Transferring active inference to newly encountered prism-shifted environments

Megumi Yoshihara[1,2*], Takuya Isomura[2*]

[1] Graduate School of Informatics, Kyoto University, 36-1 Yoshidahonmachi, Sakyo-ku, Kyoto 606-8501, Japan

[2] Brain Intelligence Theory Unit, RIKEN Center for Brain Science, 2-1 Hirosawa, Wako, Saitama 351-0198, Japan

* Corresponding authors emails: megumi.yoshihara@riken.jp, takuya.isomura@riken.jp

Abstract

Active inference has been proposed as a unified explanation of perception and action. Previous work has shown that canonical neural networks that minimize shared Helmholtz energy are cast as performing active inference of the external environment. However, how animals flexibly adapt to newly encountered environments remains to be fully addressed. To investigate the brain's generalizability and adaptability to new environments, this work develops canonical neural networks that employ multiple policy matrices in parallel. We demonstrate that the proposed model can recapitulate the prism adaptation—a form of visuomotor adaptation—under an arm-reaching task. Using policy matrices pretrained under various target positions, these networks could transfer previous experiences and exhibit faster adaptation to the prism-shifted environment than the naive networks. Furthermore, after-effects were observed following the removal of simulated prism glasses. These results suggest the biological plausibility and utility of the proposed model, providing insights into the adaptive capabilities of the brain.

Keywords: Free-energy principle, Prism adaptation, Active inference, POMDP, Transfer learning

INTRODUCTION

The brain exhibits great adaptability to newly encountered environments, and modeling it using neural networks remains a significant challenge in neuroscience. The free-energy principle (FEP) (Friston et al., 2006; Friston, 2010) has been proposed to account for perception (Lanillos et al.,

2021; Mirza et al., 2021), learning, and action(Klar et al., 2025; Lanillos et al., 2021) of biological organisms in terms of Bayesian inference and the minimization of variational free energy. Biological organisms are considered to develop the generative model—a hypothesis about the dependence structure between hidden states and sensory inputs—in the brain by minimizing variational free energy as a tractable proxy for minimizing the sensory surprise. By doing so, they perceive hidden environmental states and optimize their actions accordingly to minimize the risk associated with future outcomes, a process referred to as active inference (Da Costa et al., 2020; Friston et al., 2017).

Notably, recent works have shown that canonical neural networks that minimize Helmholtz energy can be read as following the FEP under a class of partially observable Markov decision process (POMDP) models (Isomura et al., 2022). These networks can exhibit active inference under various task settings, including causal inference (Isomura & Friston, 2020) rule learning (Tazawa & Isomura, 2024), prediction, and planning (Paul et al., 2024) These properties suggest that recapitulating the external environment in the network's internal states is an inherent property of neural networks.

Despite these successes, when encountering novel environments, these networks basically need to learn a generative model for new environments from scratch. This requires a considerable training cost, and the generalization error usually declines with a widely observed power-low order of time (i.e., $\mathcal{O}(t^{-1})$) (Hastie et al., 2009). By contrast, recent advances in machine learning suggest that transferring previously learned models may facilitate these networks to adaptation to new

environments (Kouw & Loog, 2018; Zhuang et al., 2020). However, whether canonical neural networks can exhibit such transfer learning remains largely unexplored.

To investigate the neuronal mechanisms of adaptability and generalizability, prism adaptation serves as a representative visuomotor phenomenon (Hermann & Cattell, 1898; Prablanc et al., 2020). This can be observed, for instance, when subjects perform a reaching task while wearing the prism glasses that shift their visual field. Initially, subjects wearing prism glasses struggle to complete reaching tasks as prediction misalignment causes an increase in reaching error; after a few trials, however, they gradually learn to accurately point to targets with minimal errors, demonstrating the adaptability to new environments. Upon removal of the prism glasses, initial errors are made but are immediately recovered. This task requires the readjustment and integration of visual and somatosensory information to novel sensory-motor contingencies, making it suited to model the brain's generalizability using canonical neural networks.

In doing so, some observations can be considered as a generalizability criterion: In healthy human prism exposure, the adaptation occurs within approximately 10 trials, and the reaching error approaches approximately zero (Luauté et al., 2009; Rossetti et al., 1993). This is a remarkably low number compared to the number of times a person has ever reached, indicating that biological brains can immediately correct their predictive discrepancies. However, previous modeling works have yet to fully explain the neuronal mechanisms underlying this significant gap (Sakaguchi et al., 2001; Smith et al., 2006; Inoue et al., 2015). Although Bayesian frameworks have been proposed as a promising avenue for addressing this issue (Petitet et al., 2018), concrete models that capture the underlying processes are still lacking.

To address this issue, this work developed canonical neural networks that exhibit transfer learning based on the FEP and active inference and applied them to explain the prism adaptation. We demonstrated that these neural networks can transfer previously learned policy matrices—pretrained to generate actions to reach various target positions depending on observations—to facilitate adaptation to a newly encountered environment. These networks exhibited the rapid adaptation in the presence of simulated prism glasses and the after-effects following the removal of the prism glasses, as observed in empirical prism adaptation. Mathematical analyses demonstrated that the prediction error for selecting the optimal mixture of policy matrices converges with an exponential order decay (i.e., $\mathcal{O}(e^{-\alpha t})$ with a coefficient $\alpha > 0$), which is considerably faster than the widely known t^{-1} order decay of generalization errors for conventional models. We conclude by discussing possible biological implementations of these learning mechanisms.

METHODS

Canonical neural networks

We began by outlining canonical neural networks (Isomura et al., 2022; Isomura & Friston, 2020) with a particular focus on modeling arm reaching tasks. We defined a two-layer canonical neural network (Fig. 1) of rate-coding models comprising middle $x(t) \coloneqq \left(x_1(t), \dots, x_{N_x}(t)\right)^{\mathrm{T}}$ and output $y(t) \coloneqq \left(y_1(t), \dots, y_{N_y}(t)\right)^{\mathrm{T}}$ layers. On receiving sensory inputs $o(t) \coloneqq \left(o_1(t), \dots, o_{N_o}(t)\right)^{\mathrm{T}}$, $x(t)$ yields the neural dynamics and propagates it to $y(t)$; then, $y(t)$ generates actions of the agent,

that is, a motion in either up, down, left, or right direction in a grid world. These neural activities are provided as follows:

$$\dot{x}(t) \propto -\mathrm{sig}^{-1}\big(x(t)\big) + (W_1 - W_0)o(t) + (K_1 - K_0)x(t - \Delta t) + h_1 - h_0 \tag{1}$$

$$\dot{y}(t) \propto -\mathrm{sig}^{-1}\big(y(t)\big) + (V_1 - V_0)(x(t - \Delta t) \otimes \lambda) + m_1 - m_0 \tag{2}$$

where $\lambda \coloneqq \left(\lambda_1, \dots, \lambda_{N_\lambda}\right)^{\mathrm{T}}$ denotes firing intensities of modulator neurons encoding mixture parameters; $W_1, W_0 \in \mathbb{R}^{N_x \times N_o}$ and $K_1, K_0 \in \mathbb{R}^{N_x \times N_x}$, $V_1, V_0 \in \mathbb{R}^{N_x \times N_x N_\lambda}$ are synaptic weight matrices; and $h_1 \coloneqq h_1(W_1, K_1), h_0 \coloneqq h_0(W_0, K_0), m_1 = m_1(V_1), m_0 = m_0(V_0)$ are the adaptive firing thresholds that depend on synaptic strengths. Here, $(V_1 - V_0)(x(t - \Delta t) \otimes \lambda)$ represents synaptic inputs from the middle layer that are modulated by λ.

Without loss of generality, Equations (1) and (2) can be derived as a gradient descent on Helmholtz energy $\mathcal{A}$. The functional form of $\mathcal{A}$ can be identified by computing the integral of the right-hand side of Equations (1) and (2), as follows:

$$\begin{aligned}\mathcal{A} = &\int_0^t \begin{pmatrix} x(\tau) \\ \bar{x}(\tau) \end{pmatrix}^{\mathrm{T}} \left\{ \ln \begin{pmatrix} x(\tau) \\ \bar{x}(\tau) \end{pmatrix} - \begin{pmatrix} W_1 \\ W_0 \end{pmatrix} o(\tau) - \begin{pmatrix} K_1 \\ K_0 \end{pmatrix} x(\tau - \Delta t) - \begin{pmatrix} h_1 \\ h_0 \end{pmatrix} \right\} d\tau \\ &+ \int_0^t \begin{pmatrix} y(\tau) \\ \bar{y}(\tau) \end{pmatrix}^{\mathrm{T}} \left\{ \ln \begin{pmatrix} y(\tau) \\ \bar{y}(\tau) \end{pmatrix} - \big(1 - 2\Gamma(\tau)\big) \begin{pmatrix} V_1 \\ V_0 \end{pmatrix} (x(\tau - \Delta t) \otimes \lambda) - \begin{pmatrix} m_1 \\ m_0 \end{pmatrix} \right\} d\tau + \mathcal{C}\end{aligned} \tag{3}$$

The form of $\mathcal{A}$ is equivalent to the variational free energy under a class of POMDPs (Isomura et al., 2022). This indicates that the dynamics of canonical neural networks are equivalent to Bayesian Belief updating under a class of POMDPs, supporting the validity of the free energy principle as a universal characterization of neural dynamics and self-organization. Notably, a gradient descent on the shared Helmholtz energy derives the synaptic plasticity in the output-layer synaptic weights *V*, which is expressed as neuromodulation of Hebbian plasticity in the form of a three-factor learning rule (Frémaux & Gerstner, 2016; Kuśmierz et al., 2017; Pawlak et al., 2010) (Fig. 1A). This formally

corresponds to the update of policy matrix to minimize future risks (Isomura et al., 2022). Based on this foundation, hereafter we formulated prism adaptation in terms of POMDPs (that correspond to canonical neural networks) following the FEP.

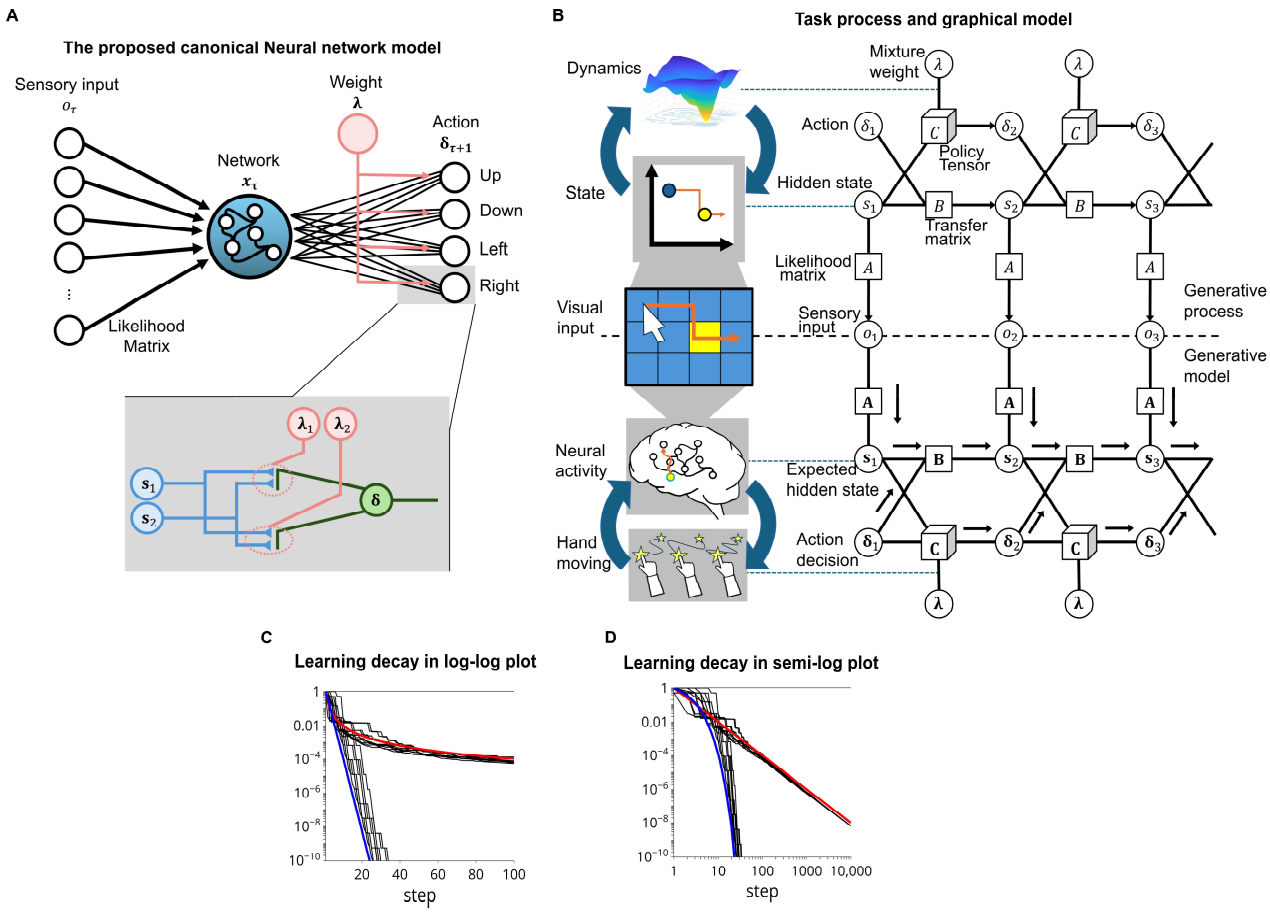

Fig. 1. Schematic of canonical neural networks performing active inference. A. Canonical neural network architecture for performing an arm reaching task (top) and possible neuronal implementation of policy learning in terms of the three-factor learning rule (bottom). Middle layer activity x_τ is generated by sensory inputs, o_τ^t and o_τ^h that carry information about target and hand positions. Action $\boldsymbol{\delta}_{\boldsymbol{\tau+1}}$ is generated by x_τ using the tensor of policy matrices, yielding a hand motion in the grid world. **B.** A POMDP model for reaching tasks expressed in the form of a

graphical model. The lower part formally corresponds to the canonical neural network shown in (A). This architecture enables transfer learning by selecting *C* from the pool via attentional switch $\boldsymbol{\lambda}$. Network weights and $\boldsymbol{\lambda}$ are updated by using $\mathbf{s}$, $\boldsymbol{\delta}$, and risk Γ. **C.** Generalization errors in **C** and $\boldsymbol{\lambda}$. **D**. Engaged view of (C) with semi-log scale.

Corresponding generative models

The behavior of an agent interacting with a discrete state space is expressed in the form of a POMDP as follows: The environment comprises binary sensory inputs (observations) $o_\tau = \left(o_\tau^1, \dots, o_\tau^{N_o}\right)^{\mathrm{T}} \in \{0,,1\}^{N_o}$ and hidden states $s_\tau = \left(s_\tau^1, \dots, s_\tau^{N_s}\right)^{\mathrm{T}} \in \{0,,1\}^{N_o}$, and their sequences are denoted as $o_{1:t} = \{o_1, \dots, o_t\}$ and $s_{1:t} = \{s_1, \dots, s_t\}$. Here, hidden states involve a factorial structure comprising two one-hot vectors of hand (s_τ^h) and target (s_τ^t) positions, which is denoted as $s_\tau = s_\tau^h \otimes s_\tau^t$ using the Kronecker product operator $\otimes$. Similarly, sensory inputs $o_\tau = o_\tau^h \otimes o_\tau^t$ comprise a product of somatosensory input for hand position o_τ^h and visual input for target position o_τ^t. Action or decision δ_τ describes the motion of the agent's hand, which can move up, down, left, or right at each time step. This makes the agent's hand position (s^h) move one grid, but it will stay if it reaches the end of the grid field.

Sensory inputs o_τ are generated based on hidden states s_τ in the form of a categorical distribution $P(o_\tau|s_\tau, A) = \mathrm{Cat}(As_\tau)$ through the likelihood matrix A. The dynamics of hidden states depend on the previous state and action, expressed as $P(s_{\tau+1}|s_\tau, \delta_\tau, B) = \mathrm{Cat}\big(B(\delta_\tau \otimes s_\tau)\big)$ using transition matrix B. The agent's action is determined based on the policy matrix C that defines the selection probability of action depending on hidden states, $P(\delta_{\tau+1}|s_\tau, C) = \mathrm{Cat}(Cs_\tau)$.

Here, a counterfactual learning of the policy matrix is adopted following the previous work (Isomura et al., 2022). Under this setting, the agent adopts a risk-dependent policy model $P(\delta_{\tau+1}|s_\tau, \gamma_{\tau+1}, C) = \mathrm{Cat}(Cs_\tau)^{\gamma_{\tau+1}}\mathrm{Cat}\big((C' \oslash C)s_\tau\big)^{\overline{\gamma_{\tau+1}}}$ to update C, in which the agent learns to retain or enhance the current strategy if the risk $\gamma_{\tau+1}$ is 0, whereas the agent forgets it if the risk is 1 (where $\overline{\gamma_{\tau+1}} = 1 - \gamma_{\tau+1}$ denotes the sign-flipped $\gamma_{\tau+1}$). Hence, the generative model is defined as follows:

$$P(\delta_{1:t}, s_{1:t}, o_{1:t}, \gamma_{1:t}, \theta) = P(\theta)P(\gamma_t)\prod_{\tau=1}^{t} P(o_\tau|s_\tau, A)P(s_\tau|s_{\tau-1}, \delta_{\tau-1}, B)P(\delta_\tau|s_{\tau-1}, \gamma_\tau, C) \qquad (4)$$

where θ is a set of parameters, $\theta = \{A, B, C\}$. Prior beliefs about parameter matrices $P(\theta) = P(A)P(B)P(C)$ follow Dirichlet distributions $P(A) = \mathrm{Dir}(a)$, $Q(B) = \mathrm{Dir}(b)$, and $Q(C) = \mathrm{Dir}(c)$ parameterized by Dirichlet parameters a, b, and c.

Implementation of transfer learning

In contrast to naive active inference agents defined above (i.e., naive learners), here we define agent exhibiting transfer learning (i.e., transfer learners). Transfer learning (Zhuang et al., 2020) enables the deployment of previously learned policy matrices to quickly adapt to newly encountered environments, including the presence of prism exposure. In this work, we model policy matrix C using a mixture of policy matrices as a function of mixture parameter λ, as follows:

$$\ln C(\lambda) = \lambda_1 \ln C_1 + \lambda_2 \ln C_2 + \cdots + \lambda_{N_C} \ln C_{N_C} \qquad (5)$$

Each fixed independent policy matrix C_k defines the selection probability of hand action in the k-th context, and λ is a mixture weight parameter or attentional switch that follows a categorical

distribution $\lambda \sim \mathrm{Cat}(\Lambda)$, analogous to a mixture of generative models(Isomura et al., 2019). A set of parameter matrices $\{C_k\}$ are generated by pretraining using naive learners. In this formulation, transfers learners employ the parameter set $\theta = \{A, B, \lambda\}$ and update them instead of updating $\theta_{\mathrm{naive}} = \{A, B, C\}$.

Variational Bayesian inference

An approximate posterior distribution is defined using a mean-field approximation as

$$Q(\delta_{1:t}, s_{1:t}, \theta_{\mathrm{naive}}) = Q(A)Q(B)Q(C)\prod_{\tau=1}^{t} Q(s_\tau)Q(\delta_\tau) \tag{6}$$

for naive learners and

$$Q(\delta_{1:t}, s_{1:t}, \theta) = Q(A)Q(B)Q(\lambda)\prod_{\tau=1}^{t} Q(s_\tau)Q(\delta_\tau) \tag{7}$$

for transfer learners. Approximate posterior beliefs about hidden states, action, and mixture parameter follow categorical distributions $Q(s_\tau) = \mathrm{Cat}(\mathbf{s}_\tau)$, $Q(\delta_\tau) = \mathrm{Cat}(\boldsymbol{\delta}_\tau)$ and $Q(\lambda) = \mathrm{Cat}(\boldsymbol{\lambda})$ parameterized by the posterior expectations $\mathbf{s}_\tau$, $\boldsymbol{\delta}_\tau$, and $\boldsymbol{\lambda}$. Moreover, approximate posterior beliefs about parameter matrices follow Dirichlet distributions $Q(A) = \mathrm{Dir}(\mathbf{a})$, $Q(B) = \mathrm{Dir}(\mathbf{b})$ and $Q(C) = \mathrm{Dir}(\mathbf{c})$ parameterized by Dirichlet parameters $\mathbf{a}$, $\mathbf{b}$, and $\mathbf{c}$.

Given the above generative model and posterior distribution, the variational free energy for naive learners is provided as follows:

$$F(o_{1:t}, \mathbf{s}_{1:t}, \boldsymbol{\delta}_{1:t}, \boldsymbol{\theta}) = \mathrm{E}_{Q,P(\gamma)}[-\ln P\,(\delta_{1:t}, s_{1:t}, o_{1:t}, \gamma_{1:t}, \theta) + \ln Q\,(\delta_{1:t}, s_{1:t}, \theta)]$$

$$= \sum_{\tau=1}^{t} \mathbf{s}_\tau \cdot (\ln \mathbf{s}_\tau - \ln \mathbf{A} \cdot o_\tau - \ln \mathbf{B}(\boldsymbol{\delta}_{\tau-1} \otimes \mathbf{s}_{\tau-1}))$$

$$+\sum_{\tau=1}^{t}\boldsymbol{\delta}_\tau\cdot(\ln\boldsymbol{\delta}_\tau-(1-2\Gamma_\tau)\ln\mathbf{C}\,\mathbf{s}_{\tau-1})$$

$$+(\mathbf{a}-a)\cdot\ln\mathbf{A}-\ln\boldsymbol{\mathcal{B}}(\mathbf{a})+(\mathbf{b}-b)\cdot\ln\mathbf{B}-\ln\boldsymbol{\mathcal{B}}(\mathbf{b})+(\mathbf{c}-\mathrm{c})\cdot\ln\mathbf{C}-\ln\boldsymbol{\mathcal{B}}(\mathbf{c}) \tag{8}$$

For transfer learners, the complexity of C in Equation (8), $(\mathbf{c}-\mathrm{c})\cdot\ln\mathbf{C}-\ln\boldsymbol{\mathcal{B}}(\mathbf{c})$, is replaced with that of λ, which is given as $\mathcal{D}_{\mathrm{KL}}[Q(\lambda)||P(\lambda)]=\boldsymbol{\lambda}\cdot(\ln\boldsymbol{\lambda}-\ln\Lambda)$. Here, $\Gamma_t\in[0,1]$ is a risk that evaluates the goodness of action for each step.

The gradient descent on F furnishes the Bayesian belief update rules. Posterior beliefs about hidden states s_τ and action δ_τ are updated for each step. These inference update rules are derived by solving the fixed point of implicit gradient descent (i.e., $\partial F/\partial\mathbf{s}_t=0$ and $\partial F/\partial\boldsymbol{\delta}_t=0$):

$$\begin{cases}\mathbf{s}_t=\sigma\big(\ln\mathbf{A}\cdot o_t+\ln\mathbf{B}\,(\boldsymbol{\delta}_{t-1}\otimes\mathbf{s}_{t-1})\big)\\ \boldsymbol{\delta}_t=\sigma(\ln\mathbf{C}\,\mathbf{s}_{t-1})\end{cases} \tag{9}$$

Dirichlet parameters a,b,c, and mixture parameter λ are updated at the end of each trial as part of the learning. These learning update rules are derived by solving $\partial F/\partial\theta=0$ as:

$$\mathbf{a}=a+\sum_{\tau=1}^{t}o_\tau\mathbf{s}_\tau^{\mathrm{T}}$$

$$\mathbf{b}=b+\sum_{\tau=1}^{t}\mathbf{s}_\tau(\boldsymbol{\delta}_{\tau-1}\otimes\mathbf{s}_{\tau-1})^{\mathrm{T}}$$

$$\mathbf{c}=c+\sum_{\tau=1}^{t}(1-2\Gamma_\tau)\boldsymbol{\delta}_\tau\mathbf{s}_{\tau-1}^{\mathrm{T}} \tag{10}$$

for naive learners. For transfer learners, the C's update is replaced with the λ's update as:

$$\boldsymbol{\lambda}_k=\sigma\left(\ln\Lambda_k+\sum_{\tau}(1-2\Gamma_\tau)\delta_\tau\cdot\ln C_k s_\tau\right) \tag{11}$$

Through Equation (11), the agent learns to select the action that minimizes the risk, while avoiding actions that increase the risk.

Simulation settings

In the simulation, a set of 100 policy matrices, $C_{(1,1)}, C_{(1,2)}, \dots, C_{(10,10)} \in [0,1]^{4\times10^4}$, are prepared for a mutually different target position, where each $C_{(n,m)}$ denotes the one pretrained to reach the target on the coordinate (n, m). During each pretraining, the target position was fixed at (n, m), and the agent continued to randomly move its hand for 100000 steps, which was followed by the learning of the posterior belief about $C_{(n,m)}$ based on Equation (10). In pre-training, we use an A matrix created based on a uniform distribution as the initial value so that the same policy can be adopted even if the target moves to a nearby location.

RESULTS

Transferring active inference to novel environments

This work aims to mathematically model the generalizability and adaptability of the brain to novel environments. An arm reaching task is a widely used visuo-motor paradigm to investigate the emergence of sensory-motor contingencies, in which subjects move their hands towards the target position based on visual input (Bourgeois et al., 2021). This can be modeled using canonical neural networks (Isomura et al., 2022), which receive visual and somatosensory inputs, infer the latent environmental states via middle-layer neural activity, and generate actions to shift the arm position (Fig. 1A, top). Risk here is characterized by the distance from the target, and the network learns the optimal action policy through the plasticity of the output layer synaptic connection weights (Fig. 1A, bottom). As previously established (Isomura et al., 2022), canonical neural

networks are read as performing active inference of the external environment to minimize risks associated with future outcomes, under a POMDP generative model (Fig. 1B). Thus, after sufficient exposure to the environment, these networks self-organize to recapitulate the interacting environment as a universal property.

In doing so, conventional models need to relearn the policy matrix (*C*) from scratch upon encountering a new environment, making it extremely time-consuming. Since the policy matrix posterior takes continuous values following a Dirichlet distribution, this method exhibits a widely observed 1/t-order generalization error decay (Ref: Hastie et al., 2009) (Fig. 1C, D, red lines). However, this indicates a large gap from humans and animals that can quickly adapt to moving their arms to a given target position (Luauté et al., 2009; Rossetti et al., 1993).

To model the quick adaptation, we constructed canonical neural networks that employ multiple policy matrices in parallel, each of which was pretrained to reach different target positions. The key advantage is the incorporation of the switching parameter λ that follows a categorical distribution. In contrast to the naive model, this network converges the generalization error considerably faster in a simple example setting (Fig. 1C, D blue lines).

Mathematical analyses revealed the distinct learning mechanisms (Appendix A). Critically, discrete switching variable λ tends to converge to a one-hot vector as evidence is accumulated. This makes the optimal policy selection from a set of policies converge with an exponential order, a speed considerably faster than the well-known 1/t order generalization error decay. Intuitively, this is because upon receiving sensory inputs, the evidence is accumulated linearly with time (i.e., $\mathcal{O}(t)$)—and the amplitude of the selected value is determined by the difference in evidence for

each component of λ. This makes the posterior expectation in the form of an exponential function $\exp(-t\langle \text{evidence gap} \rangle)$, where $\langle \text{evidence gap} \rangle$ denotes the difference in normalized evidence. Although computing this difference involves the generalization error in the $1/\sqrt{t}$ order, this is negligible compared to the order 1 amplitude of the difference itself. Therefore, the generalization error of λ is determined by the form $\exp\left(-t(X_{\text{1st}} - X_{\text{2nd}})\right)$ as the leading order, where X_{1st} and X_{2nd} are the normalized evidences for the most plausible and second most plausible policies, causing it to decrease exponentially with time. Further details are provided in Appendix A.

In essence, both mathematical analyses and numerical simulations (Fig. 1C, D) suggest that the selection mechanism of policy matrices enables more rapid adaptation than naive parameter learning, analogous to empirically observed prism adaptation. Based on these observations, in the remainder of this paper, we formally model prism adaptation using the canonical neural networks exhibiting transfer learning—and examine whether the rapid adaptation to a prism-shifted environment and the after-effects following the removal of prism glasses can be reproduced.

Modeling prism adaptation

In this section, we instantiated transfer of active inference through modeling a prism-shift paradigm of arm reaching tasks (Hermann & Cattell, 1898; Rossetti et al., 1993) (Fig. 2A). Agents perform a reaching task. For each trial, the agent's hand is initially located at a random position in a 10 × 10 grid field (Fig. 2B). At each step, the agents receive sensory information about target and hand position and infer hidden states and actions to determine the next motion from either up, down, left, or right (Equation (9)). A trial ends when either hand reaches the target or 100 steps

have passed, followed by the updates of parameters (Equations. (10) and (11)). This cycle continues for 200 trials. The risk Γ_t is defined based on the change in Manhattan distance between the hand and target: $\Gamma_t = 0.45$ when the distance decreases for each step (i.e., when the hand moved toward the target); otherwise, $\Gamma_t = 0.55$.

A simulation comprises three phases (Fig. 2C): The first 140 trials serves a "baseline" environment without prism glasses, in which the target is placed on the coordinate $(x, y) = (7,5)$ (Fig. 2C, Baseline). Starting from a flat prior $\boldsymbol{\lambda}$, the agent optimized the mixture parameter $\boldsymbol{\lambda}$ for the baseline condition (Fig. 2C). The second 30 trials correspond to the "prism exposure," in which the target is placed on (x, y) =(4,5) (a white filled grid in Fig. 2C, prism exposure phase) but its apparent position is still observed at (7,5). This prism shift was modeled by changing the likelihood mapping A, which displaces only the target position but not the hand position. While the apparent target position is unchanged from the baseline, the agent can notice the difference when it receives the risk after an action was made. In the last 30-trial "post exposure" phase, the agent returns to the environment in the absence of prism exposure, where the target is placed—and observed—at $(x, y) = (7,5)$. In each phase, the agent learns a policy that brings the hand closer to the true target by minimizing the free energy.

We demonstrated that transfer learners that employ pretrained policy matrices can replicate the key empirical properties of prism adaptation. Performance was assessed by the error between the x coordinates of the target and hand after the minimum steps to reach the target (Fig. 2C). While large errors were initially observed during both the prism exposure and post-exposure phases, they rapidly diminished to approximately zero. The transition in policy matrix $\mathbf{C}$ shows the

occurrence of adaptation in both phases (Fig. 2C, bottom), which was driven by the update of mixture weights $\boldsymbol{\lambda}$ (Fig. 2D). In each phase, transfer learners consistently learn the optimal $\boldsymbol{\lambda}$ values. Specifically, during prism exposure, $\boldsymbol{\lambda}$ is adjusted to align the policy with the actual target position, while in the post-exposure phase, $\boldsymbol{\lambda}$ reverts to the baseline policy.

We observed distinct properties between transfer and naive learners. The proposed model (transfer learner) qualitatively replicated experimental outcomes (Fig. 2E). During prism exposure, initially large duration to reach the target decreased after approximately 5 trials and converged around 10. In the post-exposure phase, transfer learners initially exhibited an after-effect akin to empirical prism adaptation and re-learn gradually. In contrast, naive learners that learn policy matrix $\mathbf{C}$ directly via Equation (10) could not reproduce the after-effect (Fig. 2E, red error bars). While the duration of naive learners increased during the prism exposure, they failed to exhibit the duration increase during post-exposure, arguably because of slow learning during the prism exposure.

In summary, the proposed model can learn to perform the arm reaching task both in the absence and presence of prism shift in the Bayes optimal manner and replicate the dynamics of empirical prism adaptation. We confirmed that these outcomes can be observed even with randomly varied target positions (Fig. 3).

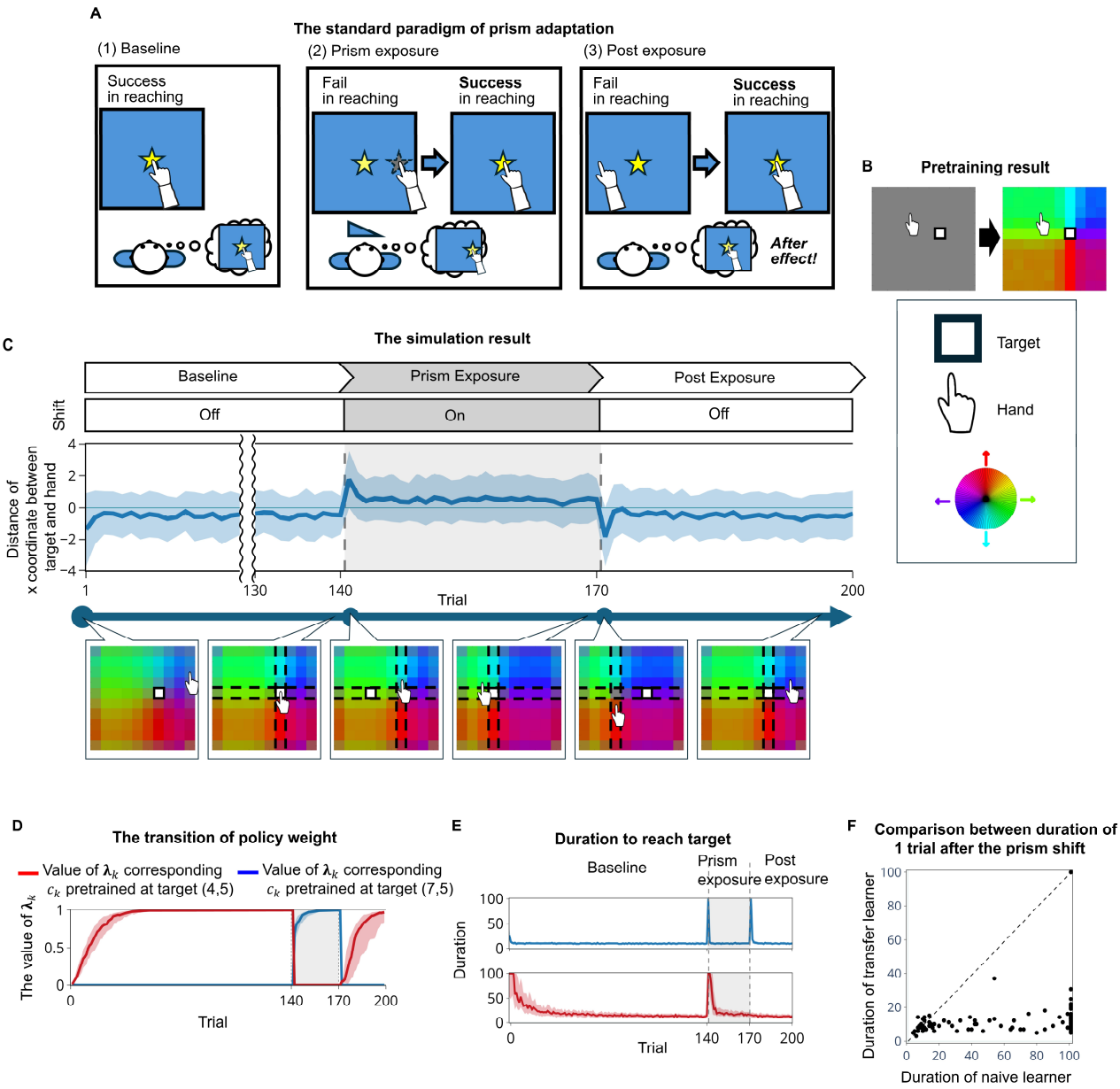

Fig. 2. Simulating prism adaptation using canonical neural networks. A. The prism adaptation paradigm: (1) a subject performs an arm reaching task in the baseline phase; (2) during the prism exposure phase, the same task is performed with wearing prism glasses; and (3) during the post-exposure phase, the prism glasses are removed. **B.** The pretraining of policy matrices. Prism adaptation is modeled with a discrete state space (POMDP) model. The hand and target are placed in a 10x10 grid world. The colors in each grid represent the hand-moving directions that correspond to the color map besides the heatmap, indicating the action selection probability for

moving up/down/left/right. **C.** Overview of the simulation setting. The upper panel shows the trajectory of the x-coordinate difference between the target and hand after the minimum steps. The bottom panel shows the corresponding changes in the policy matrix. Heatmaps illustrate how the agent's policy matrix changes over trials. While the colors are blurred before training (leftmost), the colors become high contrast through training. The fixed point (crosspoint of vertical and horizontal dashed lines) approximately pursues the target position (open square) both in the absence and presence of prism shift. **D**. Changes in the two policy mixture weights corresponding to original and new target positions. **E**. The upper and lower panels show the duration (the number of steps) for transfer learners and naive learners, respectively. **F**. Comparison of the duration at the 1 trial after prism exposure between the transfer and naive learners. In (C)–(E), the lines and shaded areas indicate the medians and interquartile ranges.

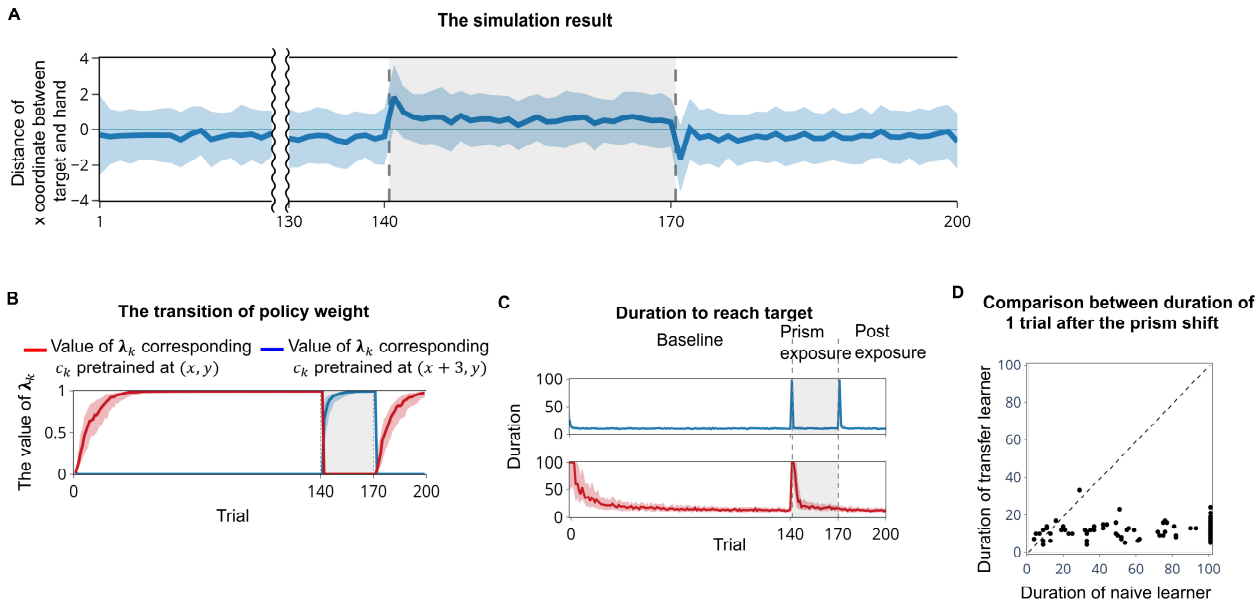

Fig. 3. Simulations with randomized target. Same as Fig.2 but the target position is changed randomly over sessions, showing the robustness of the proposed model. **A.** The x-coordinate

difference between the target and the hand after the minimum steps. **B**. Changes in the two policy mixture weights. **C**. Duration for the transfer (top) and naive learners (bottom). **D**. Comparison of the duration value between the transfer and naïve learners. In (A)–(C), the lines and shaded areas indicate the medians and interquartile ranges.

Exponential-order generalization error decay in prism adaptation

Lastly, we analytically investigated the learning speed of transfer and naive learners. For analytical tractability, here we assume that the posterior hidden state $\mathbf{s}_\tau$ well approximates the true hidden state s_τ, the action δ_τ follows a random walk, and that there exists an optimal policy matrix C^* in a set of pretrained matrices $\{C_k\}$. Under these conditions, the mixture parameter $\boldsymbol{\lambda}$ and policy matrix $\mathbf{C}(\boldsymbol{\lambda})$ converge to the Bayes optimal values $\boldsymbol{\lambda}^*$ and $\mathbf{C}^*$ with the e^{-at} order decay of generalization error, where $a > 0$. Detailed derivations are described in detail in the Appendix. This convergence speed is considerably distinct from learning speed in the policy matrix $\mathbf{C}$ of naive learners that converges to the Bayes optimal value $\mathbf{C}^*$ with the $1/t$ order decay.

To valid the change in generalization error, we simulated the transfer learners under the condition in which the action is defined as a random walk and confirmed the e^{-at} order convergence of policy mixture weights $\boldsymbol{\lambda}$ (Fig. 4A) and policy matrix $\mathbf{C}$ (Fig. 4B). The generalization error of the mixture weight $\boldsymbol{\lambda}$ in transfer learners converges in approximately 2400 steps, and its change resembles a straight line in a semi-log plot, indicating the exponential order convergence (Fig. 4A, blue line). After 2400 steps, more than half of $\boldsymbol{\lambda}$ elements become Bayes optimal one-hot vector; therefore, the median generalization error becomes zero. The gradient of the

generalization error in the semi-log plot closely matches the theoretical gradient calculated in the Appendix. The generalization error of the policy matrix $\mathbf{C}$ in the transfer learner also exhibited a similar property (Fig. 4B), in which empirical errors resembled a theoretical line (green line). In contrast, the naive learner requires a longer convergence time in the t^{-1} order (Fig. 4C), aligning a line in a log-log scale after about 500 steps.

In essence, the error of the transfer learner decreases much faster than that of the naive learner, enabling a faster optimization of the policy matrix. The faster adaptation is further confirmed under the condition where the hidden states are updated following Bayesian belief updating rules (Fig. 4D, E). At the one trial after the prism shift (i.e., trial 142), the majority of transfer learners reached the target significantly faster than the naïve learners (Fig. 2F; $p = 4.15 \times 10^{-18}$, Mann–Whitney U test) and the median duration of transfer learners was only approximately 8.3% of that of naive learners.

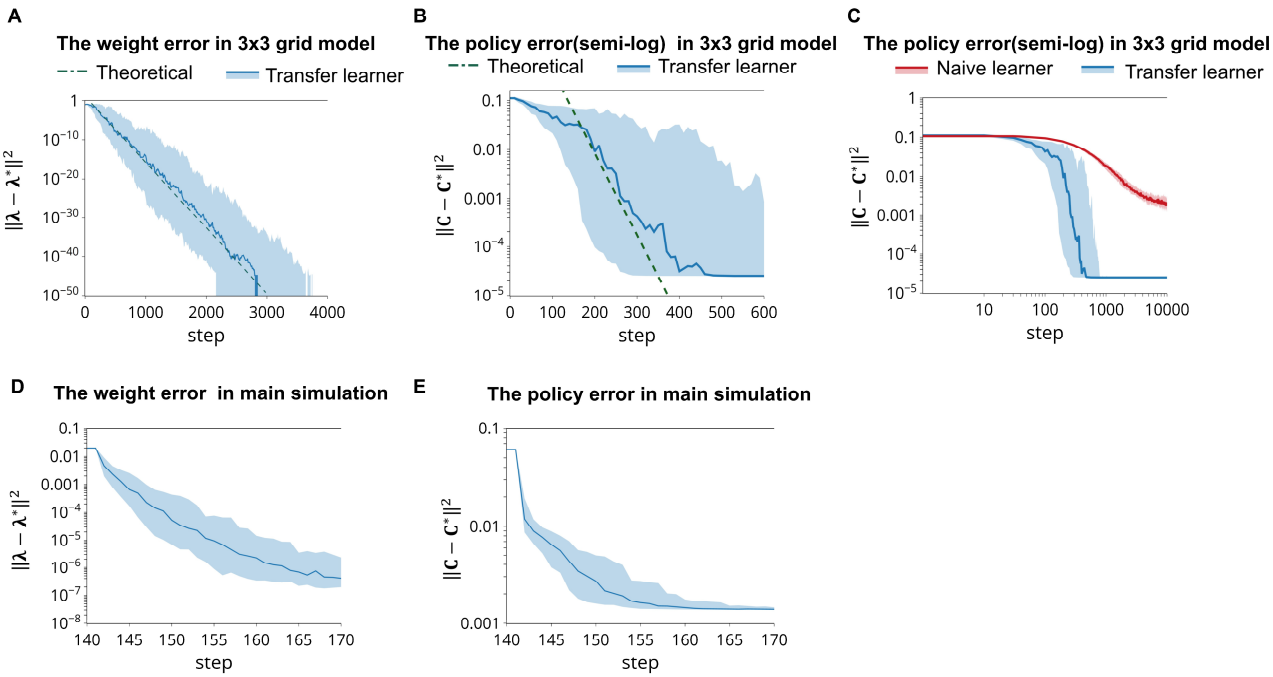

Fig. 4. Generalization error in simplified model simulation. A. Generalization error of $\boldsymbol{\lambda}$ and the theoretical value in toy model. **B**. Generalization error of $\mathbf{C}$ and theoretical value in semi-log plot in toy model. **C.** Generalization error of $\mathbf{C}$, and theoretical value in log-log plot. Both learners have a common initial value of the policy matrix in toy model. **D.** Generalization error of $\boldsymbol{\lambda}$ in main simulation. **E**. Generalization error of $\mathbf{C}$ and theoretical value in semi-log plot in main simulation.

DISCUSSION

In this work, we developed an active inference model incorporating transfer learning and applied it to model prism adaptation. Canonical neural networks employing multiple policy matrices can transfer previously learned policies to novel environments, demonstrating faster adaptation than naive learners. This has broader implications for understanding the brain's generalizability and adaptability, particularly in the context of prism adaptation. The transfer learners involve three different timescales comprising fast updates of state s and action δ, slow updates of mixture weights $\boldsymbol{\lambda}$, and ultra-slow updates of parameters $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{C}$, akin to previous computational models for prism adaptation (Inoue et al., 2015).

Two timescales in prism adaptation were suggested by several works (G. M. Redding et al., 2005; G. Redding & Wallace, 2002, 2006). On this view, fast and slow adaptations are interpreted as recalibration of frames and realignment of the visuo-motor mapping, respectively, and the latter induces after effect. Our observations were consistent with this view: only transfer learners that

update $\boldsymbol{\lambda}$, but not naive learners without $\boldsymbol{\lambda}$, exhibited after effect, indicating the importance of transfer mechanisms to reproduce empirical phenomena.

Such correspondences in timescales are beneficial to identify a precise mapping between the parameters of active inference models and corresponding brain regions. Specifically, faster timescale adaptation is attenuated by lesions in the posterior parietal cortex (Canavan et al., 1990; Clower et al., 1996; Newport & Jackson, 2006; Pisella et al., 2004; Welch & Goldstein, 1972). This impairs both error reduction and after effects. Moreover, fast adaptation involves the primary motor cortex (Danckert et al., 2008), anterior cingulate cortex (Danckert et al., 2008; Luauté et al., 2009), and cerebellum (Danckert et al., 2008; Luauté et al., 2009). In particular, cerebellar lesions result in decreased or absent prism after-effects (Calzolari et al., 2015; Hanajima et al., 2015; Pisella et al., 2005), suggesting its role in slow adaptation.

Furthermore, the underlying neuronal mechanisms can be considered in terms of neuromodulation of synaptic plasticity. The $\boldsymbol{\lambda}$ update rule in Equation (11) compares action and experience and reinforces optimal policy, providing a functional explanation of policy adaptation, which may be associated with attentional switch and working memory mediated by such as dopaminergic neurons. Then, the process of amplifying more similar connections and attenuating less similar connections can be formulated as dopaminergic modulation of synaptic plasticity.

Many traditional models in prism adaptation have posited multi-timescale mechanisms (Inoue et al., 2015; Kim et al., 2015; Kording et al., 2007). While these models can account for behavioral performance, they do not fully explain the underlying neuronal mechanisms and how different timescales contribute to learning. The proposed canonical neural networks have potential to

assimilate empirical data into modelling. Earlier work has established a reverse engineering method to estimate generative model from empirical neural activity data based on canonical neural networks (Isomura et al., 2023, 2025). This allows us to compare model predictions with empirical results (Kitazawa et al., 1995). In future work, we hope to elucidate the functional relationship between brain regions and the prism adaptation process by comparing the network properties of our model and empirical data.

CONCLUSION

In summary, we developed a transfer learning model of active inference based on canonical neural networks. Our model selected the Bayes optimal parameter from pretrained parameters and demonstrated faster adaptation to a novel prism glasses environment. Further, it successfully recapitulated features of prism adaptation such as after effect. These results support that the mixture policy matrices enable transfer of knowledge learned from experiences to quickly adapt to the environment and have the capacity to reproduce the brain's generalizability.

Appendix

Comparison of generalization errors

Here, we assume that the posterior beliefs about the hidden state and action well approximate the true hidden state and action: $\mathbf{s} = s, \boldsymbol{\delta} = \delta$. In addition, we assume that $\{\delta_\tau\}$ is a random walk, indicating that any $\delta_\tau, \delta_{\tau'}$ $(\tau \neq \tau')$ are independent and identically distributed.

We define a Hebbian product as a block vector, $\vec{X} = \frac{1}{t}\sum_{\tau=1}^{t}(1-2\Gamma_\tau)\boldsymbol{\delta}_\tau \otimes \mathbf{s}_{\tau-1}$ and its large-time limit as $\vec{X}^* = \lim_{t\to\infty}\vec{X} = \langle(1-2\Gamma_\tau)\boldsymbol{\delta}_\tau \otimes \mathbf{s}_{\tau-1}\rangle$. The generalization error of $\vec{X}$ is given by the covariance matrix of their difference $E_X = \langle\Delta\vec{X}\Delta\vec{X}^{\mathrm{T}}\rangle$, where $\Delta\vec{X} = \vec{X} - \vec{X}^*$. When the agent obeys a random work, E_X can be analytically computed, showing the t^{-1} order convergence.

The posterior belief about the policy mapping C is parameterized by the Dirichlet parameter, which is given as Equation (7), or equivalently, as a vector $\vec{\mathbf{c}} = \vec{c} + t\vec{X}$. Then, the posterior belief about C is given as $\mathbf{C} = \mathbf{c} \oslash (\vec{1}\vec{1}^{\mathrm{T}}\mathbf{c}) \Leftrightarrow \vec{\mathbf{C}} = \vec{X} \oslash \vec{Y}$ when c is negligibly small, where $\oslash$ indicates the Hadamard division operator and $\vec{Y} = \frac{1}{t}\sum_{\tau=1}^{t}(1-2\Gamma_\tau)\vec{1} \otimes \mathbf{s}_{\tau-1}$. When the agent obeys a random walk, $\vec{Y}^* = \langle(1-2\Gamma_\tau)\vec{1} \otimes \mathbf{s}_{\tau-1}\rangle = \frac{1}{N_s}(1-2\langle\Gamma_\tau\rangle)\vec{1} \otimes \vec{1}$ holds. This provides $\vec{\mathbf{C}}^* = N_s(1-2\langle\Gamma_\tau\rangle)^{-1}\vec{X}^*$.

The generalization error of $\vec{\mathbf{C}}$ is given as $E_C = \left\langle(\vec{\mathbf{C}} - \vec{\mathbf{C}}^*)(\vec{\mathbf{C}} - \vec{\mathbf{C}}^*)^{\mathrm{T}}\right\rangle$. Then, $\vec{\mathbf{C}} = \vec{X} \oslash \vec{Y} = (\vec{X}^* + \Delta\vec{X}) \oslash (\vec{Y}^* + \Delta\vec{Y})$ is approximated as $\vec{\mathbf{C}} \simeq \vec{X}^* \oslash \vec{Y}^* + \Delta\vec{X} \oslash \vec{Y}^* - \vec{X}^* \oslash \vec{Y}^{*\odot 2} \odot \Delta\vec{Y} = \vec{\mathbf{C}}^* + N_s(1-2\langle\Gamma_\tau\rangle)^{-1}(\Delta\vec{X} - \vec{\mathbf{C}}^* \odot \Delta\vec{Y})$ as the first-order approximation. Thus, E_C becomes

$$E_C = N_s^2(1-2\langle\Gamma_\tau\rangle)^{-2}\left\langle(\Delta\vec{X} - \vec{\boldsymbol{C}}^* \odot \Delta\vec{Y})(\Delta\vec{X} - \vec{\boldsymbol{C}}^* \odot \Delta\vec{Y})^{\mathrm{T}}\right\rangle \tag{12}$$

Because $\vec{Y}$ is expressed as $\vec{Y} = \Omega\vec{X}$ using a block diagonal matrix Ω, $\vec{\mathbf{C}}^* \odot \Delta\vec{Y} = \mathrm{diag}[\vec{\mathbf{C}}^*](\vec{Y} - \vec{Y}^*) = \mathrm{diag}[\vec{\mathbf{C}}^*](\Omega\vec{X} - \Omega\vec{X}^*) = \mathrm{diag}[\vec{\mathbf{C}}^*]\Omega\Delta\vec{X}$ holds. Thus, we obtain

$$E_C = N_s^2(1-2\langle\Gamma_\tau\rangle)^{-2}\left\langle\left(I-\mathrm{diag}[\vec{\mathbf{C}}^*]\Omega\right)\Delta\vec{X}\Delta\vec{X}^{\mathrm{T}}\left(I-\mathrm{diag}[\vec{\mathbf{C}}^*]\Omega\right)^{\mathrm{T}}\right\rangle$$
$$= N_s^2(1-2\langle\Gamma_\tau\rangle)^{-2}\left(I-\mathrm{diag}[\vec{\mathbf{C}}^*]\Omega\right)E_X\left(I-\mathrm{diag}[\vec{\mathbf{C}}^*]\Omega\right)^{\mathrm{T}} \quad (13)$$

In contrast, the transfer learner that equips with the mixture policies updates the mixing balance $\boldsymbol{\lambda}$ instead of updating $\mathbf{c}$ directly. In this case, the posterior belief about the policy mapping is given as follows:

$$\vec{\mathbf{C}} = \exp\left[\overrightarrow{\ln C_1}\boldsymbol{\lambda}_1 + \cdots + \overrightarrow{\ln C_{N_\lambda}}\boldsymbol{\lambda}_{N_\lambda}\right]$$
$$= \vec{\mathbf{C}}^* + \vec{\mathbf{C}}^* \odot \left(\overrightarrow{\ln C_1}\,\Delta\boldsymbol{\lambda}_1 + \cdots + \overrightarrow{\ln C_{N_\lambda}}\,\Delta\boldsymbol{\lambda}_{N_\lambda}\right) \quad (14)$$

Here, the λ update rule is expressed as $\boldsymbol{\lambda}_k = \sigma\left(\ln\boldsymbol{\Lambda}_{\mathrm{k}} + \overrightarrow{\ln C_k}\cdot t\vec{X}\right)$ and its large-time limit as $\boldsymbol{\lambda}_k^* = \sigma\left(\ln\boldsymbol{\Lambda}_{\mathrm{k}} + \overrightarrow{\ln C_k}\cdot t\vec{X}^*\right)$. The difference between $\vec{\mathbf{C}}$ and $\vec{\mathbf{C}}^*$ is computed as

$$\vec{\mathbf{C}} - \vec{\mathbf{C}}^* = \vec{\mathbf{C}}^* \odot \left(\overrightarrow{\ln C_1}, \ldots, \overrightarrow{\ln C_{N_\lambda}}\right)\Delta\boldsymbol{\lambda} \quad (15)$$

where $\Delta\boldsymbol{\lambda} = \boldsymbol{\lambda} - \boldsymbol{\lambda}^*$ holds and Λ is assumed to be a flat prior and thus omitted for simplicity.

When one of the previously experienced contexts has a considerably higher similarity to the current environment compared to other contexts, an element of $\boldsymbol{\lambda}$ that corresponds to the context monotonically converges to one. Under this condition, $\boldsymbol{\lambda}^* = \sigma\left(t\left(\overrightarrow{\ln C_1}, \ldots, \overrightarrow{\ln C_{N_\lambda}}\right)^{\mathrm{T}}\vec{X}^*\right) = \sigma\left(t\left(\overrightarrow{\ln C_1}, \ldots, \overrightarrow{\ln C_{N_\lambda}}\right)^{\mathrm{T}}\vec{X}^* - t\max_l \overrightarrow{\ln C_l}^{\mathrm{T}}\vec{X}^*\right)$ holds owing to the property of softmax function σ. Thus, by defining $l = \mathrm{argmax_i}(\overrightarrow{\ln C_l}^{\mathrm{T}}\vec{X}^*)$ and $D_i = \overrightarrow{\ln C_i}\vec{X}^*$, we obtain $\boldsymbol{\lambda}_k^* = \frac{e^{tD_k}}{\sum_i e^{tD_i}} \propto e^{t(D_k - D_l)}$. All elements of $\boldsymbol{\lambda}$ except for the largest element ($k = l$) converge to zero exponentially with time, i.e., $\boldsymbol{\lambda}_k = \exp\left(tD_k - tD_l + \mathcal{O}\left(\sqrt{t}\right)\right) = \mathcal{O}\left(e^{-(D_l - D_k)t}\right)$ as the leading order, because the generalization error of $\overrightarrow{\ln C_l}\vec{X}$ is of order $\mathcal{O}\left(1/\sqrt{t}\right)$ and negligibly smaller than $(D_l - D_k)$. Conversely, the largest element is computed as $\boldsymbol{\lambda}_k = 1 - \sum_{l\neq k}\boldsymbol{\lambda}_l = 1 - \mathcal{O}\left(e^{t(D_{l'} - D_l)}\right)$, where $l' = \mathrm{argmax}_{i\neq l}(\overrightarrow{\ln C_l}^T\vec{X}^*)$ indices the second largest element.

This provides $\boldsymbol{\lambda} - \boldsymbol{\lambda}^* = \mathcal{O}\left(e^{-t(D_l - D_{l'})}\right)$. Therefore, the generalization error $E_C = \left\langle (\vec{\mathbf{C}} - \vec{\mathbf{C}}^*)(\vec{\mathbf{C}} - \vec{\mathbf{C}}^*)^{\mathrm{T}} \right\rangle$ for the transfer learner is in $\mathcal{O}\left(e^{2t(D_l - D_{l'})}\right)$, which indicates that E_C of the transfer learner converges to zero much faster than does that of the naive learner. Thus, the generalization error of the transfer learner is determined based on the generalization error of pretrained policy matrices.

Table 1. Glossary of expressions

Expression	Definition	Explanation
t	Current time	$t \in \{0,1,\dots\}$
τ	Time index	$\tau \in \{0,1,\dots,t\}$
o_τ	Sensory input	$o_\tau \in \{0,1\}^{10000\times 1}$
s_τ	Hidden state	$s_\tau = s_\tau^h \otimes s_\tau^t$
s_τ^h	Hidden state of hand position	$s_\tau^h \in \{0,1\}^{100\times 1}$
s_τ^t	Hidden state of target position	$s_\tau^t \in \{0,1\}^{100\times 1}$
$\mathbf{s_\tau}$	Expected hidden state	Posterior belief of hidden state s_τ, $\mathbf{s_\tau} \in [0,1]^{10000\times 1}$
δ_τ	Action (up down left right)	$\delta_\tau \in \{0,1\}^{4\times 1}$
$\boldsymbol{\delta}_\tau$	Expected action	Posterior belief of action δ_τ, $\boldsymbol{\delta}_\tau \in [0,1]^{4\times 1}$

A, B, C	parameter matrices	$A \in [0,1]^{10000\times10000}, B \in [0,1]^{10000\times40000}$, $C \in [0,1]^{4\times10000}$
$\mathbf{A}, \mathbf{B}, \mathbf{C}$	Expected parameter matrices	Posterior belief of parameter matrices, A, B, C, $\mathbf{A} \in [0,1]^{10000\times10000}, \mathbf{B} \in [0,1]^{10000\times40000}$, $\mathbf{C} \in [0,1]^{4\times10000}$
λ	Mixture weight	Mixture weight for policy.
$\boldsymbol{\lambda}$	Expected mixture weight	Posterior belief of mixture parameter λ.
θ	parameters	The set of parameters.
$\boldsymbol{\theta}$	Expected parameters	Posterior belief of parameter θ.
a, b, c	Dirichlet parameter	Dirichlet parameter of matrices, A, B, C, $a \in \mathbb{R}_{>0}^{10000\times10000}, b \in \mathbb{R}_{>0}^{10000\times40000}$, $c \in \mathbb{R}_{>0}^{4\times10000}$
$\mathbf{a}, \mathbf{b}, \mathbf{c}$	Dirichlet parameter	Dirichlet parameter of matrices, $\mathbf{A}, \mathbf{B}, \mathbf{C}$, $\mathbf{a} \in \mathbb{R}_{>0}^{10000\times10000}, \mathbf{b} \in \mathbb{R}_{>0}^{10000\times40000}$, $\mathbf{c} \in \mathbb{R}_{>0}^{4\times10000}$
Γ	risk	$\Gamma \in [0,1]$
$\boldsymbol{\mathcal{B}}(\cdot)$	Beta function	$\boldsymbol{\mathcal{B}}(\mathbf{a}_{\cdot k}) = \frac{\prod_i \Gamma(\mathbf{a}_{ik})}{\Gamma(\sum_i \mathbf{a}_{ik})}$, $\Gamma(\cdot)$: Gamma function

Reference

Bourgeois, A., Schmid, A., Turri, F., Schnider, A., & Ptak, R. (2021). Visual but not auditory-verbal feedback induces aftereffects following adaptation to virtual prisms. *Frontiers in Neuroscience*, *Volume 15-2021*. https://doi.org/10.3389/fnins.2021.658353

Calzolari, E., Bolognini, N., Casati, C., Marzoli, S. B., & Vallar, G. (2015). Restoring abnormal aftereffects of prismatic adaptation through neuromodulation. *Effects of Non-Invasive Brain Stimulation on Attention: Current Debates, Cognitive Studies and Novel Clinical Applications*, *74*, 162–169. https://doi.org/10.1016/j.neuropsychologia.2015.04.022

Canavan, A. G. M., Passingham, R. E., Marsden, C. D., Quinn, N., Wyke, M., & Polkey, C. E. (1990). Prism adaptation and other tasks involving spatial abilities in patients with Parkinson's disease, patients with frontal lobe lesions and patients with unilateral temporal lobectomies. *Neuropsychologia*, *28*(9), 969–984. https://doi.org/10.1016/0028-3932(90)90112-2

Clower, D. M., Hoffman, J. M., Votaw, J. R., Faber, T. L., Woods, R. P., & Alexander, G. E. (1996). Role of posterior parietal cortex in the recalibration of visually guided reaching. *Nature*, *383*(6601), 618–621. https://doi.org/10.1038/383618a0

Da Costa, L., Parr, T., Sajid, N., Veselic, S., Neacsu, V., & Friston, K. (2020). Active inference on discrete state-spaces: A synthesis. *Journal of Mathematical Psychology*, *99*, 102447. https://doi.org/10.1016/j.jmp.2020.102447

Danckert, J., Ferber, S., & Goodale, M. A. (2008). Direct effects of prismatic lenses on visuomotor control: An event-related functional MRI study. *European Journal of Neuroscience*, *28*(8), 1696–1704. https://doi.org/10.1111/j.1460-9568.2008.06460.x

Frémaux, N., & Gerstner, W. (2016). Neuromodulated spike-timing-dependent plasticity, and theory of three-factor learning rules. *Frontiers in Neural Circuits*, *Volume 9-2015*. https://doi.org/10.3389/fncir.2015.00085

Friston, K. (2010). The free-energy principle: A unified brain theory? *Nature Reviews Neuroscience*, *11*(2), 127–138. https://doi.org/10.1038/nrn2787

Friston, K., FitzGerald, T., Rigoli, F., Schwartenbeck, P., & Pezzulo, G. (2017). Active Inference: A Process Theory. *Neural Computation*, *29*(1), 1–49. https://doi.org/10.1162/NECO_a_00912

Friston, K., Kilner, J., & Harrison, L. (2006). A free energy principle for the brain. *Theoretical and Computational Neuroscience: Understanding Brain Functions*, *100*(1), 70–87. https://doi.org/10.1016/j.jphysparis.2006.10.001

Hanajima, R., Shadmehr, R., Ohminami, S., Tsutsumi, R., Shirota, Y., Shimizu, T., Tanaka, N., Terao, Y., Tsuji, S., Ugawa, Y., Uchimura, M., Inoue, M., & Kitazawa, S. (2015). Modulation of error-sensitivity during a

prism adaptation task in people with cerebellar degeneration. *Journal of Neurophysiology*, *114*(4), 2460–2471. https://doi.org/10.1152/jn.00145.2015

Hastie, T., Tibshirani, R., & Friedman, J. H. (2009). *The elements of statistical learning: Data mining, inference, and prediction*. Springer. https://books.google.co.jp/books?id=eBSgoAEACAAJ

Hermann, V. Helmholtz., & Cattell, J. M. (1898). Handbuch der physiologischen Optik. *Science*, *8*(205), 794–796. https://doi.org/10.1126/science.8.205.794.b

Inoue, M., Uchimura, M., Karibe, A., O'Shea, J., Rossetti, Y., & Kitazawa, S. (2015). Three timescales in prism adaptation. *Journal of Neurophysiology*, *113*(1), 328–338. https://doi.org/10.1152/jn.00803.2013

Isomura, T., & Friston, K. (2020). Reverse-Engineering Neural Networks to Characterize Their Cost Functions. *Neural Computation*, *32*(11), 2085–2121. https://doi.org/10.1162/neco_a_01315

Isomura, T., Kotani, K., Jimbo, Y., & Friston, K. J. (2023). Experimental validation of the free-energy principle with in vitro neural networks. *Nature Communications*, *14*(1), 4547. https://doi.org/10.1038/s41467-023-40141-z

Isomura, T., Parr, T., & Friston, K. (2019). Bayesian Filtering with Multiple Internal Models: Toward a Theory of Social Intelligence. *Neural Computation*, *31*(12), 2390–2431. https://doi.org/10.1162/neco_a_01239

Isomura, T., Shimazaki, H., & Friston, K. J. (2022). Canonical neural networks perform active inference. *Communications Biology*, *5*(1), 55. https://doi.org/10.1038/s42003-021-02994-2

Isomura, T., Tanimoto, Y., Torigoe, M., Okamoto, H., & Shimazaki, H. (2025). Predicting individual learning trajectories in zebrafish via the free-energy principle. *bioRxiv : The Preprint Server for Biology*. https://doi.org/10.1101/2025.08.06.668947

Kim, S., Ogawa, K., Lv, J., Schweighofer, N., & Imamizu, H. (2015). Neural Substrates Related to Motor Memory with Multiple Timescales in Sensorimotor Adaptation. *PLOS Biology*, *13*(12), e1002312. https://doi.org/10.1371/journal.pbio.1002312

Kitazawa, S., Kohno, T., & Uka, T. (1995). Effects of delayed visual information on the rate and amount of prism adaptation in the human. *The Journal of Neuroscience*, *15*(11), 7644. https://doi.org/10.1523/JNEUROSCI.15-11-07644.1995

Klar, M., Stein, S., Paterson, F., Williamson, J. H., & Murray-Smith, R. (2025). An active inference model of mouse point-and-click behaviour. *arXiv Preprint arXiv:2510.14611*.

Kording, K. P., Tenenbaum, J. B., & Shadmehr, R. (2007). The dynamics of memory as a consequence of optimal adaptation to a changing body. *Nature Neuroscience*, *10*(6), 779–786. https://doi.org/10.1038/nn1901

Kouw, W. M., & Loog, M. (2018). An introduction to domain adaptation and transfer learning. *arXiv Preprint arXiv:1812.11806*.

Kuśmierz, Ł., Isomura, T., & Toyoizumi, T. (2017). Learning with three factors: Modulating Hebbian plasticity with errors. *Computational Neuroscience*, *46*, 170–177. https://doi.org/10.1016/j.conb.2017.08.020

Lanillos, P., Franklin, S., Maselli, A., & Franklin, D. W. (2021). Active strategies for multisensory conflict suppression in the virtual hand illusion. *Scientific Reports*, *11*(1), 22844. https://doi.org/10.1038/s41598-021-02200-7

Luauté, J., Schwartz, S., Rossetti, Y., Spiridon, M., Rode, G., Boisson, D., & Vuilleumier, P. (2009). Dynamic Changes in Brain Activity during Prism Adaptation. *The Journal of Neuroscience*, *29*(1), 169. https://doi.org/10.1523/JNEUROSCI.3054-08.2009

Mirza, M. B., Cullen, M., Parr, T., Shergill, S., & Moran, R. J. (2021). Contextual perception under active inference. *Scientific Reports*, *11*(1), 16223. https://doi.org/10.1038/s41598-021-95510-9

Newport, R., & Jackson, S. R. (2006). Posterior parietal cortex and the dissociable components of prism adaptation. *Neuropsychologia*, *44*(13), 2757–2765. https://doi.org/10.1016/j.neuropsychologia.2006.01.007

Paul, A., Isomura, T., & Razi, A. (2024). On Predictive Planning and Counterfactual Learning in Active Inference. *Entropy*, *26*(6). https://doi.org/10.3390/e26060484

Pawlak, V., Wickens, J. R., Kirkwood, A., & Kerr, J. N. D. (2010). Timing is not Everything: Neuromodulation Opens the STDP Gate. *Frontiers in Synaptic Neuroscience*, *2*, 146. https://doi.org/10.3389/fnsyn.2010.00146

Petitet, P., O'Reilly, J. X., & O'Shea, J. (2018). Towards a neuro-computational account of prism adaptation. *Special Issue: Lesions and Brain Mapping*, *115*, 188–203. https://doi.org/10.1016/j.neuropsychologia.2017.12.021

Pisella, L., Michel, C., Gréa, H., Tilikete, C., Vighetto, A., & Rossetti, Y. (2004). Preserved prism adaptation in bilateral optic ataxia: Strategic versus adaptive reaction to prisms. *Experimental Brain Research*, *156*(4), 399–408. https://doi.org/10.1007/s00221-003-1746-4

Pisella, L., Rossetti, Y., Michel, C., Rode, G., Boisson, D., Pélisson, D., & Tilikete, C. (2005). Ipsidirectional impairment of prism adaptation after unilateral lesion of anterior cerebellum. *Neurology*, *65*(1), 150–152. https://doi.org/10.1212/01.wnl.0000167945.34177.5e

Prablanc, C., Panico, F., Fleury, L., Pisella, L., Nijboer, T., Kitazawa, S., & Rossetti, Y. (2020). Adapting terminology: Clarifying prism adaptation vocabulary, concepts, and methods. *Neuroscience Research*, *153*, 8–21. https://doi.org/10.1016/j.neures.2019.03.003

Redding, G. M., Rossetti, Y., & Wallace, B. (2005). Applications of prism adaptation: A tutorial in theory and method. *Neuroscience & Biobehavioral Reviews*, *29*(3), 431–444. https://doi.org/10.1016/j.neubiorev.2004.12.004

Redding, G., & Wallace, B. (2002). Strategie Calibration and Spatial Alignment: A Model From Prism Adaptation. *Journal of Motor Behavior*, *34*, 126–138. https://doi.org/10.1080/00222890209601935

Redding, G., & Wallace, B. (2006). Generalization of Prism Adaptation. *Journal of Experimental Psychology. Human Perception and Performance*, *32*, 1006–1022. https://doi.org/10.1037/0096-1523.32.4.1006

Rossetti, Y., Koga, K., & Mano, T. (1993). Prismatic displacement of vision induces transient changes in the timing of eye-hand coordination. *Attention Perception & Psychophysics*, *54*, 355–364. https://doi.org/10.3758/BF03205270

Sakaguchi, Y., Akashi, Y., & Takano, M. (2001). Visuo-Motor Adaptation to Stepwise and Gradual Changes in the Environment: Relationship between Consciousness and Adaptation. *Journal of Robotics and Mechatronics*, *13*(6), 0601.

Smith, M. A., Ghazizadeh, A., & Shadmehr, R. (2006). Interacting Adaptive Processes with Different Timescales Underlie Short-Term Motor Learning. *PLOS Biology*, *4*(6), e179. https://doi.org/10.1371/journal.pbio.0040179

Tazawa, U. T., & Isomura, T. (2024). Synaptic pruning facilitates online Bayesian model selection. *bioRxiv : The Preprint Server for Biology*. https://doi.org/10.1101/2024.05.15.593712

Welch, R. B., & Goldstein, G. (1972). Prism adaptation and brain damage. *Neuropsychologia*, *10*(4), 387–394. https://doi.org/10.1016/0028-3932(72)90001-2

Zhuang, F., Qi, Z., Duan, K., Xi, D., Zhu, Y., Zhu, H., Xiong, H., & He, Q. (2020). A comprehensive survey on transfer learning. *Proceedings of the IEEE*, *109*(1), 43–76.

Acknowledgements

T.I. is supported by the Japan Society for the Promotion of Science (JSPS) KAKENHI under Grant Number JP23H04973, the Japan Science and Technology Agency (JST) CREST under Grant Number JPMJCR22P1, and the Japan Agency for Medical Research and Development (AMED) under Grant Number JP23wm0625001. The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Competing interest declaration

The authors declare no competing interests.