

# Cognitive neurorobotics and self in the shared world, a focused review of ongoing research

Adaptive Behavior

1–20

© The Author(s) 2020



Article reuse guidelines:

[sagepub.com/journals-permissions](https://sagepub.com/journals-permissions)

DOI: 10.1177/1059712320962158

[journals.sagepub.com/home/adb](https://journals.sagepub.com/home/adb)

Jun Tani<sup>1</sup> and Jeffrey White<sup>1,2</sup>

## Abstract

Through brain-inspired modeling studies, cognitive neurorobotics aims to resolve dynamics essential to different emergent phenomena at the level of embodied agency in an object environment shared with human beings. This article is a review of ongoing research focusing on model dynamics associated with human self-consciousness. It introduces the free energy principle and active inference in terms of Bayesian theory and predictive coding, and then discusses how directed inquiry employing analogous models may bring us closer to representing the sense of self in cognitive neurobots. The first section quickly locates cognitive neurorobotics in the broad field of computational cognitive modeling. The second section introduces principles according to which cognition may be formalized, and reviews cognitive neurorobotics experiments employing such formalizations. The third section interprets the results of these and other experiments in the context of different senses of self, both “minimal” and “narrative” self. The fourth section considers model validity and discusses what we may expect ongoing cognitive neurorobotics studies to contribute to scientific explanation of cognitive phenomena including the senses of minimal and narrative self.

## Keywords

Cognitive robotics, cognitive architecture, neurorobotics, free energy, predictive coding, active inference, the sense of self

Handling Editor: Massimiliano Cappuccio, University of New South Wales, Australia

## 1. Prospectus

The following article proceeds in four movements. The first section very briefly introduces cognitive neurorobotics (CNR). The second section reviews a series of humanoid CNR experiments designed to elicit specific cognitive phenomena in emergent dynamics, and the third section interprets this review in the context of minimal and narrative self. The fourth section emphasizes the potential for CNR studies of this sort to contribute to inquiry into embodied cognition and mind.

### 1.1. Introducing CNR

The main motivation of CNR is to elucidate essential mechanisms underlying embodied cognition through synthesis of analogous dynamics in various robotics experiments. CNR calls on diverse interdisciplinary knowledge including from fields of cognitive science, psychology, ethology, neuroscience, complex systems,

artificial intelligence (AI) and deep learning, artificial life, and many others. Primarily, CNR can be considered a marriage of two research fields. One is cognitive robotics (e.g. Levesque & Lakemeyer, 2008) which aims to develop human-level intelligence in robots using a rather conventional symbolism approach, and the other is neurorobotics which puts more emphasis on the realization of adaptive behaviors of biological systems using neuroscience inspired models or neuromorphic schemes.

<sup>1</sup>Cognitive Neurorobotics Research Unit, Okinawa Institute of Science and Technology (OIST), Onna-son, Japan

<sup>2</sup>Department of Philosophy, University of Twente, Enschede, The Netherlands

#### Corresponding author:

Jeffrey White, Cognitive Neurorobotics Research Unit, Okinawa Institute of Science and Technology (OIST), Onna-son 904-0495, Okinawa, Japan.  
Email: [jeffreywhitephd@gmail.com](mailto:jeffreywhitephd@gmail.com)

Initially, cognitive robotics studies proceeded with a strong conviction that formal logical descriptions of the world and rational computation for reasoning, planning, and inference could provide for human-level cognitive competency in artificial agents including robots. However, such early expectations were betrayed by the results of projects such as SRI's SHAKEY (Nilsson, 1984) which demonstrated the problems associated with the rigors of applying formal logic to real robots. Part of the reason for the trouble is that representation of the world using symbols as arbitrary tokens cannot be grounded smoothly with real-world phenomena which are fundamentally given to experience in terms of continuous sensory-motor patterns. This is the point of Harnad's (1990) famous symbol grounding problem.

Already by the end of the 1980s, a paradigm shift had been taking place in AI and robotics research with the introduction of behavior-based robotics by Rodney Brooks (1990, 1991). Brooks considered that even simple insect-like robots can exhibit extremely complex and intelligent behaviors by establishing a direct coupling between a robot's reflex-type controllers and sensations from the environment. His seminal paper, "Intelligence without representation" (Brooks, 1991) represents this thought—no representation and thus no grounding problem.

Neurorobotics in general is a growing field of especially fruitful inquiry employing biological system-inspired algorithms in a range of applications, from prosthetics with brain-machine interface technologies (Millan et al., 2010; Moxon and Foffani, 2015) to independently embodied robots with autonomous locomotion, learning, memory, value, and action selection systems (Doya et al., 2002; D. M. Kaplan, 2015; Krichmar, 2018; Kuniyoshi & Sangawa, 2006). The degree of biological precision in selection of neuronal and kinematic models depends on the degree of realism required to represent target phenomena. Extreme realism is represented in the Human Brain Project (HBP). Neurorobotics is considered a "strategic pillar" of the HBP through which biologically inspired algorithms representing levels of organization from molecular mechanism to modular function to unified cognitive architecture can be tested in simulation and then deployed in physically embodied robots sharing physical space with human beings (Knoll & Gewaltig, 2016). In general, however, more specific studies replicate focal operations in fine detail while rendering other aspects more abstract. Recognizing the impossibility of analyzing all levels of activity simultaneously for instance, Krichmar and Edelman (2002) focus on how cortical and subcortical levels interact in real time using a relatively simple embodied robot, Darwin VII.

There has been a group of researchers who have emphasized advantages of studying so-called "minimum cognition." These researchers have focused on phenomena emerging during system-level interaction

with the environment using relatively simple neuronal adaptive controllers (e.g. Barandiaran & Chemero, 2009; Beer, 2000; Froese & Ziemke, 2009; Iizuka & Di Paolo, 2007; Nolfi & Floreano, 2000; Silberstein & Chemero, 2013). Beer (2000) viewed an agent's nervous system, its body, and its environment as coupled dynamical systems. By focusing on the unfolding trajectory of the agent's system state as shaped by both forces internal to the agent and external from the environment, he attempts to extract the essential dynamic structures accounting for minimal cognition. Since most neurorobotics studies inherit the aforementioned thoughts of the behavior-based robotics, current research tends to stay close to the realization of biologically plausible adaptive behavior, focusing on sensory-motor level processing and hesitating to explore mechanisms associated with higher cognition in human beings.

On the other hand, one of the main motivations of CNR studies is to consider possible principles, algorithms, and implementation designs which can bridge the gap between higher cognition and the lower sensory-motor processing of robots. Hybrid models (e.g. Sun, 2002) attempt to combine these two levels, extracting symbolized rules and associations at one level from sensory-motor patterns of activity at another (cf. Kotseruba & Tsotsos, 2018). However, such an enterprise may suffer again from the symbol grounding problem (Harnad, 1990) since these two levels do not share the same metric space required for the dense interactions between top-down and bottom-up processes that are associated with subjective experience including the sense of self.

Recently, deep learning schemes in robots show promise in attacking this problem. It has been known that various types of deep learning networks can develop hierarchical information processing in collective spatio-temporal activities of the neural units through end-to-end learning of sensory-motor patterns. Actually, such trials have been conducted by various research groups including the authors' (Heinrich & Wermter, 2018; Levine et al., 2016; Yamada et al., 2016; Yamashita & Tani, 2008). Developmental robotics (Asada et al., 2009; Cangelosi & Schlesinger, 2015; Metta et al., 2008) is another indispensable approach to address this problem. In developmental robotics, cognitive competencies of artificial agents or robots develop gradually, supported by human tutors, with scaffolding from level to level according to fundamental theories in child development (e.g. Piaget, 1953; Vygotsky et al., 1934/1962).

Some CNR research attempts to expose connections between the phenomenology of subjective experience and embodied sensory-motor processing. Holland (2007) conjectured that building human-like bodies for robots and developing internal models for predicting body dynamics is essential for developing "machine

consciousness.” Prescott and Camilleri (2019) consider that the sense of self can be characterized as a transient process, analogous to Tani’s (1998) consideration that self becomes an object of consciousness when prediction error for the actional outcome momentarily increases (as will be described later, in Section 3). Lanillos and Cheng (2018) proposed that prediction error generated by a body’s forward model should come with the sense of what they call “enactive self” in the differentiation between inbody and other sources. And, Hafner and colleagues (Lang et al., 2018; Schillaci et al., 2016) conducted a set of robotics experiments also examining the sense of agency wherein they observed attenuation of sensory inputs to self-generated movements (in terms of prediction error minimization) but not to those of others (as these were unpredictable). Keeping this research in mind, it is becoming more crucial that the problems of cognition and of subjective experience should be investigated inseparably to gain a better understanding of the minds of both humans and artifacts.

The CNR experiments reviewed in the next section aim to uncover such structural dynamics by using an approach analogous to the free energy minimization principle (FEP) proposed by Friston (2005). The FEP approach can be interpreted in terms of Marr’s three levels (Marr, 1982) wherein first, the computational level might be represented by the FEP itself, in which the goal of computation is minimizing the free energy. Second, the representation and algorithmic level might be represented by the schemes of predictive coding (Clark, 2015; Friston, 2005; Rao & Ballard, 1999) and active inference (Clark, 2015; Friston et al., 2011; Hohwy, 2013). Finally, the implementation level might be represented by neurophysiology in biological brains or artificial neural network programs put into a robot’s head. The next section starts with a brief review of the FEP, predictive coding, and active inference, and then reviews a set of CNR experiments conducted by Tani’s group employing analogous principles. Their validity as evidence for explanations in cognitive science is recalled in Section 4.

## 2. Models and CNR experiments

First, we provide a brief introduction of the FEP (Friston, 2005). Then, some neural network models developed by Tani’s group similar to this principle are introduced along with CNR experiments using those models. These detailed descriptions of technical aspects may help to understand how this and similar ongoing research may contribute to inquiry into phenomena such as self, as explored in Section 3 and as proposed in Section 4.

### 2.1. The FEP

The FEP states that any self-organizing system at dynamic equilibrium with its environment must

minimize its free energy to maintain this equilibrium and thereby its organization in the face of otherwise disintegrative forces. The FEP is an application of Bayesian theory. Bayes’ general idea was that we can calculate the probability of something being true (before we have evidence and rather than directly testing for it), and to do so, we may use what we have already observed to predict what we should perceive next. By applying this principle to adaptation mechanisms in brains, Friston considers that all essential cognitive mechanisms including perception, action generation, and learning can be explained.

It may help to illustrate with an intuitive image of the FEP at work. Consider a snowflake with wings, fluttering about. As long as it stays in a certain conditional zone—not too turbulent, freezing air—then it continues fluttering about and is able to maintain its unique structure. If it falls out of this zone, then its integrity is lost and it dissipates. If its situation becomes too hot, then it melts, for example. In the simplest of terms, the FEP (along with active inference, also introduced below) tells us that the snowflake will do what it can to stay in this zone. As an application of Bayesian theory, the FEP attempts to describe how biological brains update beliefs (or hypotheses, what we think is the case) in light of new information, so that organisms can act on this new information toward what they wish to be the case, like snowflakes countering gusts of wind to stay in their comfort zones by anticipating from which direction the likeliest are to come next.

Formally, the relationship between the likelihood of an observation after we have experiential evidence, with what we thought that the likelihood might have been before we had such evidence, is represented in a mathematical formula, Bayes’ theorem. Bayes’ theorem updates prior probabilities (original hypotheses) with new evidence to produce new (posterior, after the evidence is gathered) probabilities that are then useful to guide the next iteration of action, for example, the snowflake’s next wing flap. The FEP tells us that biological brains do this to stay alive (like the snowflake does to maintain its unique organization), so they (in general) perceive what they need to perceive to inform beliefs that they need to believe to inform action that they need to enact to maintain their integrity in the face of dissipative change. With its brain optimized accordingly, an organism aims to minimize the difference between what it expects to happen and what it perceives, especially concerning those parameters that bear on its integrity, for example, to stay in that comfort zone. In the end, it is this adaptive updating of hypotheses in light of new information that is the focus of the CNR experiments reviewed below, and it is the temporal hierarchy characterizing biological brains grounding the anticipatory nature of the experience that their focal architectures reflect.

In terms of these robotics experiments, Bayes' theorem is shown in formula (1). By applying this formula, the robots are able to infer the hidden cause  $z$  of a sensory observation  $X$

$$p(z|X) = \frac{p(X|z)p(z)}{p(X)} \quad (1)$$

$p(z|X)$  is the posterior distribution of hidden cause  $z$  with given observation  $X$ , and  $p(X|z)$  is the likelihood which relates the sensory observation  $X$  to the hidden causes  $z$ .  $P(z)$  is the prior distribution of (probability density of) or belief in  $z$  as the latent cause before observation  $X$ .  $P(X)$  is the marginal likelihood which is obtained by marginalizing  $p(X|z)$  for all  $z$ .  $P(X)$  is obtained by considering the probabilities of  $X$  for all possible hidden causes. There is a practical problem, however, in that the direct computation of  $p(X)$  is often intractable with current methods. Simply too many possibilities must be considered in practice.

Since  $p(X)$  is necessary for the computation of the posterior distribution  $p(z|X)$ , rather than directly computing over all these possibilities, the variational free energy approach derives an approximation of  $p(z|X)$ . This scheme optimizes an auxiliary probability distribution  $q(z)$ , referred to as the recognition density, in approximation of the true posterior  $p(z|X)$  by minimizing the Kullback–Leibler (KL) divergence between the two (formula (2))

$$D_{KL}[q(z) \parallel p(z|X)] = \int q(z) \ln \frac{q(z)}{p(z|X)} dz = F + \ln p(X) \quad (2)$$

Free energy  $F$  is defined as

$$F = \int q(z) \ln \frac{q(z)}{p(z, X)} dz \quad (3)$$

Since the marginal likelihood  $p(X)$  in the second term on the right-hand side in equation (2) is independent of the recognition density  $q(z)$ , minimization of the KL divergence between the recognition density and the true posterior can be achieved by minimization of the free energy in equation (3) with respect to  $q(z)$ . This makes  $q(z)$  an adequate practical approximation of the true posterior  $p(z|X)$ .

The free energy  $F$  to be minimized can thus be rewritten in a computationally tractable form in terms of  $q(z)$ . In this equation,  $\varphi$  and  $\theta$  are model parameters (formula (4))

$$F = -E_{q_\varphi(z)}[\ln p_\theta(X|z)] + D_{KL}[q_\varphi(z) \parallel p(z)] \quad (4)$$

The first term on the right side of the equation, the accuracy term, represents the expectation of log-likelihood with respect to the approximate posterior,

which represents reconstruction of the sensory observation with the approximate posterior by generative models. The second term, the complexity term, is represented by KL divergence between the approximate posterior and the prior, which serves to regularize the model according to prior expectation.

Free energy  $F$  can be minimized with respect to  $q_\varphi(z)$  as

$$q_\varphi(z) = \arg \min F \quad (5)$$

Here, we may put the preceding in terms of predictive coding (Clark, 2015; Friston, 2005; Rao & Ballard, 1999). For a given sensory observation, the posterior inference of the latent variable (“latent” means hidden, so this is what is hypothesized to be the hidden cause) is carried out by means of minimizing the error between the sensation expected by the generative model for the latent variable and its observation under the constraint of the prior distribution of that latent variable. The idea here again is to minimize the difference between what was predicted and what is perceived.

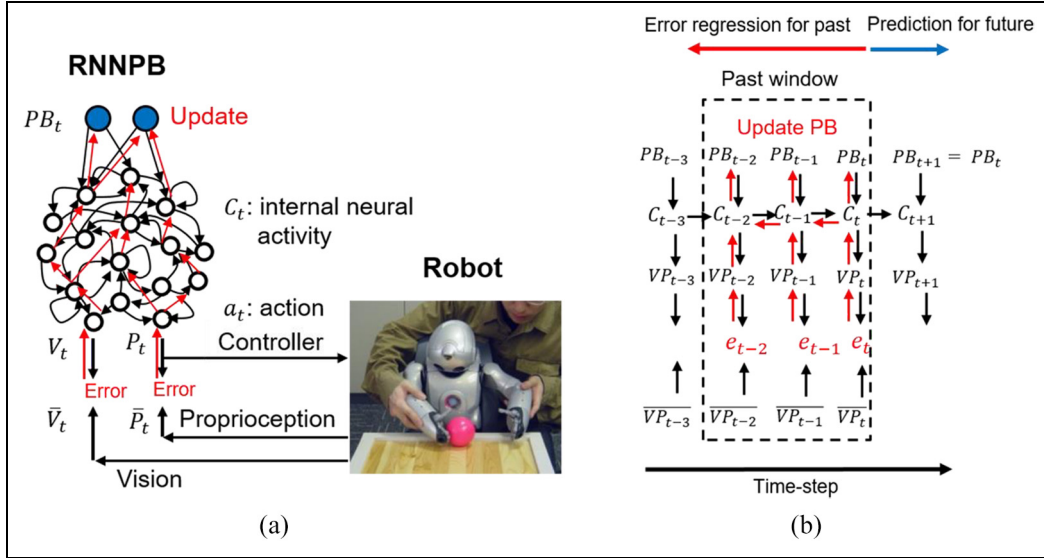
Free energy  $F$  integrated for the predicted future time period is the expected free energy  $F^E$  of a given model. We aim to minimize this value with respect to an action  $a$  by assuming a forward model that represents the likelihood of an action causing an observation, given the cause  $z$ , as shown in formulae (6) and (7)

$$F^E = -E_{q_\varphi(z)}[\ln p_\theta(X(a)|z)] \quad (6)$$

$$a = \arg \min F^E \quad (7)$$

Note that the expected free energy does not involve the KL divergence between the posterior and the prior (as in formula (4)) because KL divergence is independent of sensation (evidence,  $X$ ) and thus of action.

The preceding process represents action generation by active inference (Clark, 2015; Friston et al., 2011; Hohwy, 2013) whereby actions are selected such that the expected free energy is minimized.<sup>1</sup> (Later we introduce a more recent update of active inference (R. Kaplan & Friston, 2018) considering the epistemic value.) More intuitively, to minimize the error between the expected or preferred sensory outcome and the actual one, the actual one is modified to become closer to the preferred one by acting adequately on the environment. Finally, perception by equation (5) and action generation by equation (7) are carried out, simultaneously, thereby closing the loop between action and perception (Baltieri & Buckley, 2017). Here, it is considered that enactive cognition is a continuing trial for minimizing the conflictive error between the top-down intention projected from the latent variables and the bottom-up perception of reality through iterative



**Figure 1.** A robot controlled by an RNNPB model. (a) During movement, PB values are updated by backpropagation in the direction of minimizing error and (b) predicting the future by inferring the most likely past with black and red arrows representing the generative and error backpropagation processes during prediction and postdiction, respectively.

acts changing the external environment as well as modifying the intention within (Tani, 2016).

## 2.2. Recurrent neural network with parametric biases

Tani and colleagues (Tani, 2003; Tani et al., 2004) developed the recurrent neural network with parametric biases (RNNPB) independently from the FEP. Since that time, the RNNPB has turned out to be one possible neural network implementation of the FEP (with some simplification as detailed later). Although there have been variations in the RNNPB specific to different applications, the following describes a typical version.

**2.2.1. Model description.** In the RNNPB, the following objective function for a time series of  $L$  time steps is minimized (formula (8))

$$Obj(pb_{1...T}, W) = \frac{1}{2} \sum_{t=1}^L \left( \frac{1}{N^o \sigma^2} \|\bar{X}_t - X_t(W)\|^2 + \frac{1}{N^{pb}} pb_t^2 \right) \quad (8)$$

In equation (8),  $\bar{X}_t$  and  $X_t(W)$  are the sensory observation and the prediction outputs for the sensory observation as a function of the learning parameters  $W$  with their dimensionality  $N^o$ , respectively. Sensation may consist of exteroception such as vision  $v_t$  and proprioception  $p_t$  in a robotics application as shown in Figure 1(a). In equation (8),  $pb_{1...T}$  is a latent variable of  $N^{pb}$  dimensions.

This objective function equation (8) becomes equal to the free energy defined in equation (4) with assumptions<sup>2</sup> that the prior of the latent variable is represented by a unit Gaussian distribution, the posterior by a Gaussian distribution with its mean  $pb_t$  and its standard deviation of 1.0, and the sensory prediction outputs by a Gaussian distribution with its mean  $X_t$  and its standard deviation  $\sigma$ . Then, this objective function can be minimized with respect to  $pb_{1...T}$ , given fixed learning parameters  $W$  by the deterministic dynamics described in the following difference equations (formulas (9a) to (9d))

$$\begin{cases} C_{t+1}^i = g \left( \sum_j w_{ij}^c C_t^j + \sum_h w_{ih}^{pb} PB_{t,n}^h + b_i^c \right) & (9a) \\ pb_{t,n}^h = \alpha \left( -pb_{t,n}^h - \nabla E(pb_{t,n}^h) \right) + pb_{t,n}^h & (9b) \\ PB_{t,n}^h = g(pb_{t,n}^h) & (9c) \\ X_{t+1}^o = g \left( \sum_j w_{oj}^{out} C_t^j + b_o^{out} \right) & (9d) \end{cases}$$

In equation (9a),  $C_t$  represents the activation of internal units with recurrence at time step  $t$ .  $PB_{t,n}$  represents the outputs of the sigmoid function  $g()$  applied to the latent variable  $pb_{t,n}$  during the  $n$ th epoch iteration at time step  $t$ . The latent variable  $pb$  is inferred at each time step through  $N$  epochs of iterations of the internal computation loop. Its value is updated in the direction of minimizing prediction error in the output by following  $\nabla E(pb_{t,n}^h)$ , which is the gradient of the mean square error  $E$  of the prediction outputs for all time steps, while also considering a unit decay effect of this value

as represented in the first term on the right-hand side of equation (9b).  $w^c$ ,  $w^{pb}$ ,  $w^{out}$ ,  $b^c$ , and  $b^{out}$  are elements of the learning parameters  $W$ .

The actual computation of equation (9) is carried out for  $N$  epochs iterated in the forward computation of  $L$  time steps, with backward error regression computed for the same  $L$  time steps using a past window storing all temporal variables from time step  $t-L+1$  to the current time step  $t$  as illustrated in Figure 1(b). In the forward computation, values of the internal units as well as those of the prediction outputs from time step  $t-L+1$  to the next (anticipated future) time step  $t+1$  are computed using the PB values updated for each time step. In the backward error regression, the prediction error at each time step is backpropagated through time (BPTT) (Rumelhart et al., 1985; Werbos, 1990) from the current time step  $t$  to time step  $t-L+1$  (at the onset of the past window). The latent variable  $pb_{t-L+1\dots t}$  in the window is updated by using the error gradient information obtained by BPTT at each time step according to equation (9b).

The predicted proprioception at time step  $t+1$  is sent to the robot’s proportional–integral–derivative (PID) controller as target joint angles for the next time step, and the robot moves accordingly. Actuators receive the motor commands  $a$  generated by the PID controller, move according to these commands, and this movement generates new visual and proprioceptive (bottom-up) sensations and their (top-down) prediction errors at time step  $t+1$  again with the aim to minimize these errors. The PID controller thereby instantiates the idea of active inference, since action  $a$  is generated to minimize the error between the actual and the predicted joint angles (proprioception; Baltieri & Buckley, 2018).

Finally, the learning of the RNNPB can be carried out by minimizing the objective function equation (8) with respect to both the learning parameters  $W$  and the latent variable  $pb_{1\dots T}$

$$W_{n+1} = -\beta \nabla E(W_n) \quad (10)$$

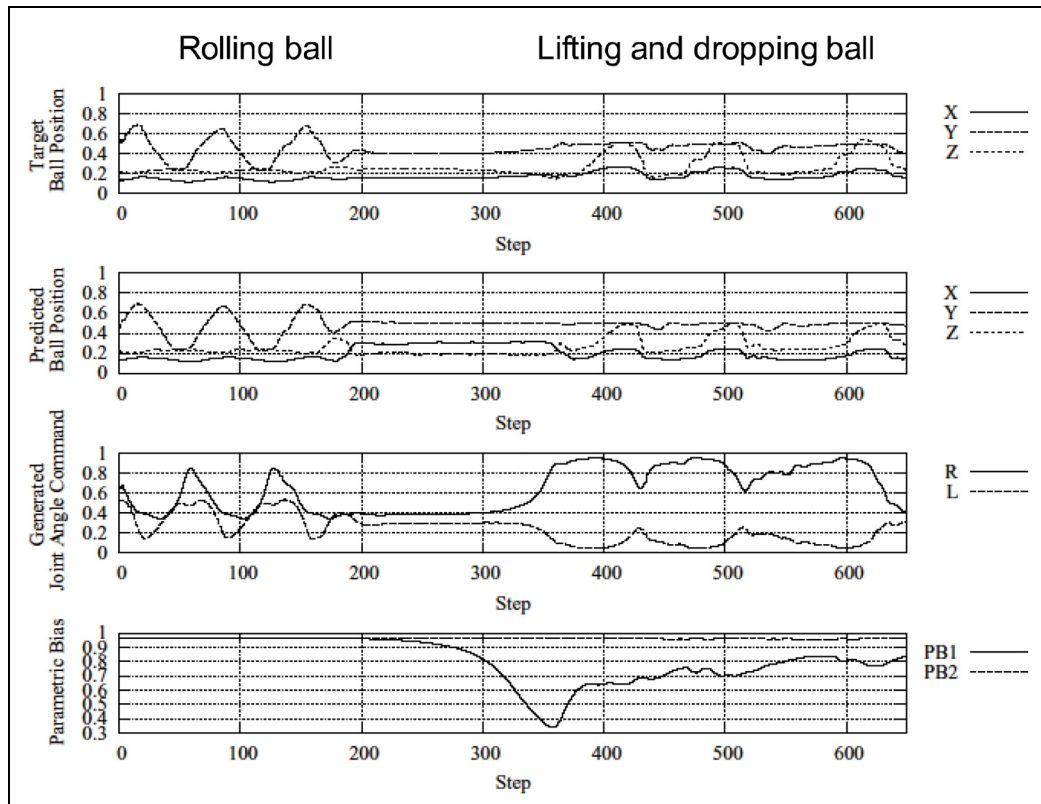
$W$  is updated at each epoch using the gradient information computed by BPTT with the learning rate  $\beta$ .

**2.2.2. Robotics experiment using RNNPB.** To evaluate the RNNPB in a robotics experiment, a humanoid robot was used which consists of three subsystems: an onboard sensory processing module with a head-mounted video camera, an RNNPB module running on an external computer, and an onboard motor control module. So configured, the RNNPB could predict two types of sensory inputs, proprioception in terms of joint angles in both arms of the robot, and visual features representing target object position (X-Y-Z) at each time step.

During ball manipulation experiments (Ito et al., 2006), human tutors trained this robot to manipulate a ball in two different sequences: repetitively pushing the ball from left to right and right to left, and repetitively grasping the ball at the center position, lifting it up, and then dropping it. After the training of the network wherein PB values adapted differently for each action sequence, the robot was tested to generate these sequences autonomously (without external help). During testing, the robot switched from one to the other intermittently, with an example presented in Figure 2 (also see a video: [https://youtu.be/a\\_auIoksGN0](https://youtu.be/a_auIoksGN0)). In this instance, the robot initially pushed the ball from left to right repeatedly until the ball bounced off of one hand too much, rolling to the center position. This unexpected movement caused prediction error. To minimize this error, the robot adopted the PB value for grasping and dropping the ball, and its behavior changed accordingly. Here, we see that the top-down intention to act with the ball, represented by the PB, shifted autonomously during iterative interaction with the bottom-up percept of the ball position from one behavioral reparatory to another in the course of minimizing the error.

Importantly, the continuous sensory flow was segmented during the error minimization process as the robot optimized coordination with the target object through autonomous shifts in PB values. Shifting PB values served as bifurcation parameters to induce transitions, effectively steering the system from one learned behavioral pattern to the other. This same mechanism for the segmentation and chunking of the continuous perceptual flow had been observed during an earlier RNNPB experiment on human–robot interaction (Ito & Tani, 2004) wherein a robot and a human participant attempted to synchronously imitate one another’s primitive movement patterns using prediction based on prior learning. With robot and human participant movements synchronized according to one of the prior learned patterns, if the human participant suddenly shifts the current movement pattern to another learned one, the synchronization breaks down thereby generating prediction error bottom-up in the RNNPB. In effort to minimize this error, the PB value is updated, and enactive synchrony with the present pattern achieved. Looking at the data from these experiments, segmentation of sensory flow from one habituated pattern (with corresponding expectations) to another can also be observed.

Observation of these phenomena suggests a general mechanism for the segmentation and chunking of the continuous sensory flow. Confirming a core tenet of predictive coding, that cognition aims to minimize prediction error in the process of interacting with changing environments, perceived error should be essential to



**Figure 2.** Dynamic generation and switching of two learned ball-handling behaviors. Top row: measured ball position. Second and third rows: predicted ball position and robot joint angles generated by the RNNPB, respectively. Bottom row: the PB as it switches from one movement sequence to another. Redrawn from (Ito et al., 2006).

mechanisms underwriting the segmentation of the continuous bottom-up sensory flow during online cognition in biological models, as well. Moreover, such mechanisms should provide for the development of compositionality in cognition, subject of the next section, because this competency requires segmentation of sensory flow into a set of reusable objects which can be mentally manipulated for combinatorial operations.

Here, it is intriguing to note that the error regression scheme for the past window in the RNNPB could provide a possible mechanism for “postdiction” (Shimojo, 2014). Postdiction is a process that is recognizable during perceptual phenomena wherein the percept of a stimulus presented earlier is affected by another stimulus presented later. Postdiction is apparent during various backward perception phenomena, including classic examples of backward masking (Raab, 1963) or the cutaneous rabbit illusion (Geldard & Sherrick, 1972). Such phenomena may be explained by the error regression mechanism assumed in the RNNPB as a model of predictive coding.

The cutaneous rabbit illusion involves tactile stimuli (taps, small electric shocks) presented to a human subject, typically on the forearm (due to this area’s relatively poor spatial acuity as mapped to the somatosensory cortex). In the simple case, three stimuli

are presented with equal temporal intervals between each. The cutaneous rabbit illusion appears when these stimuli occur with very short durations between them (less than 300 ms) and when the first and the second are given in the same position, but with the third given in a distant position (which may extend past the physical body, Miyazaki et al., 2010). The subject mislocates the second tap in the direction of the third tap, whereas the subject will not mislocate it if the third tap is given in the same position with the first or second ones.

In backward masking, the consciousness of a target presented immediately before a masking stimulus (typically something driving urgent attention, such as something scary) can be suppressed, such that subjects are unable to report having seen the first stimulus. This phenomenon may be intuitively explained, given two assumptions: one that the world as we experience it usually does not change so rapidly, and two that optimal implicit internal models of such a world should operate according to the expectation that a visual stimulus given a moment ago will be retained for a while. This gives rise to the idea that there may be two pathways operative in backward masking, one for normal operations over longer timespans, and one for surprising situations that respond rapidly to changes in the environment. In case that a stimulus presenting minor

changes in the world is suddenly followed by a surprising stimulus presenting more important changes in the world, an expectation error should be generated that effectively diminishes the previous stimulus as the agent recenters activity around the implications of the later stimulus. Thus, in the course of minimizing the error between what is expected and what is sensed, the experience of the stimulus presented in the past is “masked” by one that comes later.

The case of the cutaneous rabbit illusion can be explained similarly. After the first and the second (series of) taps are provided in the same position, an internal model implicitly expects that the third should come in the same position after the same temporal interval. However, when the third is presented in a distant position, an expectation error is generated. To minimize the error, the experience of the second tap is relocated to the midpoint between the first and the third tap positions (by means of regressing a linear model predicting the position and timing for succeeding stimuli). In sum, postdictive phenomena can be explained in terms of inference and consequent rewriting of past experience by means of the error minimization principle.<sup>3</sup>

### 2.3. Multiple timescale RNN for development of compositionality

The RNNPB experiments reviewed above demonstrate that cognitive phenomena such as segmentation of the undifferentiated perceptual flow into reusable chunks occur by minimizing conflicts between bottom-up perception and top-down intention. This section briefly reviews experiments intended to expose the role of such dynamics in the composition of novel patterns in coordination with novel task environments before turning to how these fundamental dynamics may contribute to accounts of self.

Compositionality is the ability to combine parts into wholes, evident, for example, in the abilities to determine the meanings of sentences from the structured relations between constituent parts (Costello & Keane, 2000; Evans, 1982) and to enact diverse goal-directed actions by sequentially combining reusable primitives (Arbib, 1981; Pastra & Aloimonos, 2012). Yamashita and Tani (2008) proposed a predictive RNN model characterized by multiple time constraints at different levels, the multiple timescale recurrent neural network (MTRNN) to investigate how neural networks in biological brains develop compositionality and thereby generate novel actions. The MTRNN has been used to investigate various aspects related to development of compositionality including co-development of skills between human tutors and robots (Nishimoto & Tani, 2009), goal-directed creative compositions of primitives (Arie et al., 2009), cases analogous to schizophrenic pathology including the delusion of control (Yamashita

& Tani, 2012), and imitative human–robot interactions (Hwang et al., 2018). Next, we review the MTRNN in greater detail.

**2.3.1. Model description.** The MTRNN has layers of RNNs each characterized by a different timescale constraint (see Figure 3(a) for a typical architecture). Neural activity in higher layers is slower with larger timescale constraints, while lower layers are faster with smaller timescale constraints. In the MTRNN, the following objective function for a time series of  $L$  time steps is minimized as analogous to equation (9) in the case of the RNNPB

$$Obj(c_1, W) = \frac{1}{2} \sum_{t=1}^L \left( \frac{1}{N^o \sigma^2} \|\bar{X}_t - X_t(W)\|^2 + \frac{1}{N^c} c_1^2 \right) \quad (11)$$

In equation (11),  $N^o$  and  $N^c$  represent the dimensions of the output and the internal units, respectively. Equation (11) differs from equation (9) in that  $c_1$ —which represents the potential value of the internal units with recurrence at the initial time step in equation (9)—plays the role of latent variable in equation (11).

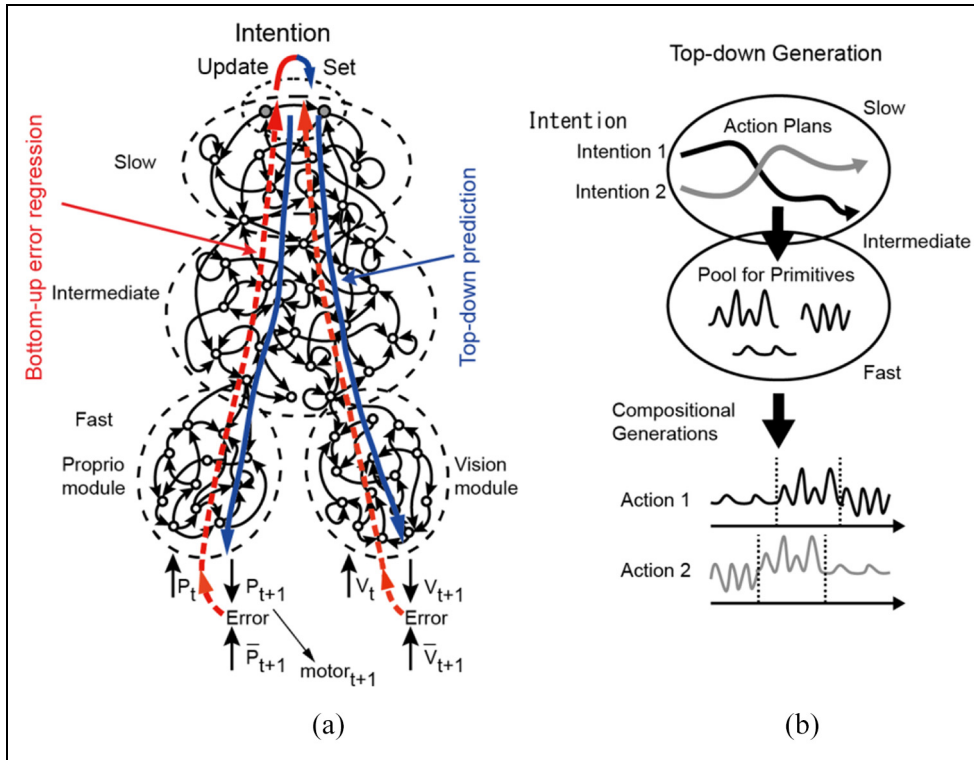
Assuming fixed learning parameters of  $W$ , this objective function can be minimized with respect to  $c_1$  by way of the following difference equations ((12a) to (12d))

$$\begin{cases} c_{t+1}^i = -(1 - \frac{1}{\tau})c_t^i + \frac{1}{\tau} \sum_j w_{ij}^c C_t^j + b_i^c & \text{for } t \geq 1 & (12a) \\ c_{1,n+1}^i = \alpha \left( -c_{1,n}^i - \nabla E(c_{1,n}^i) \right) + c_{1,n}^i & & (12b) \\ C_{t+1}^i = g(C_{t+1}^j) & \text{for } t \geq 0 & (12c) \\ X_{t+1}^o = g \left( \sum_j w_{oj}^{out} C_{t+1}^j + b_o^{out} \right) & & (12d) \end{cases}$$

In this series of equations,  $C_t$  represents activation of the internal units with recurrence at time step  $t$ .  $\tau^i$  is the time constant for each internal neural unit, and is different at each layer (in comparison, the time constant for all internal units in the RNNPB was assumed to be 1.0.) By assigning the higher layer with a larger value for the time constant, the higher layer is dominated by slower dynamics; by assigning a smaller time constant to the lower layer, it is dominated by faster dynamics. It is also noted that a bottleneck type connectivity constraint was applied to the network wherein the internal units in the higher layer were connected only with those in the intermediate layer which were again connected with only those in the lower layer.

Here, we may note that the latent variables represented as the initial states of the internal units  $c_1$  determine the time development of the whole network, including later prediction outputs, due to the sensitivity





**Figure 3.** (a) An example of an MTRNN architecture with three layers and (b) illustration of compositionality of different action plans in an MTRNN. Redrawn from (Tani, 2016).

to initial conditions that is characteristic of the deterministic RNN model. Therefore, for minimizing the reconstruction error for all time steps, the optimal values of the initial states are inferred by using the gradient information computed by BPTT through iterations of the internal computation loop using the past window in a similar way with the RNNPB case.

**2.3.2. Robotics experiment.** A robotics experiment using an MTRNN in a task of developmental tutoring of compositional object manipulation is briefly reviewed. During this experiment (Nishimoto & Tani, 2009), a Sony QRIO (as in Figure 1) controlled by an MTRNN was tutored on multiple task sequences each composed of different series of primitive actions. For example, one tutored sequence proceeded as follows. The tutor moved both of the robot’s hands toward an object located at an arbitrary position on a table, then using the hands grasped it, lifted it up and down a few times, and placed it back on the table. Another sequence involved touching the object with the left and right hands in turn, grasping the object, rotating it, and placing it back on the table. Sensors on the robot delivered simplified visual features and proprioceptive information (as in experiments described above). Tutoring proceeded gradually, that is, the robot was tutored on some tasks, then tested, and if performance was

unsatisfactory, tutoring was repeated (the corresponding video can be seen at <https://youtu.be/n9NYcG8xlyS>). It is important to emphasize that tutors directly guided the robot’s hands by feeling and correcting the “intent” of the robot with their own hands. After training, the robot was able to reliably perform all task sequences successfully and layer-specific neural regions were analyzed. It was found that each layer played a different role in action compositionality (see Figure 3(b)).

Yamashita and Tani (2008) speculated that the multiple timescale property imposed on network dynamics resulted in the emergence of a functional hierarchy in which the higher layer generated different slowly changing neural activation patterns corresponding to a scenario or plan for each task sequence, whereas the lower layer developed precise skillful control for each behavior primitive. Recalling experiments involving the RNNPB, the slowly changing neural activity from the higher layer served as a source of bifurcation parameters, steering activity in the lower layer, while the faster lower layer learned to develop a set of behavior primitives, for example, grasping, lifting up, or moving the object left and right, in complementary ways. After the development of these dynamics through learning, the MTRNN-driven robot became able to generate different ways of combining behavior primitives. Yamashita and Tani (2008) interpreted these results by

hypothesizing that the robot developed compositionality in generating actions as its functional hierarchy self-organized. To test this interpretation, they manipulated the timescale parameters of the MTRNN to explore the role of the multiple timescales in structuring cognitive dynamics. Compositional representations (layer-specific stable structural dynamics) could not be developed when all layers shared the same time constant. The same results were repeatedly confirmed in other robotics experiments conducted under more complex conditions, for example, while using pixel-level vision (Hwang et al., 2018).

#### 2.4. Goal-directed planning

With this understanding of compositionality as achieved through the segmentation of fluid experience, predictive coding and active inference (see Section 2.1) can be applied also to the problem of the goal-directed planning of action sequences by robots.

Tani (1996) conducted experiments on the goal-directed navigation of a mobile robot, Yamabico, which was equipped with a range sensor using a single-layer RNN. Yamabico was developed with a lower level automatic control scheme using range sensors which can perceive images of 24 angular directions covering the front of the body to travel smoothly in a collision-free manner in a workspace. Basically, it moves toward the largest open space in a forward direction by maneuvering between obstacles on its left and right sides. When a new open space appears, a decision is made on whether to pursue the current open space direction or to branch to the new one. This branching decision is made by the RNN in the higher cognitive level as described next.

In the experiment, Yamabico explored the obstacle workspace under collision-free maneuvering control during daytime for the purpose of gathering sensory-action data. When it encountered a branching point, a branching decision was made arbitrarily (with an equal chance for either option) by the experimenter. At this time, the sensory inputs in terms of the range image, the travel distance (as indicated on the odometer) from the previous branching point to the current one, and the action in terms of the branching decision were recorded. Yamabico explored the environment experiencing around 200 successive branches until its battery was depleted. This resulted in a sampling of a sensory-action sequence of around 200 branching steps. During nighttime while the battery was charging, the RNN was trained in the form of the forward model (Kawato, 1999; Miall & Wolpert, 1996) so that it could predict the sensory input at a next branching step  $X_{\tau+1}$  when presented with current sensory input  $X_{\tau}$  and branching action  $a_{\tau+1}$  by developing an adequate dynamic structure for the latent state transition from  $Z_{\tau}$  to  $Z_{\tau+1}$

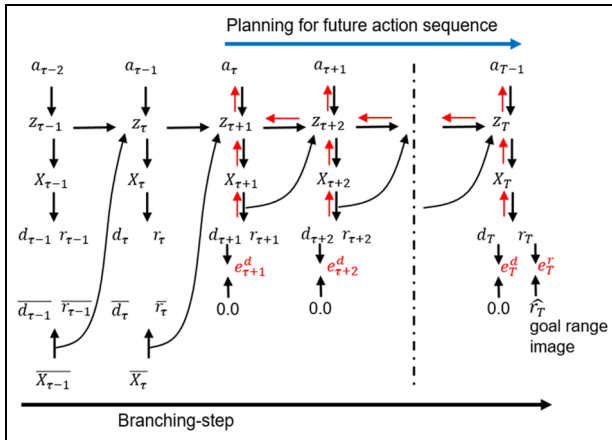
based on the branching scheme using the sampled sensory-action sequence.

After this training, a test of goal-directed planning was conducted through the following procedure. First, Yamabico traveled for several steps by randomly branching out from an arbitrary position for the purpose of inferring the latent state by way of the observed sensory sequence.<sup>4</sup> Several branching steps of travel were necessary because this navigation problem involves the sensory aliasing problem; the current sensory inputs cannot uniquely identify the current latent state. Then, after inference of the latent state in the current branching step, Yamabico generated goal-directed planning under the constraint of minimum travel distance to reach a branching point specified as a goal by its expected sensory inputs.

Analogous to active inference as in equation (7), an optimal action sequence minimizing the error between the preferred and the predicted sensory outcomes was identified, specifically minimizing the error between the goal range image (image of the goal as it is ideally achieved) and the predicted one in the distal (final) goal step (given the current action plan) while also minimizing the predicted travel distance at each branching step. This can be carried out by means of BPTT applied to the trained RNN as illustrated in Figure 4. In the test experiment, a set of possible action sequence plans was searched (including suboptimal ones) through iterative computation. Generated action plans were actually executed by the robot.

Results of goal-directed plan generation are shown in Figure 5. Figure 5(a) shows the designated starting position and the goal position. Figure 5(b) to (d) shows the actual trajectories generated by executing three different action sequence plans, with Figure 5(b) showing the optimal trajectory that minimizes the travel distance, and Figure 5(c) and (d) showing suboptimal plans. Especially, it is noted that the robot had never enacted the trajectory in Figure 5(d) during the exploration phase before learning. This result implies that Yamabico became capable of mentally imaging novel compositions of experienced parts of trajectories by extracting the hidden structure of the environment through consolidative learning of diverse sensory-action trajectories sampled.

To examine the internal structure developed by the RNN, a phase space analysis was conducted. The RNN generates mental simulations of thousands of consecutive steps of random branching sequences. Figure 5(e) shows the phase space plot wherein each point represents an internal state (projected in two-dimensional space) when visiting a particular branching point. It can be seen that the points are all clustered into a set of segments. It was found that each segment corresponds to a particular branching point, and that the points inside each segment form a Cantor set-like assembly. These observations inform us how the compositionality is



**Figure 4.** Goal-directed planning using RNN. The future action sequence  $a_{t \dots T-1}$  is optimized for minimizing the expected travel distance  $d_{t+1 \dots T}$  at each branching step as well as the error between the goal range image at the distal step  $\hat{r}_t$  and its expected value  $r_t$ . Both errors for the travel distance at each step  $e^d_{t+1 \dots T}$  and the goal range image at the distal step  $e^r_t$  are backpropagated.

represented in the internal state of the RNN. When the robot moves from one branching point to another by choosing a particular branching action, the internal state transits from one segment to another. Therefore, a graph-like state transition structure can be seen by considering those segments as nodes of a state transition diagram.

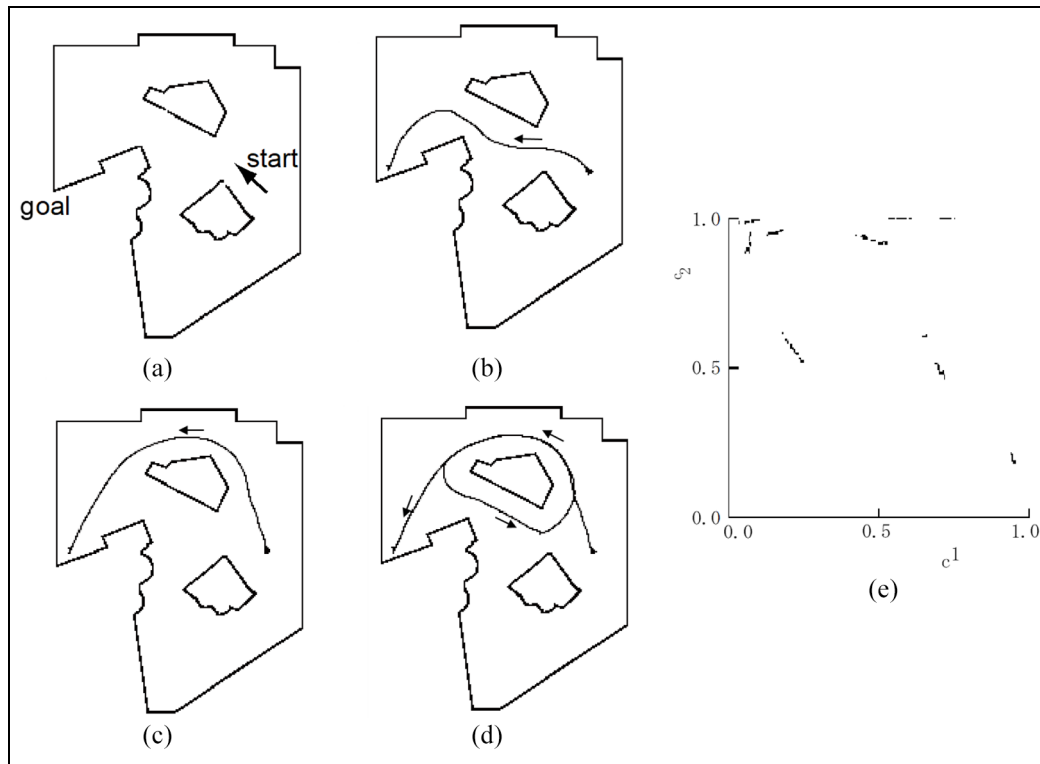
Importantly, each node is represented not by a point, but by an assembly of points in the continuous state space of the RNN, because each point within a segment can have a different history of branching in arriving there. If two points are neighboring each other, two branching sequences of reaching these two points may be exactly the same for a long past history. On the other hand, if two points are distant in the same segment, two branching sequences reaching these points will be quite different in the immediate past history. Since an infinite number of different branching sequences can be composed to reach a branching point, and all points corresponding to those different compositions should be embedded in a segment of finite length, Cantor set-like assembly is organized within each segment. Cantor sets are interesting in this context because they represent boundary points. They are perfect sets, meaning that a set is equal to its derivative set, which means that it is equal to its limit points. These findings are analogous to what Elman (1991) and Pollack (1991) showed in investigating the capability of RNNs to learn word or symbol sequences regulated by grammar.

Further analysis revealed that the whole assembly of points in the phase plot represents an invariant set of a global attractor. When mental simulation is perturbed with external noise added to the network activity, prediction goes wrong and the internal state falls out of

the invariant set.<sup>5</sup> However, if the noise is removed, the predictability can be recovered after several steps of mental simulation, and the trajectory of the internal state returns back to the invariant set. In this way, the invariant set may represent the boundary of cognition (Maturana & Varela, 1991) which is structurally stable against perturbation due to the nature of an attractor.

One drawback in this study (Tani, 1996) is that the exploration of the environment was conducted randomly, independent of the process of learning about the environment to satisfy a purpose. Human infants or artificial agents may explore the world by seeking some intrinsic rewards such as novelty in interacting with their environments (Oudeyer et al., 2007; Schmidhuber, 1991; Tschantz et al., 2020). Tani and Yamamoto (2002) extended the study in Tani (1996) to investigate the issue of intrinsic motivation by adding a novelty rewarding mechanism as a drive for seeking novel experiences during exploration. The exploration and the learning phases of this experiment were interleaved, with each happening after the other. In the exploration phase, action plans for branching sequences were generated and executed such that the sum of expected prediction error at each branching point was maximized (rather than minimized). And in the learning phase, the training of the network was conducted by using two types of sensory-action sequence data. One was the sequence which had been experienced in the last travel and which had been stored in the short-term memory, and the other was the set of rehearsed sequences generated by using the same planning scheme for maximizing the novelty (in which the sum of expected prediction error at each branching point was maximized). By extending the model thusly, Yamamoto and Tani showed that compositionality—in terms of the number of different combinations of branching—increases both in the physically generated trajectories and in the rehearsed sequences during the learning phase over the course of development. Finally, a phase space analysis of the internal state at each stage of development indicated that a segment-wise invariant set similar to the one shown in Figure 5(e) appeared, but only during the end period of the development, when the robot had completely learned about all possible branching consequences. Such dynamics illustrate how the aforementioned boundary of cognition might emerge during developmental processes similarly motivated by exploration in biological models including human beings, as well.

Recently, Friston and colleagues (R. Kaplan & Friston, 2018) took a generic Bayesian perspective on the exploration–exploitation trade-off by considering that the expected free energy is composed of novelty, salience, and prior preferences which together constitute optimal beliefs on policies guiding action. On this scheme, novelty and salience represent intrinsic values where novelty is for epistemic gain, as described



**Figure 5.** Trajectories generated by goal-directed planning and the phase plot. (a) Starting and goal positions, (b–d) three different trajectories generated by Yamabico based on different action plans, and (e) the phase plot of the internal state projected in a two-dimensional space.

previously, salience is for the gain of certainty about the latent state, and prior preferences serve the role of extrinsic values, such as for achievement of predefined rewards or goals. As such, R. Kaplan and Friston (2018) assume that actions are selected both to gain knowledge about the world and to achieve predefined preferences. One point of interest in this context, however, is the relatively high computational cost required to optimize such action plans. In reality, it is not necessary to always compute optimal action plans, since action can typically be generated according to the learned routine, or habit. On this point, Maisto et al. (2019) proposed that an active inference agent caches the probabilities of policies from previous trials in memory as habits to reduce the computational costs for re-calculating them at each new trial. Simply put, a probability is only re-calculated when encountering a new context, and is then kept cached as long as the context does not change. Testing this sort of idea in robots should be practically beneficial in, for example, reducing real-time computational burdens of robots in operational contexts.

### 3. The sense of self

Building from the preceding introduction, this section speculates on how such cognitive robotics experiments

employing fundamental principles may contribute to ongoing inquiry into the sense of self and related psychological phenomena. Although the problem of self has been addressed in various ways, such as Neisser's (1988) five different types of selves or Kohut's (2013) bipolar self from psychoanalysis, the current article focuses on two distinct types of self, minimal self and narrative self. This section begins with Gallagher's (2000) concept of the minimal self, and then examines how his notions may correspond to the phenomena observed in CNR experiments such as those introduced in the preceding section.

#### 3.1. The minimal self

Gallagher (2000) argues that after all the unessential features of experience are stripped away, we still have a feeling of a basic, immediate, or primitive "something" that we can call the "minimal self." He further contemplates that this sort of non-reflective self is associated with two different types of senses, one is a sense of ownership and the other is a sense of agency. According to Gallagher (2000), the sense of ownership is the sense that I am the one who is undergoing an experience. For example, a sense that this is my body moving regardless of whether the movement is caused by me or others. The sense of agency, on the other hand, refers to

congruence between an agent's intention or belief in an action and its anticipated outcome, which endows the agent with the sense that "I am the one generating this action."

Both cases may be explained in terms of internal models for predicting perceptual outcomes. For example, Hohwy (2013) showed that the sense of body ownership in the rubber hand illusion (Botvinick & Cohen, 1998) can be explained by using predictive coding that models the probabilities of the dummy hand being mine or another's. During the rubber hand illusion, experience of the temporally synchronized multimodal sensation of touching, one from a tactile stimulus and the other from visual observation, results in a causal inference that these two occur at the same location even though they actually do not. This is because the temporal correlation entails more dominant effects on the inference than the spatial one does (Hohwy, 2013; Limanowski & Blankenburg, 2013). This leads to the illusion of the dummy hand being mine.

Concerning the sense of agency, Gallagher (2000) suggests that a possible underlying mechanism may be conceivable by considering neurocognitive models accounting for some cases of schizophrenia such as discussed by Feinberg (1978) and Frith (1992). They proposed that the delusion of control as a characteristic of schizophrenia may occur when a mismatch takes place between an intended state and the anticipated state produced by the forward model. As the forward model is informed by the motor efference copy, they proposed that the mismatch may be caused by either the failure of the forward model or the fact that the efference copy cannot be sent to the forward model because the motor controller is disconnected from it. Gallagher (2000) suggested that the sense of agency, which remains implicit in a normal condition, can be disturbed in such a case, resulting in the feeling that "somebody is controlling me" which is common to self-reports in some cases of schizophrenia.

The RNNPB robotics experiments (reviewed in Section 2) may help to account for the emergence of minimal self accordingly. Minimal self should develop implicitly as an aspect of the sense of agency derived from causality between an agent's intention driving action (encoded in the PB) and the affected perceptual reality, similarly to the sentiment expressed in Hohwy (2007, p. 5):

mineness is the feeling of already being familiar with the movement's sensory consequences when they actually occur, we are so to speak already "at home" in the movement because the incoming signals are predicted through habituation, and therefore it is regarded as an implicit sense of self.

Accordingly, when a robot's prediction was accurate, action proceeded smoothly and automatically without distinction between synchronized embodied self and the

external objects with which it was interacting, such as a ball or human counterpart.

Our proposal in the context of minimal self is that, when such synchrony breaks down due to miscellaneous unpredictable influences including noise in the physical system or a human participant's sudden intentional change, the otherwise implicit sense of minimal self becomes an object of consciousness. This is because the consequent effort required to minimize the reconstruction error in the immediate past window, by inferring an optimal intention state in the PB, is accompanied by a focal awareness of the gap between embodied routines and the capacity for embodied routines to successfully meet environmental demands. At this very moment of the unified structure breaking down, the independence of each element becomes noticeable. Consequently, this experience of self is formally articulated as *minimal self-consciousness* (Tani, 1998; 2016, pp. 169–172).

Moreover, on this account, the sense of minimal self should intermittently shift between an unconscious phase (predictable phase, when intention guides action without undue error) and a conscious phase (unpredictable phase, when error forces reformulation of intentions guiding action going forward) as had been observed in a vision-based robot in Tani (1998), as well as in the humanoid robot experiments involving ball handling described in Section 2. Why do those system dynamics once converged into an attractor basin, such as a predictable or routine interaction in a region, get destabilized again, and move out to another basin of attractor? One possibility is the inherently indeterministic nature of embodied cognitive systems due to the circular causality established in the enactment loop (Tani, 1998; 2009).

Circular causality describes the embedded and embodied agent's situation. An agent acts on the world, and a sequence of causes and effects returns back to the original cause, possibly altering it, whereby another sequence is produced in an ongoing interactive feedback loop. For example, in the case of the ball-handling humanoid robot, when a prediction error is generated for the ball position, the intention of the robot in terms of the PB value is updated in the direction of minimizing the error. This intention generates the next-step prediction of proprioception which, in the case of tracking a ball, turns out to be new target joint angles. These angles are fed into the robot motor controllers. Then, both hands of the robot move to push the ball further, for example.

When every process in this loop proceeds ideally, the whole system dynamics stay always in the same attractor basin by successfully minimizing prediction error, unless exceptionally large noise comes from the external world. However, in reality, under resource-bounded situations, this cannot be guaranteed. Inherent indeterminism will appear. The predictability of the neural

network is limited because it is trained with only a finite amount of sensory-motor experience. The inference of PB values cannot be guaranteed to be optimal in a real-time situation, and physical movements of the robot body as well as the resultant movement of the ball should contain some margin of unpredictability because of nonlinearity in both static and dynamic friction and contact dynamics. All these contingencies (to a very large degree) are due to embodiment. They provide potential instability to the system dynamics.

In fact, two flows in opposing directions coexist. One is the inflow converging to attractors or thermodynamic equilibria by minimizing prediction error (or free energy), and the other is the outflow destabilizing the convergence by means of embodiment and circular causality. In a macroscopic sense, the coexistence of inflow and outflow in the phase space makes attractors only marginally stable, wherein the state trajectories tend to visit multiple pseudoattractors one by one itinerantly (known as chaotic itinerancy, Kaneko, 1990; Tsuda et al., 1987). This may correspond to the “momentary self” contemplated by William James (1890) who wrote that “When we take a general view of the wonderful stream of our consciousness, what strikes is the pace of its parts. Like a bird’s life, it seems to be an alternation of flights and perchings” (p. 243).<sup>6</sup>

Such unsteady dynamics resulting from potential indeterminism provide an inherent autonomy to the minimal self in terms of their spontaneous shifts between their unconscious phase (staying inside basins of attractors or habitual regions) and conscious phase (transition to another attractor passing through less familiar regions). Froese and Taguchi (2019) present an analogous argument that artificial as well as living agents may make sense of their interactions with the world, provided that there is some room for indeterminism or incompleteness in the causal closure of these interactions. One difference, however, between Froese and Taguchi’s (2019) and our consideration is that they attribute the origin of indeterminacy to quantum mechanics at the micro level, which is supposed to be amplified through the enactment loop at the macroscopic level. On the contrary, we presume that the sensitivity to the initial state caused by chaos, or the structural instability observed in chaotic itinerancy, may account for the origin of indeterminacy without resorting to (what currently remain) mystic propositions.

### 3.2. Narrative self

Gallagher (2000) considers narrative self as “a more or less coherent self (or self-image) that is constituted with a past and a future in the various stories that we and others tell about ourselves” (p. 15). He contrasts two distinct ways of representing such a sense of narrative self. One is offered by Dennett Daniel (1992) who characterizes self as the constant locus of experience and

center of “narrative gravity.” The other is a more distributed model inspired by Paul Ricoeur’s (1984) philosophy of narrative self. Ricoeur considers a hermeneutic cycle of movements from *prefiguration* of phenomena in the world to their *refiguration* or restoration back into the real world through (communicative) action via *configuration* of interpreted narratives, with configuration of narratives playing the role of mediation between prefiguration and refiguration, and the three together constituting a process through which the agent gains a better understanding of its self and its place in the social and natural world. Especially, he considers the human experience of *aporia*, when phenomena in the world are experienced as incomprehensibly contradictory. Ricoeur emphasizes that human beings compose fictive as well as true narratives to reconcile this feeling of *aporia*. Furthermore, he considers that one’s own self-narratives are configured in a way that they are intermingled with those communicated by others. Following Ricoeur, Gallagher (2000) proposes that narrative self might be developed as a mixing of stories about one’s self, including conflictive and irresolvable ones which an individual might tell about herself or himself or others might tell about her or him (cf. “pernicious misunderstandings” in Gallagher & Allen, 2018, pp.14–15). Thus, on Gallagher’s account, the sense of narrative self might be considered a center of narrative gravity that is more distributed, representing the reconciliation of narratives normalized in communication with others in interaction with the shared object environment.

We can find some analogy with what Ricoeur and Gallagher consider narrative, including this more distributed sense of narrative self, in the results of the CNR experiments described previously. In effect, the robots used in experiments by Tani and colleagues were frequently confronted with incomprehensible and irresolvable situations during tutoring. For example, the humanoid robot implemented with the MTRNN (see Section 2.3) was tutored to generate inconsistent primitive movement sequences; after grasping an object, it was tutored to lift the object up in one instance, and in another it was tutored to rotate the same object. The mobile robot Yamabico was tutored to branch to the left during one trial, and during another to branch right at the same branching point. Each tutoring trajectory never repeated exactly the same sensory-motor sequence pattern as another, even though the tutors attempted to do so, due to noise and fluctuation—indeterminacy (cf. Froese & Taguchi, 2019; Tani, 1998)—inevitable during embodied interaction with the physical world. In light of Gallagher and Ricoeur’s insights into narrative self, we may say that these robots could develop compositionality for generating diverse actions because it was necessary to deal with the inconsistency presented in the tutoring trajectories. Through consolidative learning, the RNNs could self-

organize finite-state machine-like state transition structures in their latent state space by extracting relational structures from among the tutored set of inconsistent sequences.

The segmentation and chunking of the continuous sensory flow observed in robotic experiments using the RNNPB and MTRNN are mechanized by means of the error regression employed to minimize the error signal generated in prediction and reconstruction at the moment of transition from one primitive to another. The inconsistency brought to the robots especially by human tutors in demonstration and tutoring of movement patterns plays an important role here, also. If we tutor the robots with a continuous sensory flow consisting of a primitive sequence like  $A \rightarrow B \rightarrow C$ , repeated without inconsistencies, this sensory flow will be learned as a big chunk without segmentation into primitives. But, if the robots are tutored with a set of sensory flows consisting of inconsistent primitive sequences like  $A \rightarrow B \rightarrow C$ ,  $A \rightarrow B \rightarrow A$ , and  $A \rightarrow C \rightarrow B$ , these sensory flows are learned as compositions of reusable primitives with each of them segmented. We presume that the cognitive competency of compositionality for segmenting the continuous sensory flow into a set of reusable primitives and composing and decomposing a whole action sequence using these primitives can provide a basis for development of the narrative self.

Furthermore, some robots (or more generally agents; R. Kaplan & Friston, 2018; Oudeyer et al., 2007; Schmidhuber, 1991; Tani & Yamamoto, 2002) are motivated to learn to predict unpredictable situations by seeking novelty, with the objective of which also seeming rather contradictory or conflictive at first glance. And, to reconcile such conflicts, agents seem to be required to generate creative or fictive mental images accounting for the hidden causal structure in the world. Here, we might note some analogy between the compositionality developed by means of a self-organizing finite-state machine-like structure in distributed neural activation in the RNN during such conflictive situations, and the sense of narrative self which develops distributedly, with the mixing of diverse inputs, including inconsistent ones from the outside as described above. The compositionality developed in the robots enables them to mentally simulate future actions, including counterfactual or fictional ones, as well as to rehearse what has been experienced in the past to prepare for the uncertain future.

At this point, it is natural to consider that the development of narrative requires linguistic competency for telling stories. On this front, there have been some efforts in attempting to ground linguistic expressions in sensory-motor modalities by using RNNPB and MTRNN architectures by some groups (Heinrich & Wermter, 2018; Peniak et al., 2011; Sugita & Tani, 2005; Yamada et al., 2016). With a vision-based mobile

robot using the predictive coding framework, Sugita and Tani (2005) showed that an RNNPB can learn to bind a set of simple imperative sentences consisting of verbs and nouns, for example, Point-Red, Push-Blue, Hit-Green, with corresponding sensory-motor behavioral patterns. An analysis on the experimental results of learning and action generation for given imperative sentences showed that the compositionality in combining verbs and nouns in the linguistic modality and the one in combining actions and objects developed as a unified structure in the RNNPB. Peniak et al. (2011) and Heinrich and Wermter (2018) showed scaling of such language-behavior binding using extended MTRNN models. Also extending the MTRNN, Yamada et al. (2016) presented a continuing human-robot interaction experiment using both linguistic and behavioral modalities. These experimental results revealed that the contextual flow corresponding to successive human-robot interactions was represented in the higher level latent variables in the MTRNN. These results suggest that the internal structures developed in the latent space in these RNN models, via continuing human-robot interaction using both linguistic and behavioral modalities, bring us *closer* to realizing a narrative self as articulated by Ricoeur and by Gallagher and colleagues in an embodied cognitive neurorobot.

#### 4. Discussion

The preceding article introduced CNR, the principles of prediction error minimization, and backpropagation as implemented in different RNN architectures; related these with free energy and active inference; and reviewed selected CNR experiments employing these principles in greater detail. In terms of Marr's (1982) three levels introduced in Section 1, the error minimization principle appears at the computational level, prediction and active inference at the algorithm level, and RNNs embodied by robots at the implementation level as an example. This leaves open questions, for instance, how these implementations can be validated, and then, to what extent we may expect them to contribute to scientific explanation of cognitive phenomena including the senses of minimal and narrative self as reported by human beings.

First, in the review of the CNR experiments using the RNNPB (Ito et al., 2006; Ito & Tani, 2004), we explained that the RNNPB learns multiple behaviors in the course of prediction error minimization as embedded in different attractor basins which represent habitual regions for the robot. It was also explained that behavior patterns of the robot shift from a learned one, by means of the error regression accompanied by segmentation of continuous sensory-motor flows, to another due to either external forces or internal fluctuation.

Next, in the review of the CNR experiments using the MTRNN (Nishimoto & Tani, 2009; Yamashita & Tani, 2008), it was explained that compositionality as a cognitive competency for composing/decomposing the whole action from/into behavior primitives can develop gradually in the course of iterative tutoring of the robot by using the error minimization principle. We explained that such compositionality can develop by means of self-organization of functional hierarchy using the prior constraints applied to the network, including layer-wise timescale differences and information bottlenecks in the connectivity between layers.

In the review of experiments on goal-directed planning in robot navigation tasks, it was explained that robots can learn compositional structures latent in the obstacle environment through either supervised tutoring (Tani, 1996) or self-exploration (Tani & Yamamoto, 2002). Further, it was noted that such compositional structures develop by self-organizing global attractors of Cantor set-like assembly in the latent state phase space. Although its appearance seems analogous to finite-state machines at first glance, these two are crucially different. As for the attractors developed in the latent state phase space, they represent the boundary of cognition (Maturana & Varela, 1991) wherein prediction goes well as habituated within the invariant set of the attractors, and prediction goes wrong once the state trajectory goes out of the invariant by possible permutation. However, the state trajectory can come back to the invariant set as long as it is formed as a global attractor. On the other hand, in the case of a finite-state machine, there is no mechanism for such auto-recovery unless some external programs for this purpose are provided. The CNR experiments reviewed above demonstrate that robots can generate both fictional and factual action plans by using compositional structures developed in both goal-directed planning and novelty rewarding schemes.

Further analysis contemplated possible accounts for subject experience including senses of minimal and narrative self. First, let us revisit our account of the sense of minimal self. When action goes smoothly and an agent remains in habitual regions by minimizing the prediction error, the sense of minimal self is present but only implicitly. However, a breakdown of such a steady phase comes inevitably because of the inherent indeterminacy due to the circular causality established in the enactment loop. In such an instance, the minimal self should become an object of conscious awareness with the effort to return from unfamiliar regions to a routine one by minimizing prediction error. We see the structure of the minimal self in this autonomy of spontaneous shifts between unconscious and conscious phases analogous to James' (1890) "wonderful stream of our consciousness."

Next, let us revisit the sense of narrative self. We found a good analogy in the results of the CNR experiments with what Ricoeur and Gallagher's socially

distributed sense of narrative self. Robotic experiments using the RNNPB and MTRNN showed that compositionality can be naturally developed, provided that the robots are tutored with a set of inconsistent sensory-motor sequences, corresponding with Ricoeur's thought that humans compose both fictive and true narratives in the process of resolution of aporia. Indeed, the CNR experiments reviewed in the current article showed that these robots can generate both fictive and factual compositions of primitive action both in physical execution and in mental planning and rehearsing. And, by briefly introducing the ongoing research on embodied language using various RNN models, discussion extended to possibilities of how such narratives initially represented in distributed neural activation patterns can be transformed into linguistic representations for sharing stories.

Although space forbids the present review, other ongoing work "breaks" these and complimentary architectures (cf. Glennan, 2005), tests them against biological models in similar abnormal conditions, and in this way aims to inform accounts of psychiatric conditions, for example, schizophrenia and autism understood as self-disturbances (Idei et al., 2018; Yamashita & Tani, 2012). It is expected that such research contributes to our understanding of otherwise difficult to resolve cognitive phenomena in two directions. For one, emergent dynamics when analogous to biological model behavior may inform researchers working at lower levels of organization about how different operations may be related, for example, temporally. In such an instance, prediction error minimization inspires neural network algorithm design (backpropagation, error regression, etc.). With successful demonstration of functional dynamics in a real-world context, such a synthetic architecture can be correlated with biological models (e.g. the higher, intermediate, and the lower level correspond to the prefrontal cortex (PFC), parietal, and S1 + V1 + M1) in exploration of possible explanations for (normal and abnormal) biological structural dynamics. An example of such work includes the HBP neurorobotics group's effort to render a computational model of an embodied mouse that may replace biological models in psychological studies (Falotico et al., 2017).

On the other hand, the neurorobotics studies reviewed in this article target invariant structures arising in what Sun et al. (2005) call the "causal nexus" between top-down and bottom-up processes, for example, selves, internal world models. Informed by phenomenological and neurocognitive research, these studies aim to contribute to explanations in cognitive science by articulating architectures which generate target emergent phenomena through their dynamic interaction with the world. In such an experiment, for example, the prediction error minimization principle shapes the cognitive architecture according to biological and psychological constraints. This architecture is tested in robot experiments, and emergent phenomena are



recorded. Correspondences to phenomenology (e.g. minimal self), neuroscience (e.g. mirror neurons), and psychiatry (e.g. autism) are considered. Limitations of the model inform ongoing inquiry, and insights drawn from these studies are offered.

Although physical robot experiments might be replaced by simulations in some contexts, it must be stressed that embodied humanoid robots are important to the success of CNR research such as that reviewed in Section 2. Practically, the results of physical robot experiments are more robust than simulated variants. Because simulation experiments typically employ thousands of trials that establish optimal parameters within narrow ranges, the resulting networks often become too rigid to perform in real-world embodied robots because robots themselves are quite noisy (in the informational sense) in their mechanics and physical interactions. With this in mind, we may point to an epistemic upshot to embodied robotics experiments over simulated variants. Humanoid CNR experiments open phenomena to investigation which remain inaccessible to simulations, especially those which emerge through direct human interaction as in the experiments reviewed above. During these experiments, trajectories of every learned movement were co-developed by a tutor and robot through their interaction. Accordingly, we speculate that successful results were achieved not only due to essential structural dynamics captured by biologically inspired cognitive architectures, but also due to the intuitive interaction with both tutor and humanoid robot aiming to minimize error in the embodied manipulation of common objects—including each other—in a shared space of action. Such phenomena cannot be (easily) simulated. And, given the fundamental role of interaction with others in the development of self (cf. Bolis & Schilbach, 2020), the socially situated nature of embodied humanoid CNR experiments presents special potential for ongoing inquiry into phenomena associated with self in human beings.

## 5. Conclusion

Guided by the intuition that higher level cognitive phenomena including different senses of self should emerge from the effort to minimize conflicting interactions between top-down and bottom-up information processes, Tani and colleagues have been refining cognitive neurodynamic models since the mid-1990s to articulate structural dynamics native to living systems in artificial ones. The preceding article reviewed synthetic neurorobotics experiments using analog devices (RNNs) directly sharing the same analog metric space with human beings with the expectation that such continuous spatio-temporal dynamics can both avoid the notorious symbol grounding problem (Harnad, 1990)

as well as inform our understanding of cognitive phenomena such as self in human beings.

## Declaration of Conflicting Interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

## Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This study has been partially supported by Grant-in-Aid for Scientific Research (A) in Japan, 20H00001, “Phenomenology of Altered Consciousness: An Interdisciplinary Approach through Philosophy, Mathematics, Neuroscience, and Robotics.”

## ORCID iD

Jun Tani  <https://orcid.org/0000-0002-9131-9206>

## Notes

1. It is interesting in this context to consider Bruineberg et al. (2018) who modeled foraging agents learning most efficient paths to prior-specified goals, “desire paths.” These models extended the free energy minimization principle (FEP) and active inference to show that expectations are shaped by changes to the environment, resulting in reduced free energy for the agent–environment system.
2. In the interests of simplicity, we neglect here discussion of merits of variational Bayes models such as those representing the strength of belief or an estimate of precision in prediction. Recent developments using recurrent neural network (RNN) models by the author’s group (Ahmadi & Tani, 2019; Murata et al., 2015) address these issues.
3. Here, we may understand “inference” specifically, as search for an optimal latent variable to minimize the error between the expected sensation and the actual one.
4. This process can be related to the process of perception by predictive coding shown in (5). However, in this case, the latent variable is not inferred by minimizing prediction error. Instead, the latent variable is updated by means of the entrained (learned) sensory input sequence which turns out to minimize the prediction error.
5. The network loses its “grip” cf. Bruineberg et al. (2014).
6. In terms of compositionality, this image also recalls Plato’s aviary from the Theaetetus, in which a person collects birds in the cage of the mind (where they may be imagined to flit from perch to perch), representing expressions of knowledge that can be taken when the knowledge is useful, or mistaken when the knowledge is not, cf. discussion beginning 197d, Plato Cooper and Hutchinson (1997, p. 218).

## References

- Ahmadi, A., & Tani, J. (2019). A novel predictive-coding-inspired variational RNN model for online prediction and recognition. *Neural Computation*, *31*, 2025–2074.

- Arbib, M. A. (1981). Perceptual structures and distributed motor control. In V. B. Brooks (Ed.), *Handbook of physiology, the nervous system II. Motor control* (pp. 1448–1480). American Physiological Society.
- Arie, H., Endo, T., Arakaki, T., Sugano, S., & Tani, J. (2009). Creating novel goal-directed actions at criticality: A neuro-robotic experiment. *New Mathematics and Natural Computation*, 5(1), 307–334.
- Asada, M., Hosoda, K., Kuniyoshi, Y., Ishiguro, H., Inui, T., Yoshikawa, Y., & . . . Yoshida, C. (2009). Cognitive developmental robotics: A survey. *IEEE Transactions on Autonomous Mental Development*, 1(1), 12–34.
- Baltieri, M., & Buckley, C. L. (2017). An active inference implementation of phototaxis. In *Artificial Life Conference Proceedings* (pp. 36–43). MIT Press.
- Baltieri, M., & Buckley, C. L. (2018). A probabilistic interpretation of PID controllers using active inference. In *International Conference on Simulation of Adaptive Behavior* (pp. 15–26). Springer.
- Barandiaran, X. E., & Chemero, A. (2009). Animats in the modeling ecosystem. *Adaptive Behavior*, 17(4), 287–292.
- Beer, R. D. (2000). Dynamical approaches to cognitive science. *Trends in Cognitive Sciences*, 4(3), 91–99.
- Bolis, D., & Schilbach, L. (2020). ‘I Interact Therefore I Am’: The self as a historical product of dialectical attunement. *Topoi*, 39(3), 521–534.
- Botvinick, M., & Cohen, J. (1998). Rubber hands “feel” touch that eyes see. *Nature*, 391(6669), 756–756.
- Brooks, R. A. (1990). Elephants don’t play chess. *Robotics and Autonomous Systems*, 6, 3–15.
- Brooks, R. A. (1991). Intelligence without representation. *Artificial Intelligence Journal*, 47, 139–159.
- Bruineberg, J., Rietveld, E., Bruineberg, J., Rietveld, E., Bruineberg, J., & Rietveld, E. (2014). Self-organization, free energy minimization, and optimal grip on a field of affordances. *Frontiers in Human Neuroscience*, 8, Article 599.
- Bruineberg, J., Rietveld, E., Parr, T., van, M. L., & Friston, K. J. (2018). Free-energy minimization in joint agent-environment systems: A niche construction perspective. *Journal of Theoretical Biology*, 455, 161–178.
- Cangelosi, A., & Schlesinger, M. (2015). *Developmental robotics: From babies to robots*. MIT Press.
- Clark, A. (2015). Radical predictive processing. *The Southern Journal of Philosophy*, 53, 3–27.
- Costello, F., & Keane, M. (2000). Efficient creativity: Constraint-guided conceptual combination. *Cognitive Science*, 24(2), 299–349.
- Dennett Daniel, C. (1992). The self as a center of narrative gravity. In F. Kessel, P. Cole, & D. Johnson (Eds.), *Self and consciousness: Multiple perspectives* (pp. 275–288). Lawrence Erlbaum.
- Doya, K., Samejima, K., Katagiri, K. I., & Kawato, M. (2002). Multiple model-based reinforcement learning. *Neural Computation*, 14(6), 1347–1369.
- Elman, J. L. (1991). Distributed representations, simple recurrent networks, and grammatical structure. *Machine Learning*, 7(2–3), 195–225.
- Evans, G. (1982). *The varieties of reference*. Clarendon Press.
- Falotico, E., Vannucci, L., Ambrosano, A., Albanese, U., Ulbrich, S., Vasquez, T. J. C., Hinkel, G., & . . . Gewaltig, M. O. (2017). Connecting artificial brains to robots in a comprehensive simulation framework: The neurorobotics platform. *Frontiers in Neurobotics*, 11, Article 2.
- Feinberg, I. (1978). Efference copy and corollary discharge: Implications for thinking and its disorders. *Schizophrenia Bulletin*, 4(4), 636.
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions: Biological Sciences*, 360(1456), 815–836.
- Friston, K., Mattout, J., & Kilner, J. (2011). Action understanding and active inference. *Biological Cybernetics*, 104(1–2), 137–160.
- Frith, C. D. (1992). *The cognitive neuropsychology of schizophrenia*. Psychology Press.
- Froese, T., & Taguchi, S. (2019). The problem of meaning in AI and robotics: Still with us after all these years. *Philosophies*, 4(2), 1–14.
- Froese, T., & Ziemke, T. (2009). Enactive artificial intelligence: Investigating the systemic organization of life and mind. *Artificial Intelligence*, 173(3–4), 466–500.
- Gallagher, S. (2000). Philosophical conceptions of the self: Implications for cognitive science. *Trends in Cognitive Sciences*, 4(1), 14–21.
- Gallagher, S., & Allen, M. (2018). Active inference, enactivism and the hermeneutics of social cognition. *Synthese*, 195(6), 2627–2648.
- Geldard, F. A., & Sherrick, C. E. (1972). The cutaneous “rabbit”: A perceptual illusion. *Science*, 178(4057), 178–179.
- Glennan, S. (2005). Modeling mechanisms. *Studies in History and Philosophy of Biological and Biomedical Sciences*, 36, 443–464.
- Harnad, S. (1990). The symbol grounding problem. *Physica D: Nonlinear Phenomena*, 42(1-3), 335–346.
- Heinrich, S., & Wermter, S. (2018). Interactive natural language acquisition in a multi-modal recurrent neural architecture. *Connection Science*, 30(1), 99–133.
- Hohwy, J. (2007). The sense of self in the phenomenology of agency and perception. *Psyche*, 13(1), 1–20.
- Hohwy, J. (2013). *The predictive mind*. Oxford University Press.
- Holland, O. (2007). A strongly embodied approach to machine consciousness. *Journal of Consciousness Studies*, 14(7), 97–110.
- Hwang, J., Kim, J., Ahmadi, A., Choi, M., & Tani, J. (2018). Dealing with large-scale spatio-temporal patterns in imitative interaction between a robot and a human by using the predictive coding framework. *IEEE Transactions on Systems, Man, and Cybernetics: Systems*, 50, 1918–1931.
- Idei, H., Murata, S., Chen, Y., Yamashita, Y., Tani, J., & Ogata, T. (2018). A neurobotics simulation of autistic behavior induced by unusual sensory precision. *Computational Psychiatry*, 2, 164–182.
- Iizuka, H., & Di Paolo, E. A. (2007). Toward Spinozist robotics: Exploring the minimal dynamics of behavioral preference. *Adaptive Behavior*, 15(4), 359–376.
- Ito, M., Noda, K., Hoshino, Y., & Tani, J. (2006). Dynamic and interactive generation of object handling behaviors by a small humanoid robot using a dynamic neural network model. *Neural Networks*, 19, 323–337.
- Ito, M., & Tani, J. (2004). On-line imitative interaction with a humanoid robot using a dynamic neural network model of a mirror system. *Adaptive Behavior*, 12(2), 93–115.

- James, W. (1890). *The principles of psychology: In two volumes*. Holt.
- Kaneko, K. (1990). Clustering, coding, switching, hierarchical ordering and control in a network of chaotic elements. *Physica D*, *41*, 137–172.
- Kaplan, D. M. (2015). Moving parts: The natural alliance between dynamical and mechanistic modeling approaches. *Biology and Philosophy*, *30*(6), 757–786.
- Kaplan, R., & Friston, K. J. (2018). Planning and navigation as active inference. *Biological Cybernetics*, *112*(4), 323–343.
- Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*, *9*(6), 718–727.
- Knoll, A., & Gewaltig, O. (2016). Neurorobotics: A strategic pillar of the Human Brain Project. In S. Sanders, & J. Oberst (Eds.), *Brain inspired intelligent robotics: The intersection of robotics and neuroscience* (pp. 25–34). Science/AAAS.
- Kohut, H. (2013). *The analysis of the self: A systematic approach to the psychoanalytic treatment of narcissistic personality disorders*. University of Chicago Press.
- Kotseruba, I., & Tsotsos, J. K. (2018). 40 years of cognitive architectures: Core cognitive abilities and practical applications. *Artificial Intelligence Review*, *53*, 17–94. <https://doi.org/10.1007/s10462-018-9646-y>
- Krichmar, J. L. (2018). Neurorobotics — A thriving community and a promising pathway toward intelligent cognitive robots. *Frontiers in Neurobotics*, *12*, Article 42. <https://doi.org/10.3389/fnbot.2018.00042>
- Krichmar, J. L., & Edelman, G. M. (2002). Machine psychology: autonomous behavior, perceptual categorization and conditioning in a brain-based device. *Cerebral Cortex*, *12*(8), 818–830.
- Kuniyoshi, Y., & Sangawa, S. (2006). Early motor development from partially ordered neural-body dynamics: Experiments with a cortico-spinal-musculo-skeletal model. *Biological Cybernetics*, *95*(6), 589–605.
- Lang, C., Schillaci, G., & Hafner, V. V. (2018). A deep convolutional neural network model for sense of agency and object permanence in robots. In *8th Joint IEEE International Conference on Development and Learning and Epigenetic Robotics (ICDL-EpiRob)* (pp. 260–265). IEEE.
- Lanillos, P., & Cheng, G. (2018). Adaptive robot body learning and estimation through predictive coding. In *IEEE International Conference on Intelligent Robots and Systems* (pp. 4083–4090). IEEE.
- Levesque, H., & Lakemeyer, G. (2008). Cognitive robotics. *Foundations of Artificial Intelligence*, *3*, 869–886.
- Levine, S., Finn, C., Darrell, T., & Abbeel, P. (2016). End-to-end training of deep visuomotor policies. *The Journal of Machine Learning Research*, *17*(1), 1334–1373.
- Limanowski, J., & Blankenburg, F. (2013). Minimal self-models and the free energy principle. *Frontiers in Human Neuroscience*, *7*, Article 547.
- Maisto, D., Friston, K., & Pezzulo, G. (2019). Caching mechanisms for habit formation in active inference. *Neurocomputing*, *359*, 298–314.
- Marr, D. (1982). *Vision: A computational investigation into the human representation and proceedings of visual information*. Freeman.
- Maturana, H. R., & Varela, F. J. (1991). *Autopoiesis and cognition: The realization of the living (Vol. 42)*. Springer Science & Business Media.
- Metta, G., Sandini, G., Vernon, D., Natale, L., & Nori, F. (2008). The iCub humanoid robot: An open platform for research in embodied cognition. In *Proceedings of the 8th workshop on performance metrics for intelligent systems* (pp. 50–56). ACM.
- Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural Networks*, *9*(8), 1265–1279.
- Millan, J. D. R., Leeb, R., Rupp, R., Muller-Putz, G. R., Neuper, C., Murray-Smith, R., Giugliemma, C., & . . . Kubler, A. (2010). Combining brain-computer interfaces and assistive technologies: State-of-the-art and challenges. *Frontiers in Neuroscience*, *4*, Article 161.
- Miyazaki, M., Hirashima, M., & Nozaki, D. (2010). The “Cutaneous Rabbit” hopping out of the body. *Journal of Neuroscience*, *30*(5), 1856–1860.
- Moxon, K. A., & Foffani, G. (2015). Brain-machine interfaces beyond neuroprosthetics. *Neuron*, *86*(1), 55–67.
- Murata, S., Yamashita, Y., Arie, H., Ogata, T., Sugano, S., & Tani, J. (2015). Learning to perceive the world as probabilistic or deterministic via interaction with others: A neuro-robotics experiment. *IEEE Transactions on Neural Networks and Learning Systems* *28*(4), 830–848.
- Neisser, U. (1988). Five kinds of self-knowledge. *Philosophical Psychology*, *1*, 35–59.
- Nilsson, N. J. (1984). *Shakey the robot*. Sri International.
- Nishimoto, R., & Tani, J. (2009). Development of hierarchical structures for actions and motor imagery: A constructivist view from synthetic neuro-robotics study. *Psychological Research*, *73*, 545–558.
- Nolfi, S., & Floreano, D. (2000). *Evolutionary robotics: The biology, intelligence, and technology of self-organizing machines*. MIT Press.
- Oudeyer, P. Y., Kaplan, F., & Hafner, V. V. (2007). Intrinsic motivation systems for autonomous mental development. *IEEE Transactions on Evolutionary Computation*, *11*(2), 265–286.
- Pastra, K., & Aloimonos, Y. (2012). The minimalist grammar of action. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *367*(1585), 103–117.
- Peniak, M., Marocco, D., Tani, J., Yamashita, Y., Fischer, K., & Cangelosi, A. (2011). Multiple time scales recurrent neural network for complex action acquisition. *Proceedings of ICDL-EPIROB*, *5*, 1–7.
- Piaget, J. (1953). *The origins of intelligence*. Routledge.
- Plato Cooper, J. M., & Hutchinson, D. S. (1997). *Plato: Complete works*. Hackett.
- Pollack, J. B. (1991). The induction of dynamical recognizers. *Machine Learning*, *7*, 227–252.
- Prescott, T. J., & Camilleri, D. (2019). The synthetic psychology of the self. In M. I. Aldinhas Ferreira, J. Silva Sequeira, & R. Ventura (Eds.), *Cognitive architectures* (pp. 85–104). Springer.
- Raab, D. H. (1963). Backward masking. *Psychological Bulletin*, *60*(2), 118–129.
- Rao, R., & Ballard, D. H. (1999). Predictive coding in the visual cortex: A functional interpretation of some extra-classical receptive-field effects. *Nature Neuroscience*, *2*(1), 79–87.
- Ricoeur, P. (1984). *Time and narrative (3 Vols)*. University of Chicago Press.

- Rumelhart, D. E., Hinton, G. E., & Williams, R. J., & University of California, San Diego. (1985). Learning internal representations by error propagation. In D. Rumelhart, & F. McClelland (Eds.), *Parallel distributed processing (Vol. 1)*, pp. 318–362. MIT Press.
- Schillaci, G., Ritter, C. N., Hafner, V. V., & Lara, B. (2016). Body representations for robot ego-noise modelling and prediction. Towards the development of a sense of agency in artificial agents. In *Proceedings of the Artificial Life Conference 2016* (pp. 390–397). MIT Press.
- Schmidhuber, J. (1991). *Adaptive confidence and adaptive curiosity*. Institut für Informatik, Technische Universität München.
- Shimojo, S. (2014). Postdiction: Its implications on visual awareness, hindsight, and sense of agency. *Frontiers in Psychology*, 5, Article 196.
- Silberstein, M., & Chemero, T. (2013). Constraints on localization and decomposition as explanatory strategies in the biological sciences. *Philosophy of Science*, 80(5), 958–970.
- Sugita, Y., & Tani, J. (2005). Learning semantic combinatoriality from the interaction between linguistic and behavioral processes. *Adaptive Behavior*, 13(1), 33–52.
- Sun, R. (2002). Duality of the mind: A bottom-up approach toward cognition. *London: Routledge*.
- Sun, R., Coward, L. A., & Zenzen, M. J. (2005). On levels of cognitive modeling. *Philosophical Psychology*, 18(5), 613–637.
- Tani, J. (1996). Model-based learning for mobile robot navigation from the dynamical systems perspective. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, 26(3), 421–436.
- Tani, J. (1998). An interpretation of the self from the dynamical systems perspective: A constructivist approach. *Journal of Consciousness Studies*, 5(5), 516–542.
- Tani, J. (2003). Learning to generate articulated behavior through the bottom-up and the top-down interaction processes. *Neural Networks*, 16(1), 11–23.
- Tani, J. (2009). Autonomy of “Self” at criticality: The perspective from synthetic neuro-robotics. *Adaptive Behavior*, 17(5), 421–443.
- Tani, J. (2016). *Exploring robotic minds. Actions, symbols, and consciousness as self-organizing dynamic phenomena*. Oxford University Press.
- Tani, J., Ito, M., & Sugita, Y. (2004). Self-organization of distributedly represented multiple behavior schemata in a mirror system: Reviews of robot experiments using RNNPB. *Neural Networks*, 17, 1273–1289.
- Tani, J., & Yamamoto, J. (2002). On the dynamics of robot exploration learning. *Cognitive Systems Research*, 3(3), 459–470.
- Tschantz, A., Seth, A., & Buckley, C. (2020). Learning action-oriented models through active inference. *PLoS Computational Biology*, 16(4), Article e1007805.
- Tsuda, I., Körner, E., & Shimizu, H. (1987). Memory dynamics in asynchronous neural networks. *Progress of Theoretical Physics*, 78, 51–71.
- Vygotsky, L. (1962). *Thought and language*. MIT Press. (Original work published 1934)
- Werbos, P. (1990). Backpropagation through time: What it does and how to do it. *Proceedings of the IEEE*, 78(10), 1550–1560.
- Yamada, T., Murata, S., Arie, H., & Ogata, T. (2016). Dynamical integration of language and behavior in a recurrent neural network for human–robot interaction. *Frontiers in Neurobotics*, 10, Article 5.
- Yamashita, Y., & Tani, J. (2008). Emergence of functional hierarchy in a multiple timescale neural network model: A humanoid robot experiment. *PLoS Computational Biology*, 4(11), Article e1000220.
- Yamashita, Y., & Tani, J. (2012). Spontaneous prediction error generation in schizophrenia. *PLOS ONE*, 7(5), Article e37843.

## About the Authors



**Jun Tani** received the D.Eng. degree from Sophia University, Tokyo in 1995. He started his research career with Sony Computer Science Lab in 1993. He became a Team Leader of the Laboratory for Behavior and Dynamic Cognition, RIKEN Brain Science Institute, Saitama, Japan, in 2001. He became a Full Professor with the Electrical Engineering Department, Korea Advanced Institute of Science and Technology, Daejeon, South Korea in 2012. He is currently a Full Professor with the Okinawa Institute of Science and Technology, Okinawa, Japan. His current research interests include cognitive neuroscience, developmental psychology, phenomenology, complex adaptive systems, and robotics. He is an author of “Exploring Robotic Minds: Actions, Symbols, and Consciousness as Self-Organizing Dynamic Phenomena.” published from Oxford Univ. Press in 2016.



**Jeffrey White** took his PhD (Philosophy) from the University of Missouri (Columbia) in 2006 having developed an information processing model of cognition as an extension of situated, embodied agency grounded in complex dynamic systems and integrating traditional moral psychology with hybrid computational models of human learning into the ACTWith, and later in terms of the inchworm and bivalve, models of moral cognition. Work with the Tani group focuses on understanding the development of biological cognition in terms of various computational architectures in order to advance research into autonomy, self, and the human condition.