86. doi:10.1214/aoms/1177729694.

Kuśmierz, Ł., Isomura, T., & Toyoizumi, T. (2017). Learning with three factors: modulating Hebbian plasticity with errors. *Current Opinion in Neurobiology*, *46*, 170–177. doi:10.1016/j.conb.2017.08.020. PMID:28918313.

Laje, R., Gardner, T. J., & Mindlin, G. B. (2002). Neuromuscular control of vocalizations in birdsong: A model. *Physical Review E*, *65*, 051921. doi:10.1103/PhysRevE .65.051921. PMID:12059607.

Laje, R., & Mindlin, G. B. (2002). Diversity within a birdsong. *Physical Review Letters*, *89*, 288102. doi:10.1103/PhysRevLett.89.288102. PMID:12513182.

LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*, 436–444. doi:10.1038/nature14539. PMID:26017442.

Lee, T. W., Lewicki, M. S., & Sejnowski, T. J. (2000). ICA mixture models for unsupervised classification of non-gaussian classes and automatic context switching in blind signal separation. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *22*, 1078–1089. doi:10.1109/34.879789.

Lipkind, D., Marcus, G. F., Bemis, D. K., Sasahara, K., Jacoby, N., Takahasi, M., & Tchernichorski, O. (2013). Stepwise acquisition of vocal combinatorial capacity in songbirds and human infants. *Nature*, *498*, 104–108. doi:10.1038/nature12173. PMID:23719373.

Lipkind, D., Zai, A. T., Hanuschkin, A., Marcus, G. F., Tchernichovski, O., & Hahnloser, R. H. (2017). Songbirds work around computational complexity by learning song vocabulary independently of sequence. *Nature Communications*, *8*, 1247. doi:10.1038/s41467-017-01436-0. PMID:29089517.

Long, M. A., & Fee, M. S. (2008). Using temperature to analyse temporal dynamics in the songbird motor pathway. *Nature*, *456*, 189–194. doi:10.1038/nature07448. PMID:19005546.

Luck, S. J., Woodman, G. F., & Vogel, E. K. (2000). Event-related potential studies of attention. *Trends in Cognitive Sciences*, *4*, 432–440. doi:10.1016/S1364-6613(00) 01545-X. PMID:11058821.

Malenka, R. C., & Bear, M. F. (2004). LTP and LTD: An embarrassment of riches. *Neuron*, *44*, 5–21. doi:10.1016/j.neuron.2004.09.012. PMID:15450156.

Mann, R. P., & Garnett, R. (2015). The entropic basis of collective behaviour. *J. Royal Soc. Interface.*, *12*, 20150037. doi:10.1098/rsif.2015.0037. PMID:25833243.

Mante, V., Sussillo, D., Shenoy, K. V., & Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *Nature*, *503*, 78–84. doi:10.1038/nature12742. PMID:24201281.

Markram, H., Lübke, J., Frotscher, M., & Sakmann, B. (1997). Regulation of synaptic efficacy by coincidence of postsynaptic APs and EPSPs. *Science*, *275*, 213–215. doi:10.1126/science.275.5297.213. PMID:8985014.

Moutoussis, M., Trujillo-Barreto, N. J., El-Deredy, W., Dolan, R. J., & Friston, K. J. (2014). A formal model of interpersonal inference. *Front. Hum. Neurosci.*, *8*, 160. doi:10.3389/fnhum.2014.00160. PMID:24723872.

Ogawa, S., Cambon, B., Leoncini, X., Viltot, M., Castillo-Negrete, D., Dif-Pradalier, G., & Garbet, X. (2016). Full particle orbit effects in regular and stochastic magnetic fields. *Physics of Plasmas*, *23*, 072506. doi:10.1063/1.4958653.

Paille, V., Fino, E., Du, K., Morera-Herreras, T., Perez, S., Kotaleski, J. H., & Venance, L. (2013). GABAergic circuits control spike-timing-dependent plasticity. *Journal of Neuroscience*, *33*, 9353–9363. doi:10.1523/JNEUROSCI.5796-12.2013. PMID:23719804.

Parkinson, C., & Wheatley, T. (2015). The repurposed social brain. *Trends in Cognitive Sciences*, *19*, 133–141. doi:10.1016/j.tics.2015.01.003. PMID:25732617.

Pawlak, V., Wickens, J. R., Kirkwood, A., & Kerr, J. N. (2010). Timing is not everything: Neuromodulation opens the STDP gate. *Frontiers in Synaptic Neuroscience*, *2*, 146. doi:10.3389/fnsyn.2010.00146. PMID:21423532.

Perl, Y. S., Arneodo, E. M., Amador, A., Goller, F., & Mindlin, G. B. (2011). Reconstruction of physiological instructions from zebra finch song. *Physical Review E*, *84*, 051909. doi:10.1103/PhysRevE.84.051909. PMID:22181446.

Rao, R. P., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*, 79–87. doi:10.1038/4580. PMID:10195184.

Reader, S. M., & Laland, K. N. (2002). Social intelligence, innovation, and enhanced brain size in primates. *Proceedings of the National Academy of Sciences of the USA*, *99*, 4436–4441. doi:10.1073/pnas.062041299. PMID:11891325.

Reynolds, J. N., Hyland, B. I., & Wickens, J. R. (2001). A cellular mechanism of reward-related learning. *Nature*, *413*, 67–70. doi:10.1038/35092560. PMID:11544526.

Roweis, S., & Ghahramani, Z. (1999). A unifying review of linear gaussian models. *Neural Computation*, *11*, 305–345. doi:10.1162/089976699300016674. PMID:9950734.

Saalmann, Y., & Kastner, S. (2009). Gain control in the visual thalamus during perception and cognition. *Current Opinion in Neurobiology*, *19*, 408–414. doi:10.1016/j.conb.2009.05.007. PMID:19556121.

Salgado, H., Köhr, G., & Treviño, M. (2012). Noradrenergic "tone" determines dichotomous control of cortical spike-timing-dependent plasticity. *Scientific Reports*, *2*, 417. doi:10.1038/srep00417. PMID:22639725.

Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics*, *6*, 461–464. doi:10.1214/aos/1176344136.

Schwing, R., Nelson, X. J., Wein, A., & Parsons, S. (2017). Positive emotional contagion in a New Zealand parrot. *Current Biology*, *27*, R213–R214. doi:10.1016/j.cub.2017.02.020. PMID:28324733.

Seol, G. H., Ziburkus, Huang, S., Song, L., Kim, I. T., Takamiya, K., . . . Kirkwood, A. (2007). Neuromodulators control the polarity of spike-timing-dependent synaptic plasticity. *Neuron*, *55*, 919–929. doi:10.1016/j.neuron.2007.08.013. PMID:17880895.

Shipp, S. (2016). Neural elements for predictive coding. *Frontiers in Psychology*, *7*, 1792. doi:10.3389/fpsyg.2016.01792. PMID:27917138.

Shultz, S., & Dunbar, R. I. M. (2010). Species differences in executive function correlate with hippocampus volume and neocortex ratio across nonhuman primates. *Journal of Comparative Psychology*, *124*, 252–260. doi:10.1037/a0018894. PMID:20695656.

Suh, S., Chae, D. H., Kang, H. G., & Choi, S. (2016). Echo-state conditional variational autoencoder for anomaly detection. In *Proceedings of the International Joint*

Conference on Neural Networks (pp. 1015–1022). Piscataway, NJ: IEEE. doi:10.1109 /IJCNN.2016.7727309.

Suzuki, T. N., Wheatcroft, D., & Griesser, M. (2016). Experimental evidence for compositional syntax in bird calls. *Nature Communications*, 7, 10986. doi:10.1038 /ncomms10986. PMID:26954097.

Taborsky, B., & Oliveira, R. F. (2012). Social competence: An evolutionary approach. *Trends in Ecology and Evolution*, 27, 679–688. doi:10.1016/j.tree.2012.09 .003. PMID:23040461.

Tani, J., & Nolfi, S. (1999). Learning to perceive the world as articulated: An approach for hierarchical learning in sensory-motor systems. *Neural Networks*, 12, 1131–1141. doi:10.1016/S0893-6080(99)00060-X. PMID:12662649.

Tchernichovski, O., Mitra, P. P., Lints, T., & Nottebohm, F. (2001). Dynamics of the vocal imitation process: How a zebra finch learns its song. *Science*, 291, 2564–2569. doi:10.1126/science.1058522. PMID:11283361.

Vossel, S., Bauer, M., & Mathys, C. (2014). Cholinergic stimulation enhances Bayesian belief updating in the deployment of spatial attention. *Journal of Neuroscience*, 34, 15735–15742. doi:10.1523/JNEUROSCI.0091-14.2014. PMID:25411501.

Whittington, J. C., & Bogacz, R. (2017). An approximation of the error backpropagation algorithm in a predictive coding network with local Hebbian synaptic plasticity. *Neural Computation, 29*, 1229–1262. doi:10.1162/NECO_a_00949. PMID:28333583.

Womelsdorf, T., & Fries, P. (2006). Neuronal coherence during selective attentional processing and sensory-motor integration. *Journal of Physiology Paris, 100*, 182–193. doi:10.1016/j.jphysparis.2007.01.005. PMID:17317118.

Woolley, S. (2012). Early experience shapes vocal neural coding and perception in songbirds. *Developmental Psychobiology, 54*, 612–631. doi:10.1002/dev.21014. PMID:22711657.

Wurtz, R. H. (2008). Neuronal mechanisms of visual stability. *Vision Research, 48*, 2070–2089. doi:10.1016/j.visres.2008.03.021. PMID:18513781.

Yagishita, S., Hayashi-Takagi, A., Ellis-Davies, G. C., Urakubo, H., Ishii, S., . . . Kasai, H. (2014). A critical time window for dopamine actions on the structural plasticity of dendritic spines. *Science, 345*, 1616–1620. doi:10.1126/science .1255514. PMID:25258080.

Yanagihara, S., & Yazaki-Sugiyama, Y. (2016). Auditory experience-dependent cortical circuit shaping for memory formation in bird song learning. *Nature Communications, 7*, 11946. doi:10.1038/ncomms11946. PMID:27327620.

Yu, A. J., & Dayan, P. (2005). Uncertainty, neuromodulation and attention. *Neuron, 46*, 681–692. doi:10.1016/j.neuron.2005.04.026. PMID:15944135.

Zeki, S., & Shipp, S. (1988). The functional logic of cortical connections. *Nature, 335*, 311–317. doi:10.1038/335311a0. PMID:3047584.

Zelinsky, G. J., & Bisley, J. W. (2015). The what, where, and why of priority maps and their interactions with visual working memory. *Annals of the New York Academy of Sciences, 1339*, 154–164. doi:10.1111/nyas.12606. PMID:25581477.

Zenke, F., Poole, B., & Ganguli, S. (2017). Continual learning through synaptic intelligence. *In Proceedings of the International Conference on Machine Learning* (pp. 3987–3995). http://proceedings.mlr.press/v70/zenke17a.html.

Zhang, J. C., Lau, P. M., & Bi, G. Q. (2009). Gain in sensitivity and loss in temporal contrast of STDP by dopaminergic modulation at hippocampal synapses. *Proceedings of the National Academy of Sciences of the USA, 106*, 13028–13033. doi:10.1073 /pnas.0900546106. PMID:19620735.