

RESEARCH

Open Access



# Enhanced “learning to learn” through a hierarchical dual-learning system: the case of action video game players

Yu-Yan Gao<sup>1,2,3,4</sup> , Zeming Fang<sup>2</sup> , Qiang Zhou<sup>4\*†</sup>  and Ru-Yuan Zhang<sup>1,2\*†</sup> 

## Abstract

In contrast to conventional cognitive training paradigms, where learning effects are specific to trained parameters, playing action video games has been shown to produce broad enhancements in many cognitive functions. These remarkable generalizations challenge the conventional theory of generalization that learned knowledge can be immediately applied to novel situations (i.e., immediate generalization). Instead, a new “learning to learn” theory has recently been proposed, suggesting that these broad generalizations are attained because action video game players (AVGPs) can quickly acquire the statistical regularities of novel tasks in order to increase the learning rate and ultimately achieve better performance. Although enhanced learning rate has been found for several tasks, it remains unclear whether AVGPs efficiently learn task statistics and use learned task knowledge to guide learning. To address this question, we tested 34 AVGPs and 36 non-video game players (NVGPs) on a cue-response associative learning task. Importantly, unlike conventional cognitive tasks with fixed task statistics, in this task, cue-response associations either remain stable or change rapidly (i.e., are volatile) in different blocks. To complete the task, participants should not only learn the lower-level cue-response associations through explicit feedback but also actively estimate the high-level task statistics (i.e., volatility) to dynamically guide lower-level learning. Such a dual learning system is modelled using a hierarchical Bayesian learning framework, and we found that AVGPs indeed quickly extract the volatility information and use the estimated higher volatility to accelerate learning of the cue-response associations. These results provide strong evidence for the “learning to learn” theory of generalization in AVGPs. Taken together, our work highlights enhanced hierarchical learning of both task statistics and cognitive abilities as a mechanism underlying the broad enhancements associated with action video game play.

**Keywords** Action video game, Learning to learn, Generalization, Hierarchical Gaussian Filter

<sup>†</sup>Qiang Zhou and Ru-Yuan Zhang co-senior authors.

\*Correspondence:

Qiang Zhou

zq@wmu.edu.cn

Ru-Yuan Zhang

ruyuanzhang@sjtu.edu.cn

<sup>1</sup> School of Psychology, Shanghai Jiao Tong University, Shanghai 200030, China

<sup>2</sup> Shanghai Mental Health Center, School of Medicine, Shanghai Jiao Tong University, Shanghai 200030, China

<sup>3</sup> Cixi Biomedical Research Institute, Wenzhou Medical University, Zhejiang 315302, China

<sup>4</sup> Department of Psychology, Wenzhou Medical University, Wenzhou 325035, China



## Introduction

Humans possess impressive and adaptable learning abilities, as evidenced by the rapid learning of diverse cognitive tasks and the flexible application of learned knowledge to unfamiliar scenarios. Optimizing learning and facilitating generalization has been a fundamental challenge in cognitive science. Traditional cognitive training often exhibits specificity to the training settings (tasks or parameters)—the improvement in learning are greatly reduced in previously unseen situations [1, 2]. If the benefits of cognitive training cannot efficiently generalize across different application situations, its real-world applicability is significantly diminished. Action video game training has been shown a unique training regime that can overcome such limitations. A large body of cognitive science research have shown that playing action video games can directly enhance a wide range of seemingly unrelated cognitive functions, such as attention [2, 3], memory [4–6], perception [2, 7, 8], and reasoning [9]. Importantly, players are not directly trained on these specific cognitive tasks when playing action video games. Because of these astonishingly broad generalizations, action video games have also been suggested as a useful paradigm for cognitive training [2] and even for therapeutic purposes [10]. As generalization is the key for observers to learn infinite knowledge based on finite learning samples, it is of paramount importance to understand the neuro-computational mechanisms of broad generalization induced by action video game play.

Why can action video game play lead to broad generalization in stark contrast to conventional training approaches? Classic theories of learning generalization postulate that an observer generalizes learned knowledge to novel cases by inferring the common constructs between training and application situations [11–13]. This view assumes that, once common constructs are identified, improvement on novel tasks is immediately achievable. This classic view is often referred to as “immediate generalization” [14, 15]. More recently, a new mechanism of generalization has been proposed, which suggests that action video game play induces broad generalizations by enabling observers to “learning to learn” [10]. In contrast to the “immediate generalization” theory, the “learning to learn” theory predicts that avid action video game players (AVGPs) can quickly capture the underlying structural knowledge of new tasks and thus accelerate learning. Faster learning (i.e., taking less time to achieve good performance) on new tasks, as a hallmark of “learning to learn”, has been found in several recent studies of action video games [16] and classical perceptual learning [7–9].

Although “learning to learn” is an elegant theory that can potentially explain the remarkable generalization

afforded by action video game play, two issues remain unresolved. First, in addition to predicting faster learning of novel tasks, the “learning to learn” theory has two other key predictions — (1) action video game players (AVGPs) can estimate and understand task statistics more quickly and accurately, and (2) the learned task statistics can in turn guide faster learning of a task. However, the enhanced ability of AVGPs to learn the statistical structure of tasks has not been directly investigated. Second, the “learning to learn” theory also implicitly assumes that, even in an apparently simple task, a hierarchical dual learning system operates: a high-level system for learning task statistics and a lower-level system for learning appropriate responses. Previous studies only assessed observers’ learning behavior as a result of the low-level learning system. It remains unclear whether a high-level learning system exists and how it supports the lower-level response learning. To address these two questions directly, two factors should be considered. First, to demonstrate the superior ability of AVGPs to extract task statistics, we need a task with systematic variation in stimulus statistical regularities and test whether AVGPs are indeed sensitive to such variation. Second, the “learning to learn” ability should be explicitly formulated. In other words, a computational framework is needed to explicitly specify how the correct decisions emerge according to the interactions within the dual-learning system in an online fashion.

In this study, we aim to directly test the “learning to learn” theory using a volatile reversal learning task [17, 18]. In this task, participants learn the associations between a visual cue and its corresponding response through trial-by-trial feedback. Importantly, such cue-response associations either remain stable over several trials (i.e., stable block) or change rapidly on other trials (i.e., volatile block, see Methods for details). This volatility variation allows us to assess participants’ ability to learn such task statistics, and, unlike classic learning tasks [19–21], such an associative learning task also allows us to explicitly estimate participants’ learning rate at both levels. Furthermore, we used the Hierarchical Gaussian Filter (HGF, [22]) to formulate the “learning to learn” process. In particular, unlike classical reinforcement learning models that only formulate the learning of cue-response associations [23, 24], the HGF also specifies a high-level learning process of task statistics (i.e., association volatility). Importantly, changes in the lower-level cue-response associations lead to trial-by-trial updates in the high-level belief of association volatility, and the high-level estimates of association volatility in turn adjust the rate of the lower-level association learning. These bidirectional interactions between a hierarchical

dual-learning systems exactly corresponds to the “learning to learn” hypothesis.

Our results show that AVGPs display higher learning rates in the volatile reversal learning task, consistent with previous studies. Most importantly, this higher learning rate is a result of an efficient representation of the association volatility, as evidenced by a higher estimate of association volatility in the AVGPs. All these results are consistent with the “learning to learn” theory of action video game play.

## Materials and methods

### Ethics and participants

All experimental protocols were approved by the institutional review board of Shanghai Jiao Tong University. All research was conducted in accordance with relevant guidelines and regulations. Informed written consent was obtained from all participants.

We firstly administered the Chinese version of the Video-Game-Expertise Classification Scheme [25] to screen for action video game players (AVGPs) and non-video game players (NVGPs). Both English and Chinese versions of the video game questionnaire can be downloaded from <https://www.unige.ch/fapse/brainlearning/vgq/>. The basic inclusion criteria require participants to have Chinese as a first or second language; normal or corrected-to-normal vision; no history of mental disorders; not taking significant psychiatric medications; and an age range of 18 to 40 years old. NVGPs need to meet the following criteria: (1) play first/third-person shooter, action/sports, real-time strategy/ Multiplayer Online Battle Arena (MOBA) games, or simulation games for no more than 1 h/week in the past year and the year before; (2) play any other type of games for no more than 3 h/week in the past year; (3) play any other type of games for no more than 5 h/week a year ago. AVGPs need to meet any of the following criteria: (1) play other games for no more than 3 h/week, but play action games for at least 5 h/week in the past year; (2) play action games for at least 3 h/week in the past year, with other games not exceeding 3 h/week, and play action games for at least 5 h/week a year ago; (3) play other games for no more than 3 h/week in the past year, with at least 3 h/week for action games and at least 5 h/week for sports/driving games; (4) play other games for no more than 3 h/week in the past year, with at least 3 h/week for action games and at least 5 h/week for real-time strategy/MOBA games. There exist other inclusion criteria for both groups. More detailed screening criteria can be found in the questionnaire above.

Previous studies have documented several important ingredients of AVGs that enable generalization effect,

including (i) decision-making under time constraints, (ii) maintaining divided attention, and (iii) the necessity for prompt transitions between two distinct attentional states (focused and divided) [1, 26, 27]. These factors have also been incorporated into a number of other game genres, including sports and driving games, as well as real-time strategy and MOBA games. We thus also include these genres in the screening of AVGPs.

Based on the filtering criteria, 34 AVGPs (24 males and 10 females) and 36 NVGPs (12 males and 24 females) were recruited to participate in the formal experiment after obtaining their consents. All participants were right-handed and had normal or corrected-to-normal vision. After excluding the subjects who exhibited extreme performance (see data analysis below), data from 33 AVGPs (23 males and 10 females) and 34 NVGPs (12 males and 22 females) were included for further analysis.

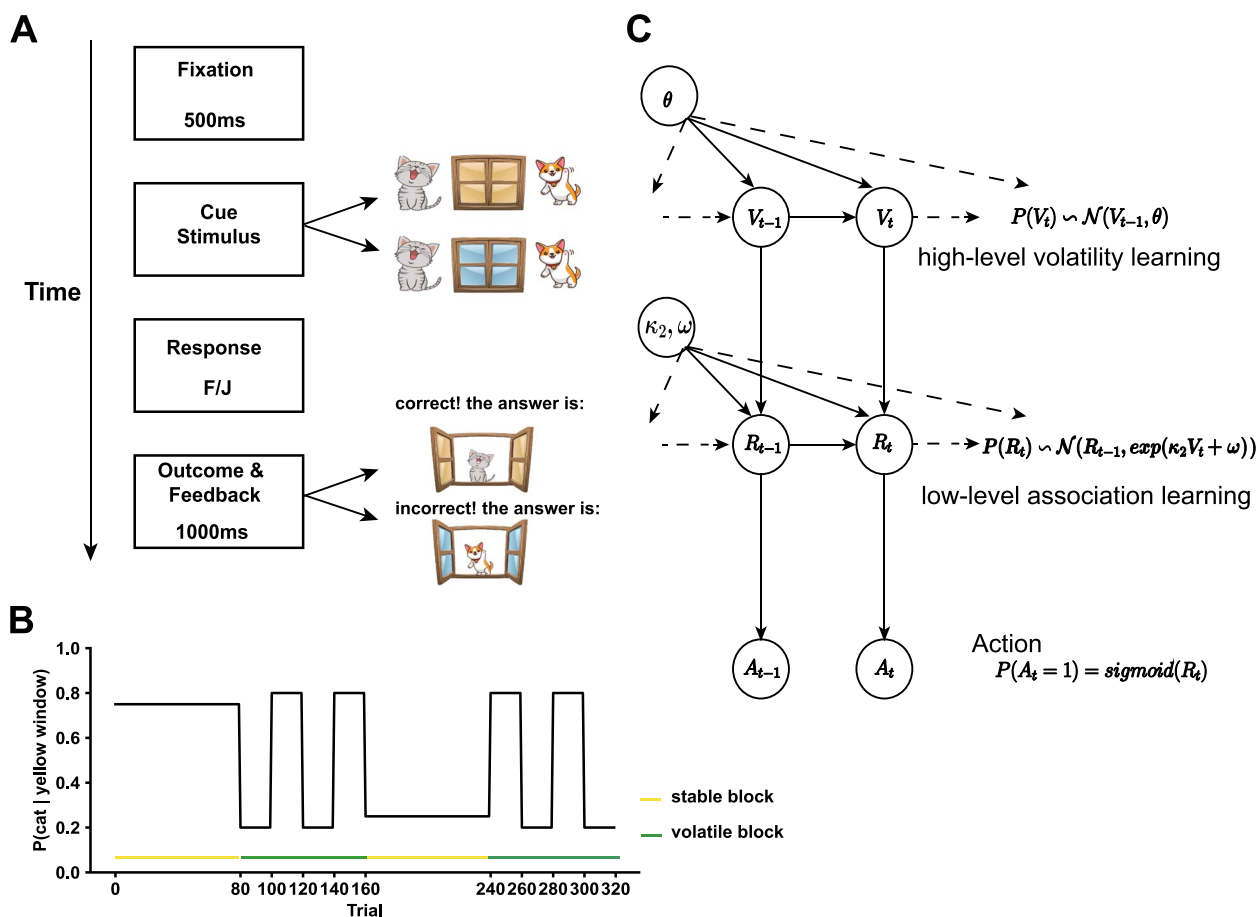
### Stimulus and task

This experiment was hosted on the Naodao platform (<https://www.naodao.com/>). Participants accessed the task remotely and completed it online. They received the corresponding participant compensation after the experiment.

Both AVGPs and NVGPs performed the same volatile reversal learning task (Fig. 1A). Each trial began with a 500 ms fixation period. A cue stimulus (i.e., a yellow or a blue window) was presented. The cue stimulus disappeared after the participant made a keypress response to predict which outcome stimulus (i.e., a cat or a dog) was more likely to appear after the cue stimulus. After the keypress response, an outcome stimulus was presented for 1000 ms. The whole experiment consists of four blocks (80 trials per block) with a total of 320 trials. In each block, the association settings between the cue and outcome stimuli were changed (Fig. 1B).

Here, association is defined as the probability of a cue-response pair. For example, in the first 80 trials, the outcome stimulus cat (or dog) appeared after the cue stimulus yellow window with a probability of 0.75 (or 0.25, respectively). Similarly, the association “blue window-dog” is 0.75. The association settings changed in each block (Fig. 1B). The key point here is that the association setting is stable (i.e., stable condition) in Block 1 (i.e., trials 1–80) and Block 3 (i.e., trials 161–240) but switches rapidly between 0.8 and 0.2 (i.e., volatile condition) in Block 2 (i.e., trials 81–160) and Block 4 (i.e., trials 241–320).

The stimulus materials for this task were created using Photoshop, and each stimulus material has a resolution of 1080×720. The presentation order of the stimuli was pseudorandomized and generated in MATLAB 2020a



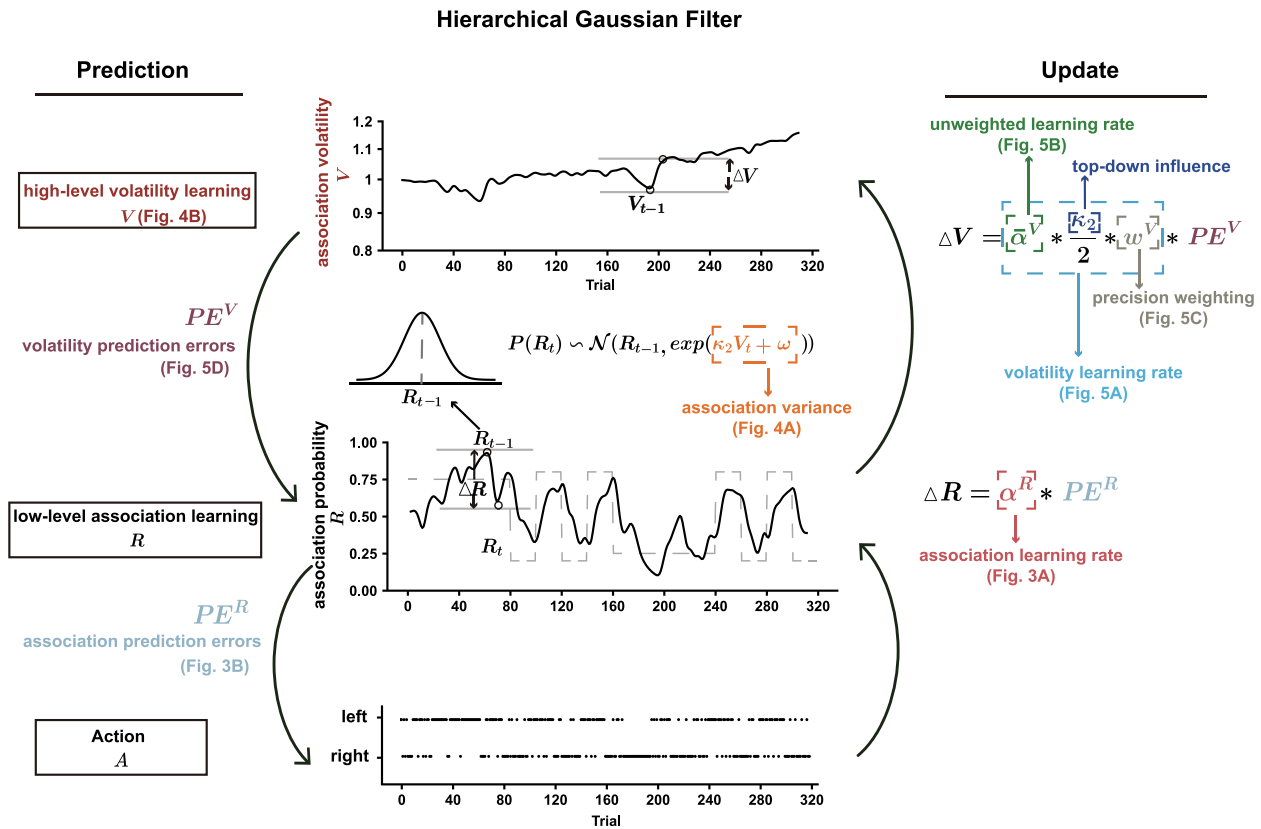
**Fig. 1** Task design and model. **A** Each trial started with a fixation cross in the center of the screen. After a delay of 500 ms, a stimulus was presented on the screen. Participants were instructed to predict the animal behind the window based on the current yellow or blue window and press the ‘F’ key for a cat or the ‘J’ key for a dog. Immediate feedback and outcome stimuli were provided after each response, lasting for 1000 ms before proceeding to the next trial. **B** The experiment was divided into four blocks based on the probability of cue-response association: stable (trials 1–80,  $p=0.75$ )—volatile (trials 81–160, with a switching sequence of  $p$  values: 0.2–0.8–0.2–0.8)—stable (trials 161–240,  $p=0.25$ )—volatile (trials 241–320 with a switching sequence of  $p$  values: 0.8–0.2–0.8–0.2). The yellow line parallel to the x-axis represents trials in the stable blocks, the green line represents trials in the volatile blocks. In the stable blocks, the association probability remained constant within 80 trials, while in the volatile blocks, the probability changed every 20 trials. **C** Generative process of the HGF.  $A$  represents action;  $R$  indicates the estimated association probability between the given window cue and the corresponding animal response;  $V$  represents the estimated association volatility.  $t$  denotes each time point.  $A_t$  depends on  $R_{t-1}$ ,  $V_{t-1}$ , and parameters  $\theta, \kappa_2, \omega$ . The interconnection between levels is achieved through uncertainty

according to the number of trials in each experimental block and the four cue-response association probabilities. The presentation order of the cues within the experimental block was fixed by a predetermined shuffled order. Thus, each participant received the same stimulus sequence, allowing for a comparable learning process and model parameter estimation. The experimental procedure was developed using jsPsych-6.3.0 (<https://www.jspsych.org/6.3/>). Participants were informed that these probabilities would change, but were not given with specific

information about the four blocks and the exact values of the probabilities.

**Computational modeling**

The HGF [22] model is used to analyze the participants’ behavior. We plotted and compared the trial-by-trial generated data from two groups of participants. At the same time, we used t-tests to compare the parameters of the two groups of participants.



**Fig. 2** Overview of the HGF model. The probability at each level is determined by the previous level and parameters. Throughout the paper, we analyzed several key variables of this model. We color labeled the variables of interest and illustrate the figure number where the group differences in the variables are compared to facilitate reading

**Generative model**

The HGF can be understood via two distinct components: prediction and update. Briefly, this model formulates the prediction and update process in a two-level hierarchy (Fig. 2). The prediction (i.e., generative) process can be seen in Fig. 1C and the left part of Fig. 2. Specifically, the higher level of the model represents the estimated association volatility ( $V$ ) (i.e., how quickly the cue-response associations switch), which is updated by

$$P(V_t|V_{t-1}) = \mathcal{N}(V_t; V_{t-1}, \theta) \tag{1}$$

where  $\theta$  is a constant parameter which determines the variance of estimated association volatility (the high-level,  $V$ ). Estimated association volatility  $V$  determines the magnitude for updating the lower-level cue-response association ( $R$ , the estimated association probability between the given window cue and the corresponding behavioral choices in the logarithmic domain).

$$P(R_t|R_{t-1}, V_t) = \mathcal{N}(R_t; R_{t-1}, \exp(\kappa_2 V_t + \omega)) \tag{2}$$

where  $\kappa_2$  is the top-down influence factor that determines the coupling strength between the association probability (the low-level,  $R$ ) and the estimated association volatility (the high-level,  $V$ );  $\omega$  is a constant component of the association variance ( $\kappa_2 * V_t + \omega$ ), independent of the state of the estimated association volatility (the high-level,  $V$ ). The behavioral action  $A$  is generated by the association probability ( $R$ ), and  $\mu$  (i.e., correct or incorrect) is the actual outcome the participant received.

$$P(A_t|R_t) = \text{Bernouli}(A_t; s(\kappa_1 * R_t)) \tag{3}$$

$$P(\mu_t|A_t) = (\mu_t)^{A_t} (1 - \mu_t)^{1-A_t} \tag{4}$$

where the function  $s(\cdot)$  is the sigmoid function with  $\kappa_1$  as the inverse temperature. To simplify our modeling, we fixed the coupling factor controlling the influence of association probability (the low-level,  $R$ ) on action (i.e.,  $\kappa_1$ ) to 1.

This model has three free parameters : $\theta, \kappa_2$ , and  $\omega$ .

**Trial-by-trial update rule of model parameters**

The detailed trial-by-trial update rule of model parameters in HGF has been documented in Mathys, et al. [22]. Furthermore, this update process is illustrated in the right part of Fig. 2. Here we provide a short overview and an introduction of the variables and free parameters.

On the  $t$ -th trial, the action ( $A_t$ ) is determined by the actual outcome the subject received ( $\mu_t$ ), where  $\mu_t \in \{0,1\}$  indicates the correct/incorrect feedback.

$$A_t = \mu_t \tag{5}$$

The update of estimated association probability ( $R_t$ ) depends on the association learning rate ( $\alpha_t^R$ ) and the association prediction errors ( $PE_t^R$ ).

$$\Delta R_t = \alpha_t^R * PE_t^R \tag{6}$$

Note that the association learning rate ( $\alpha_t^R$ ) varies trial-by-trial and is determined by association expectation ( $\hat{\alpha}_t^R$ ) and action expectation ( $\hat{\alpha}_t^A$ ). The superscript  $R$  denotes the variables as the ones operating at the low-level association learning.

$$\alpha_t^R = \frac{1}{\frac{1}{\hat{\alpha}_t^R} + \hat{\alpha}_t^A} \tag{7}$$

The association expectation ( $\hat{\alpha}_t^R$ ) *per se* also varies trial-by-trial and is determined by the learning rate of the last trial ( $\alpha_{t-1}^R$ ) and the upper-level estimated association volatility ( $V_{t-1}$ ), where  $\kappa_2$  and  $\omega$  are free parameters.

$$\hat{\alpha}_t^R = \alpha_{t-1}^R + e^{(\kappa_2 * V_{t-1} + \omega)} \tag{8}$$

The action expectation ( $\hat{\alpha}_t^A$ ) *per se* also varies trial-by-trial and is determined by the action of the last trial ( $A_{t-1}$ ).

$$\hat{\alpha}_t^A = A_{t-1} * (1 - A_{t-1}) \tag{9}$$

the association prediction errors ( $PE_t^R$ ) is given by:

$$PE_t^R = A_t - \text{sigmoid}(R_{t-1}) \tag{10}$$

The update of the estimated volatility ( $V_t$ ) depends on the volatility learning rate ( $\alpha_t^V$ ) and the volatility prediction errors ( $PE_t^V$ ) The superscript  $V$  denotes the variables as the ones operating at the high-level volatility learning.

$$\Delta V_t = \alpha_t^V * PE_t^V \tag{11}$$

where the volatility learning rate ( $\alpha_t^V$ ) consists of three components:

$$\alpha_t^V = \bar{\alpha}_t^V * \frac{\kappa_2}{2} * w_t^V \tag{12}$$

Here,  $\bar{\alpha}_t^V$  represents unweighted volatility learning rate of  $V$  and varies trial-by-trial:

$$\bar{\alpha}_t^V = \frac{1}{\bar{\alpha}_{t-1}^V + \theta} + \frac{\kappa_2^2}{2} * w_t^V * (w_t^V + r_t^V * PE_t^V) \tag{13}$$

where  $\theta$  is a free parameter.  $w_t^V$  denotes a precision weighting factor.

$$w_t^V = \frac{e^{(\kappa_2 V_{t-1} + \omega)}}{\alpha_{t-1}^R + e^{(\kappa_2 V_{t-1} + \omega)}} \tag{14}$$

$$r_t^V = \frac{e^{(\kappa_2 V_{t-1} + \omega)} - \alpha_{t-1}^R}{\alpha_{t-1}^R + e^{(\kappa_2 V_{t-1} + \omega)}} \tag{15}$$

the volatility prediction errors ( $PE_t^V$ ) is given by:

$$PE_t^V = \frac{\alpha_t^R + (R_t - R_{t-1})^2}{\alpha_{t-1}^R + e^{(\kappa_2 V_{t-1} + \omega)}} - 1 \tag{16}$$

In summary, the estimated free parameters for each participant are  $\kappa_2$ ,  $\omega$ , and  $\theta$ . The variables with subscript “ $t$ ” change from trial to trial, and the three free parameters without subscript “ $t$ ” are fixed values that hold for all trials.

The analysis was performed using the HGF toolbox in MATLAB (<https://translationalneuromodeling.github.io/tapas>). The `tapas_fitModel` function was used to iteratively fit the model 100 times for each participant, using the Maximum A Posteriori (MAP) method for parameter estimation. Configuration settings, facilitated by functions such as `tapas_hgf_binary_config`, `tapas_unitsq_sgm_config`, and `tapas_quasineutron_optim_config`, were used to set prior ranges for the parameters. The ranges of priors for the parameters to be fitted are as follows: top-down factor  $\log(\kappa_2) \sim \mathcal{N}(\log(1), 4)$ ; association constant uncertainty  $\omega \sim \mathcal{N}(-3, 16)$ ; volatility constant uncertainty  $\log(\theta) \sim \mathcal{N}(-6, 16)$ . All other parameters involved in the code, including their ranges and initial values, follow the default settings in the toolbox.

**Statistical analysis**

Linear mixed model analysis was performed in JASP 0.18.1.0 (<https://jasp-stats.org/>), and all multiple comparisons were corrected using the Holm correction in JASP. All t-tests were performed using the *Pingouin* package in Python and were all two-tailed. In this experiment, participants with an average association learning rate exceeding (or fall below) the mean plus (or minus) two

standard deviations of the overall sample were excluded. A total of 4 participants met these criteria. 33 AVGPs and 34 NVGPs were included in the reported results.

## Results

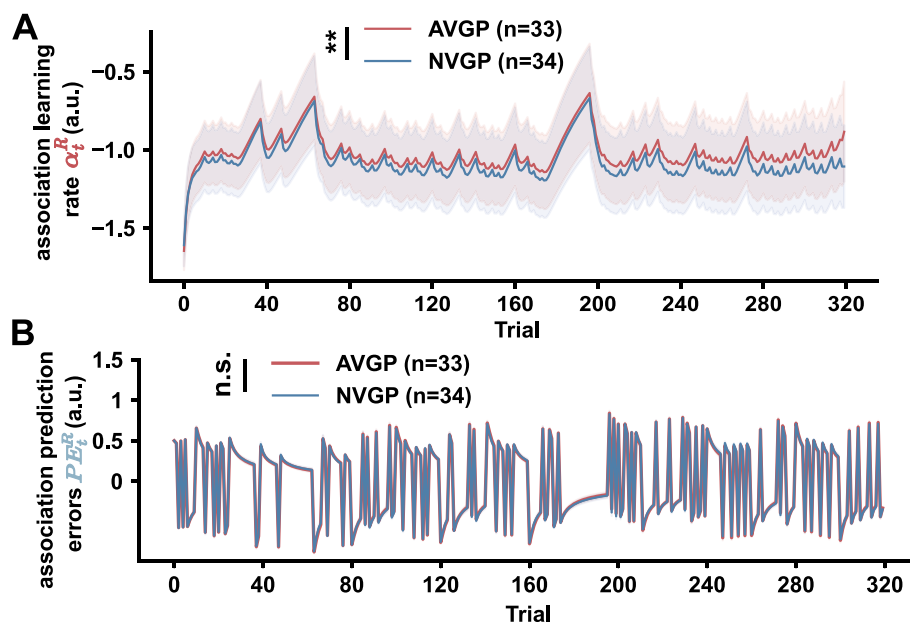
### Superior low-level learning rate of cue-response associations in AVGPs

Participants performed a volatile reversal learning task (Fig. 1A). On each trial, a fixation was shown for 500 ms and followed by a cue stimulus (i.e., a yellow window or a blue window). Participants were asked to predict the subsequent outcome stimulus (i.e., a cat or a dog) associated with the cue. Following a keypress response, an outcome stimulus was presented for 1000 ms as feedback. The two cue stimuli and the two outcome stimuli were paired. For example, within a stable block, the cat (or dog) appeared after the yellow window (or blue) window in 75% (or 25%, respectively) of the trials. Such cue-response associations varied across blocks. Importantly, the task statistic is defined as the changing rate of such cue-response associations (i.e., volatility). In particular, in the two stable blocks (Block 1, trials 1–80; Block 3, trials 161–240), the cue-response association settings remained constant. In contrast, in the two volatile blocks (Block 2, trials 81–160; Block 4, trials 241–320), the cue-response associations switched

between 0.8 and 0.2 every 20 trials. The key question here is whether participants can learn the stability and volatility of the associations and use this information to guide their learning. Followed by the conventional approach [17, 28], we directly fitted computational models (see below) to represent participants' learning process in this task.

We first asked whether we could replicate the finding that AVGPs learn a novel task faster than NVGPs [7, 10, 16, 29]. Unlike the conventional reinforcement learning approach that estimates a single learning rate parameter throughout the task [30, 31], HGF assumes that participants' learning rate also varies from trial to trial based on updated beliefs about the task statistics (i.e., volatility). In this task, participants learned the cue-response associations. The trial-by-trial association learning rate ( $\alpha_t^R$ , Eqs. 6–8) in both groups is plotted as a function of trials in Fig. 3A.

A linear mixed model (LMM) was built in JASP with *Trial* as a random effect factor, *Group* (AVGPs/NVGPs) as a fixed effect factor, *Log Association Learning Rate* ( $\alpha_t^R$ , Eqs.7–9) in each trial as the dependent variable. We found that the effect of *Group* is significant, indicating the overall higher learning rate of the AVGPs than that of the NVGPs ( $t_{(21119)}=2.852$ ,  $p=0.004$ ,  $Estimate=0.055$ ,  $SE=0.019$ ,  $CI=[0.017, 0.093]$ ). In summary, Fig. 3A



**Fig. 3** Comparison of association learning rate between two groups. **A** The log association learning rate ( $\alpha_t^R$ , Eqs.7, 8, 9) required for updating the estimated association probability for each participant. The x-axis represents the trial sequence ( $t$ ), and the y-axis illustrates participants' log association learning rate ( $\alpha_t^R$ ). The red line represents AVGPs, and the blue line represents NVGPs. The shaded area represents S.E.M across all participants within each group (33 AVGPs, 34 NVGPs). Significance symbol conventions is \*\*:  $p < 0.01$ . **B** Two groups' association prediction errors ( $PE_t^R$ , Eqs. 6&10) across trials. The x-axis represents the trial sequence ( $t$ ), the y-axis illustrates association prediction errors ( $PE_t^R$ ). Significance symbol convention is n.s.: non-significant

shows that the AVGPs indeed had a generally higher learning rate than the NVGPs, although the learning rate in both groups varied from trial to trial in both groups.

Because the trial-by-trial update of the association probability ( $\Delta R_t$ , Eq. 6) is determined by both association learning rate ( $\alpha_t^R$ ) and association prediction errors ( $PE_t^R$ , Eqs. 6&10), we also analyzed the association prediction errors ( $PE_t^R$ ) in both groups and plotted them as a function of trials in Fig. 3B. A LMM was performed with the Association Prediction Errors ( $PE_t^R$ ) as the dependent variable; Group (AVGPs/NVGPs) as a fixed effect factor and Trial as a random effect factor. We found no significant effect of Group ( $t_{(21119)} = -0.036$ ,  $p = 0.971$ , Estimate = -0.001,  $SE = 0.002$ ,  $CI = [-0.003, 0.003]$ ), suggesting the superior learning in AVGPs is mostly due to the association learning rate rather than association prediction errors.

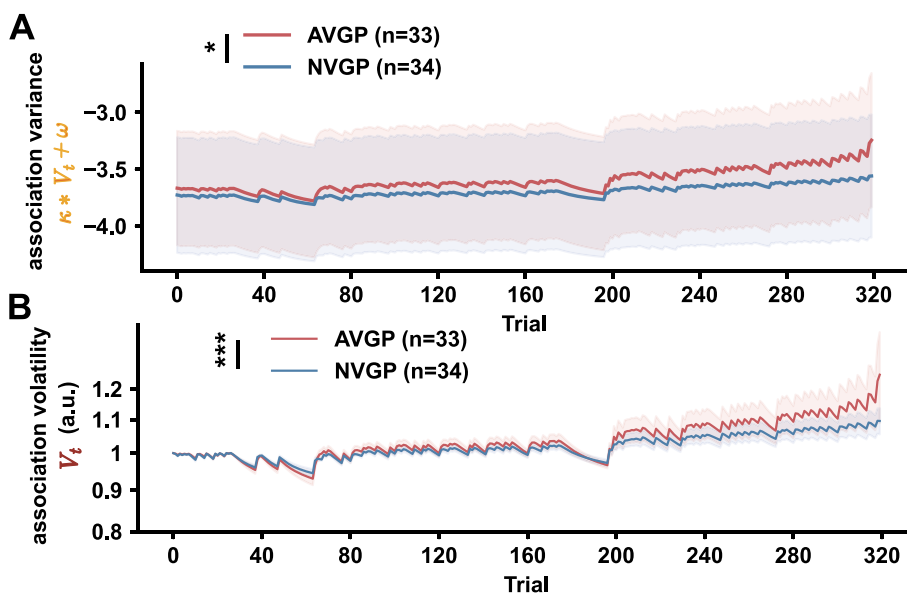
**Higher low-level learning rate in AVGPs is due to high-level association volatility**

We have confirmed the overall higher association learning rate in AVGPs. A higher association learning rate ( $\alpha_t^R$ ) leads to a larger update ( $\Delta R_t$ ) of the estimated association probability. But how did AVGPs develop an overall higher association learning rate in the volatile reversal task? The key aspect of the HGF is that association learning rate is determined by association variance in the last

trial ( $\kappa_2 * V_{t-1} + \omega$ ), which is further controlled by high-level volatility  $V_{t-1}$  in the last trial (Eqs. 7–9). Here, we examine whether higher association variance leading to an increased association learning rate in the AVGPs.

A LMM was performed with Association variance ( $\kappa_2 * V_t + \omega$ ) as the dependent variable; Group (AVGPs/NVGPs) as a fixed effect factor and Trial as a random effect factor. We found that the effect of the Group was significant, indicating overall greater association variance of AVGPs compared to NVGPs ( $t_{(21119)} = 2.516$ ,  $p = 0.012$ , Estimate = 0.100,  $SE = 0.040$ ,  $CI = [0.022, 0.179]$ , Fig. 4A). For completeness, in addition to the association variance ( $\kappa_2 * V_t + \omega$ ) and the association learning rate from the previous trial ( $\alpha_{t-1}^R$ ), we also compared action expectation ( $\hat{\alpha}_t^A$ , Eqs. 6&7) that contribute to the update of association learning rate (Eq. 7). We found no significant effect of Group ( $t_{(21119)} = -0.071$ ,  $p = 0.944$ , Estimate = -0.001,  $SE = 0.005$ ,  $CI = [-0.010, 0.009]$ ). This suggests that the higher association learning rate ( $\alpha_t^R$ ) observed in AVGPs is likely due to their overall higher association variance ( $\kappa_2 * V_t + \omega$ ).

The association variance ( $\kappa_2 * V_t + \omega$ ) is determined by the linear addition of two components: a top-down component ( $\kappa_2 * V_t$ ) and a constant component ( $\omega$ ). The top-down component indicates that a higher estimated association volatility ( $V_t$ ) leads to a larger update of the association learning rate, where  $\kappa_2$  is the



**Fig. 4** Association variance and estimated association volatility in two groups. **A** Participants’ association variance ( $\kappa_2 * V_t + \omega$ ) across trials. The x-axis represents the trial sequence ( $t$ ), and the y-axis illustrates participants’ association variance ( $\kappa_2 * V_t + \omega$ ). The red line represents AVGPs, and the blue line represents NVGPs. The shaded area represents S.E.M across all participants within each group (33 AVGPs, 34 NVGPs). Significance symbol conventions is \*:  $p < 0.05$ . **B** Participants’ estimated log association volatility ( $V_t$ ) across trials. The x-axis represents the trial sequence ( $t$ ), and the y-axis illustrates participants’ estimated association volatility ( $V_t$ ). The y-axis is plotted on a logarithmic scale. The red line represents AVGPs, and the blue line represents NVGPs. Significance symbol conventions is \*\*\*:  $p < 0.001$



top-down coupling factor. The constant step indicates the default magnitude of the update in the subject. Note that the top-down factor  $\kappa_2$  and the association constant step  $\omega$  are considered as traits of each subject and are fixed across trials, while the high-level estimated association volatility  $V_t$  varied across trials.

Next, we sought to understand which factor of association variance contributed most to the increased association learning rate. There were no significant differences in both  $\kappa_2$  ( $t_{(58.569)} = -0.236$ ,  $p = 0.814$ , *Cohen's*  $d = 0.058$ ,  $CI = [-0.320, 0.250]$ ) and  $\omega$  ( $t_{(64.677)} = -0.597$ ,  $p = 0.552$ , *Cohen's*  $d = 0.146$ ,  $CI = [-1.740, 0.940]$ ). A LMM was performed with *Estimated Association Volatility* ( $V_t$ ) in each trial as the dependent variable, *Group* (AVGPs/NVGPs), *Block Type* (stable/volatile), and their interaction as the fixed effect factors, and *Trial* as a random effect factor. The “learning to learn” theory predicts that AVGPs should be more sensitive to task statistics (i.e., volatility). Indeed, we found that AVGPs estimated higher association volatility than NVGPs ( $t_{(21116)} = 8.453$ ,  $p < 0.001$ , *Estimate* = 0.073,  $SE = 0.009$ ,  $CI = [0.056, 0.090]$ ). Post-hoc pairwise comparisons revealed that AVGPs had significantly higher estimated association volatility ( $V_t$ ) than NVGPs in the second stable block (stable block 2,  $t_{(21116)} = 3.737$ ,  $p < 0.001$ , *Estimate* = 0.016,  $SE = 0.004$ ,  $CI = [0.008, 0.025]$ ) and two volatile blocks (volatile block 1,  $t_{(21116)} = 2.378$ ,  $p = 0.017$ , *Estimate* = 0.010,  $SE = 0.004$ ,  $CI = [0.002, 0.019]$ ; volatile block 2,  $t_{(21116)} = 11.1778$ ,  $p < 0.001$ , *Estimate* = 0.048,  $SE = 0.004$ ,  $CI = [0.040, 0.057]$ ) but not in the first stable block (stable block 1,  $t_{(21116)} = -0.387$ ,  $p = 0.698$ , *Estimate* = -0.002,  $SE = 0.004$ ,  $CI = [-0.011, 0.007]$ , Fig. 4B). This may be because the first block was a stable block. These results show that the AVGPs can detect relatively higher association volatility ( $V_t$ ) as the task proceeds and then produce a greater trial-by-trial update of the association learning rate, resulting in faster learning of low-level associations. This process is consistent with the “learning to learn” theory that AVGPs can quickly adapt to ever-changing task environments.

Furthermore, we found that the estimated association volatility  $V_t$  during the volatile blocks was significantly higher than that during the stable blocks in both groups ( $t_{(316.125)} = 19.862$ ,  $p < 0.001$ , *Estimate* = 0.183,  $SE = 0.009$ ,  $CI = [0.164, 0.201]$ ). This result indicates that both groups can indeed recognize the different levels of volatility of the task. This is also consistent with the well-established theory in reinforcement learning that an agent should relatively increase learning rate in a volatile reward environment [32].

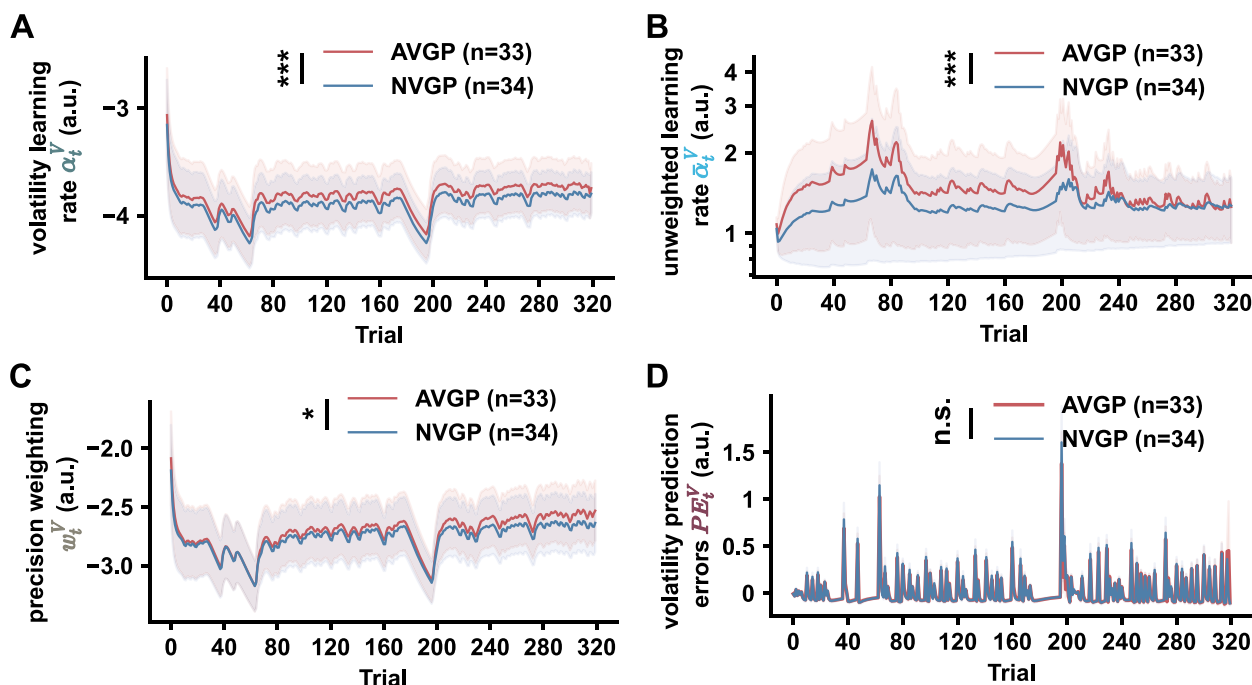
### Superior high-level learning rate of tasks statistics in AVGPs

The above results suggest that AVGPs subjectively experience a higher high-level association volatility ( $V_t$ ) and use this information to increase the low-level association learning rate ( $\alpha_t^R$ ). Here, we further asked how AVGPs learn the task statistics and obtain the higher association volatility. Again, we examined the volatility learning rate ( $\alpha_t^V$ , Eq. 12), which indicates how quickly the association volatility ( $V_t$ ) evolves across trials. The volatility learning rate is plotted as a function of trials in Fig. 5A. A LMM was performed with *Log Volatility Learning Rate* as the dependent variable; *Group* (AVGPs/NVGPs) as a fixed effect factor and *Trial* as a random effect factor. We found that the volatility learning rate of AVGPs consistently exceeded that of NVGPs' ( $t_{(211119)} = 3.995$ ,  $p < 0.001$ , *Estimate* = 0.081,  $SE = 0.020$ ,  $CI = [0.041, 0.120]$ ).

It was mentioned earlier that an advantage of the HGF model over traditional reinforcement learning models is that the precision-weighted learning rates (including the association learning rate and the volatility learning rate) in HGF can vary from trial to trial, allowing more flexible adaptation of individual beliefs to volatilities. According to the HGF model (Eq. 12,  $\alpha_t^V = \bar{\alpha}_t^V * \frac{\kappa_2}{2} * w_t^V$ ), the volatility learning rate ( $\alpha_t^V$ ) is determined by three factors: the unweighted volatility learning rate  $\bar{\alpha}_t^V$  (see Eq. 13), the top-down factor  $\kappa_2$  introduced above, and the precision weighting factor ( $w_t^V$ , Eq. 14) of the volatility prediction errors ( $PE_t^V$ , Eq. 16). Note that  $\bar{\alpha}_t^V$  and  $w_t^V$  varied from trial to trial but  $\kappa_2$  is a fixed value in each subject.

The trial-by-trial unweighted volatility learning rate, precision weighting factor, and volatility prediction errors are plotted as function of trials in Fig. 5B-D. Three LMMs were performed with *Unweighted Volatility Learning Rate* ( $\bar{\alpha}_t^V$ ), *Precision Weighting Factor* ( $w_t^V$ ), and *Volatility Prediction Errors* ( $PE_t^V$ ) as the dependent variables; *Group* (AVGPs/NVGPs) as the fixed effect factor and *Trial* as a random effect factor. We found that AVGPs had an overall higher unweighted learning rate ( $t_{(21119)} = 5.142$ ,  $p < 0.001$ , *Estimate* = 0.219,  $SE = 0.043$ ,  $CI = [0.136, 0.303]$ ) and an overall higher precision weighting ( $t_{(21119)} = 2.459$ ,  $p = 0.014$ , *Estimate* = 0.048,  $SE = 0.020$ ,  $CI = [0.010, 0.087]$ ) than NVGPs. However, there was no group difference on the volatility prediction errors ( $t_{(21119)} = -0.767$ ,  $p = 0.443$ , *Estimate* = -0.003,  $SE = 0.004$ ,  $CI = [-0.010, 0.005]$ ).

Taken together, we found that AVGPs can perceive higher association volatility because they can learn volatility *per se* faster (i.e., higher volatility learning rate) rather than because of higher volatility prediction errors. This higher volatility learning rate is augmented by more optimal uncertainty processing (i.e., higher precision weighting factor).



**Fig. 5** Volatility learning in two groups. **A** The log volatility learning rate ( $\alpha_t^V$ ) over all trials of the two groups. The x-axis represents the trial sequence, and the y-axis reflects the volatility learning rate. The red line represents the AVGPs, and the blue line represents the NVGPs. The shaded area represents S.E.M across all participants within each group (33 AVGPs, 34 NVGPs). Significance symbol conventions is \*\*\*:  $p < 0.001$ . **B** The unweighted volatility learning rate ( $\bar{\alpha}_t^V$ ) of the two groups across trials. The y-axis is plotted on a logarithmic scale. **C** The precision weighting factor ( $w_t^V$ ) of the association prediction errors of the two groups across trials. Significance symbol conventions is \*:  $p < 0.005$ . **D** the volatility prediction errors ( $PE_t^V$ ) of the two groups across trials. Significance symbol convention is n.s.: non-significant

## Discussion

The theory of “learning to learn” has recently been proposed as a novel mechanism of learning generalization [10], in particular the broad cross-task generalizations found in avid AVGPs. In this study, we proposed that enhanced “learning to learn” in AVGPs is achieved by an improved hierarchical dual learning system that takes into account both low-level cue-response associations and high-level task statistics (i.e., volatility). 34 AVGPs and 36 NVGPs completed a volatile reversal learning task in which participants should learn both cue-response associations and the temporal volatility of these associations (i.e., task statistics). We used Hierarchical Gaussian Filter (HGF) to quantify both low-level association learning and high-level volatility learning in the two groups and made three main observations. First, consistent with “learning to learn” and previous results, we found that AVGPs indeed exhibit a higher low-level learning rate of cue-response associations. Second, the higher low-level learning rate of associations is primarily driven by a higher high-level volatility on a trial-by-trial basis. Third, we further investigated the evolution of estimated volatility and found that the high-level learning rate of volatility *per se* is also higher in the AVGP group. These results

strongly support the “learning to learn” theory of action video game play and show that AVGPs can quickly learn the intrinsic statistics of novel tasks and use the learned task knowledge to guide low-level learning of correct responses. Our work sheds new light on generalization in action video games and, more broadly, on cognitive training in general.

### Two aspects of “learning to learn”

“Learning to learn” has two key components—enhanced learning rate and multi-level hierarchical learning.

Within the framework of “learning to learn”, enhanced learning rate in novel tasks is a new form of learning generalization. The classical theory of learning generalization posits that observers immediately and directly generalize what they have learned by inferring the shared constructs of the trained and generalization task contexts. This classical view is often referred to as immediate generalization [14, 15]. However, immediate generalization highly depends on the recognition of shared constructs between training and generalization. This means that learned experience may be limited to some specific task components. In contrast, the “learning to learn” theory emphasizes the general ability to quickly acquire task

statistics and facilitate learning in real time [7, 10, 16]. Most importantly, this “learning to learn” ability should not be specific to a particular task component and thus has the potential to produce broad generalizations across different types of tasks. This new form of generalization has recently been discovered in sequential perceptual learning [33] and has also been proposed to underlie broad generalization associated with action video game play [10, 34]. Both cross-sectional and intervention studies have identified the increased learning rate, as a hallmark of “learning to learn”, associated with action video game play in perceptual [7, 16], cognitive [16], and motor learning tasks [35].

“Learning to learn” also proposes that high-level statistical learning of task structure is the underlying mechanism for increasing learning rate. Hierarchical learning allows individuals to flexibly adjust their learning rates in response to changing environments. The environments we face are often filled with different types of uncertainty [17, 36], such as uncertainty about how a reward is obtained and uncertainty about how tasks may evolve. A lack of flexibility in responding to environmental changes is likely to be associated with psychiatric disorders, such as social anxiety disorder and major depressive disorder [37, 38]. Traditional reinforcement learning often assumes that the learning rate is a fixed property of an agent [39]. This means that an agent has the same learning rate across throughout the task, which is obviously suboptimal and inflexible [40, 41]. A better approach is to adjust the learning rate according to task statistics. For example, if the task statistics (e.g., the probabilistic mapping between action and reward) change rapidly, an agent needs to increase the learning rate to adapt quickly to the changes. However, if this task statistics are stable, individuals should decrease the learning rate to avoid overfitting to noise [36, 42, 43]. In other words, the hierarchical form of “learning to learn” allows an agent to flexibly adjust learning speed accordingly in different tasks.

#### **The underlying mechanisms associated with enhanced “learning to learn” in AVGPs**

We speculate that several unique characteristics of action video games may be the reasons.

First, the fast pace of action video games may lead to superior cognitive functions. Fast-paced games require players to switch quickly between different scenarios or tasks [10, 26]. Several studies have shown that AVGPs have greater task switching abilities [34, 44–46]. Given limited cognitive resources [47, 48], the reduced cognitive cost of task switching allows AVGPs to allocate more cognitive resources to hierarchical learning, leading to better “learning to learn”. The fast pace of action video games also requires players to simultaneously track and

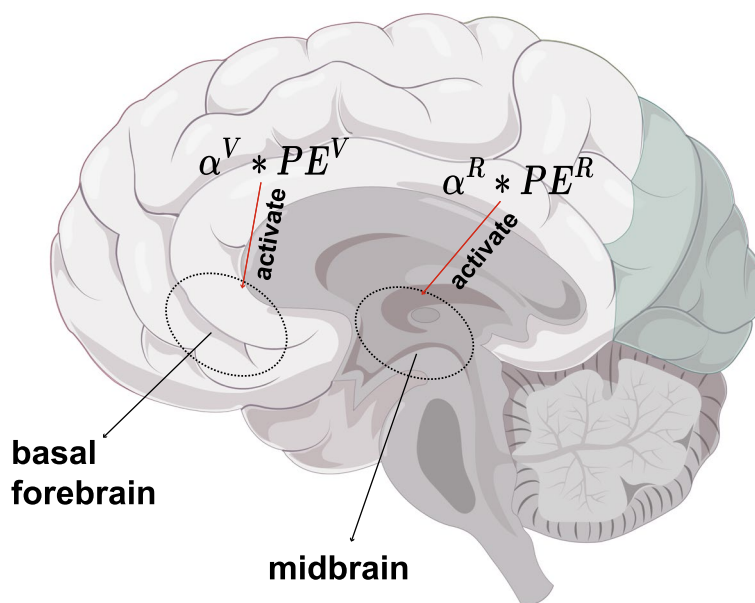
store multiple rapid processes and predict future game events in real time. For example, in a first-person shooting game (i.e., *Overwatch*), a player must quickly determine where other players have previously attacked and predict their possible current and next locations. Training to track and store information is associated with improved working memory in AVGPs [49, 50]. Improved working memory allows players to retain task statistics during sequential tasks and respond more quickly and accurately.

Second, the complex spatial environments of action video games promote perceptual sensitivity. Action video games tend to contain highly complex and realistic spatial environments, and this is associated with increased perceptual sensitivity to external sensory events [51]. Enhanced perceptual sensitivity allows AVGPs to quickly and accurately detect real-time fluctuations or changes in new tasks, thereby improving “learning to learn”.

However, this is a cross-sectional study, and we cannot exclude the possibility that the people with enhanced “learning to learn” are more attracted by action video games such that they are related. Researchers [52, 53] postulated that the capacity to make multilevel predictions and to learn from uncertainties that emerge during gameplay will facilitate the expeditious and efficacious reduction of prediction errors in game scenarios. This will enable players to “feel good” and, as a result, select and persist with such games.

#### **Neural mechanisms underlying enhanced “learning to learn”**

What are the neural mechanisms underlying enhanced “learning to learn”? Previous studies have shown that hierarchical learning exist in the human brain. Existing studies have focused on the neural mechanisms associated with different levels of learning rates and prediction errors (PEs). A study combining HGF modeling with electroencephalogram (EEG) found that beta power in the sensorimotor cortex is negatively correlated with volatility learning rate before action execution and positively correlated with association learning rate after action execution [54]. Another EEG study found that the P300 response in the frontal and central scalp regions is positively correlated with the absolute values of low-level PEs and negatively correlated with high-level PEs [43]. In other words, beta power in sensorimotor cortex and P300 responses in the frontal and central scalp may serve as neural markers of hierarchical learning. In this study, we found both increased volatile and association learning rate. Our results predict that enhanced “learning to learn” may produce a weaker and stronger beta wave in sensorimotor cortex before and after action execution. Interestingly, these predictions are consistent with two



**Fig. 6** Corresponding brain regions for learning rate-weighted prediction errors at different levels demonstrated in the previous studies [58, 59]

recent EEG studies of AVGPs. In the two EEG studies, the researchers did not find the changes in beta wave power in the frontal lobes before and after movement but found that the variation of beta-wave power is greater before and after action execution in AVGPs [55]. In addition, beta wave power has been shown to increase significantly during high-intensity action video game activities [56]. Our findings also predict a stronger P300 response in the frontal and central scalp regions associated with enhanced “learning to learn.” This prediction is consistent with a recent EEG study that identified a greater amplitude of the task-evoked P300 component in AVGPs [57].

The studies combining HGF modeling with functional magnetic resonance imaging (fMRI) have shown that low-level PEs are encoded in dopamine-related regions of the midbrain, including the ventral tegmental area (VTA) and substantia nigra (SN). These regions have been shown to regulate dopamine release [60–62]. In contrast, high-level PEs are encoded in the basal forebrain, which regulates acetylcholine release [58, 59]. These results predict stronger activities in the midbrain VTA and SN (Fig. 6). These predictions are consistent with several recent fMRI studies of AVGPs. One fMRI study found stronger activation of reward-related midbrain structures in AVGPs [63]. Another longitudinal fMRI study showed that action video games can increase functional connectivity within the basal ganglia [64]. Similarly, some fMRI studies have found elevated activity in the striatum, as part of the basal forebrain, of AVGPs [65, 66]. All of these studies suggest that enhanced “learning to learn” is likely to be

associated with stronger activation or inhibition in the midbrain and basal forebrain.

## Conclusion

In conclusion, this study employed a Hierarchical Gaussian Filter (HGF) model to test 34 AVGPs and 36 NVGPs in a volatile reversal learning task. The results of the study demonstrate that AVGPs indeed rapidly extract volatility information and utilize the estimated higher volatility to accelerate learning of cue-response associations. These findings provide strong evidence for the “learning to learn” theory of generalization in AVGPs.

## Acknowledgements

The authors thank the participants for their support to this study.

## Authors' contributions

R-Y.Z. and Y.G. conceived and designed the study. Y.G. prepared the computer program for the Behavioral task and collected the data. Y.G. and Z.F. analyzed the data. Y.G. wrote the first draft of the manuscript. R-Y.Z., Y.G., Q.Z., and Z.F. revised the manuscript.

## Funding

This work was supported by the National Natural Science Foundation of China (32100901) and Natural Science Foundation of Shanghai (21ZR1434700) to R-Y.Z.

## Availability of data and materials

The source code of Hierarchical Gaussian Filter (HGF) can be downloaded from <https://translationalneuromodeling.github.io/tapas>. The HGF task and data for each group of subjects, as well as the code used for analysis and plotting, can be downloaded from <https://osf.io/sk82r/>.

## Declarations

### Ethics approval and consent to participate

All experimental protocols were approved by the institutional review board of Shanghai Jiao Tong University (I20211151). All research was conducted in accordance with relevant guidelines and regulations. Informed written consent was obtained from all participants.

### Consent for publication

Not applicable.

### Competing interests

The authors declare no competing interests.

Received: 23 April 2024 Accepted: 14 August 2024

Published online: 30 August 2024

## References

- Bavelier D, Green CS, Pouget A, Schrater P. Brain plasticity through the life span: learning to learn and action video games. *Annu Rev Neurosci*. 2012;35:391–416.
- Green CS, Gorman T, Bavelier D. Action video-game training and its effects on perception and attentional control. *Cognitive training: An overview of features and applications*. 2016;107–116.
- Dye MW, Green CS, Bavelier D. The development of attention skills in action video game players. *Neuropsychologia*. 2009;47:1780–9.
- Blacker KJ, Curby KM, Klobusicky E, Chein JM. Effects of action video game training on visual working memory. *J Exp Psychol Hum Percept Perform*. 2014;40:1992–2004.
- McDermott AF, Bavelier D, Green CS. Memory abilities in action video game players. *Comput Hum Behav*. 2014;34:69–78.
- Gonçalves, E. d. S. & Castilho, G. M. d. Effects of action video game engagement on attention and working memory. *Psychology & Neuroscience* (2024).
- Bejjanki VR, et al. Action video game play facilitates the development of better perceptual templates. *Proc Natl Acad Sci U S A*. 2014;111:16961–6.
- Green CS, Li R, Bavelier D. Perceptual learning during action video game playing. *Top Cogn Sci*. 2010;2:202–16.
- Green CS, Pouget A, Bavelier D. Improved probabilistic inference as a general learning mechanism with action video games. *Curr Biol*. 2010;20:1573–9.
- Bavelier D, Green CS, Pouget A, Schrater P. Brain plasticity through the life span: learning to learn and action video games. *Annu Rev Neurosci*. 2012;35:391–416.
- Poggio T, Bizzi E. Generalization in vision and motor control. *Nature*. 2004;431:768–74.
- Shepard RN. Toward a universal law of generalization for psychological science. *Science*. 1987;237:1317–23.
- Schulz, E. Towards a unifying theory of generalization. UCL (University College London). (2017).
- Johnson BP, et al. Generalization of procedural motor sequence learning after a single practice trial. *NPJ Sci Learn*. 2023;8:45.
- Liu Z, Weinsall D. Mechanisms of generalization in perceptual learning. *Vision Res*. 2000;40:97–109.
- Zhang R-Y, et al. Action video game play facilitates “learning to learn.” *Communications biology*. 2021;4:1154.
- Behrens TE, Woolrich MW, Walton ME, Rushworth MF. Learning the value of information in an uncertain world. *Nat Neurosci*. 2007;10:1214–21.
- Browning M, Behrens TE, Jocham G, O’Reilly JX, Bishop SJ. Anxious individuals have difficulty learning the causal statistics of aversive environments. *Nat Neurosci*. 2015;18:590–6.
- Balleine BW, O’Doherty JP. Human and rodent homologies in action control: corticostriatal determinants of goal-directed and habitual action. *Neuropsychopharmacology*. 2010;35:48–69.
- Owen AM, McMillan KM, Laird AR, Bullmore E. N-back working memory paradigm: A meta-analysis of normative functional neuroimaging studies. *Hum Brain Mapp*. 2005;25:46–59.
- Baddeley A. Working memory: looking back and looking forward. *Nat Rev Neurosci*. 2003;4:829–39.
- Mathys C, Daunizeau J, Friston KJ, Stephan KE. A bayesian foundation for individual learning under uncertainty. *Front Hum Neurosci*. 2011;5:39.
- Glascher J, Daw N, Dayan P, O’Doherty JP. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron*. 2010;66:585–95.
- Dayan P, Daw ND. Decision theory, reinforcement learning, and the brain. *Cogn Affect Behav Neurosci*. 2008;8:429–53.
- Green CS, et al. Playing some video games but not others is related to cognitive abilities: A critique of Unsworth et al.(2015). *Psychol Sci*. 2017;28:679–82.
- Novak, E. & Tassell, J. in 2017 IEEE 17th International Conference on Advanced Learning Technologies (ICALT). 142–144 (IEEE).
- Dale G, Kattner F, Bavelier D, Green CS. Cognitive abilities of action video game and role-playing video game players: Data from a massive open online course. *Psychology of Popular Media*. 2020;9:347.
- Gagne C, Zika O, Dayan P, Bishop SJ. Impaired adaptation of learning to contingency volatility in internalizing psychopathology. *Elife*. 2020;9:e61387.
- Bejjanki VR, Sims CR, Green CS, Bavelier D. Evidence for action video game induced ‘learning to learn’ in a perceptual decision-making task. *J Vis*. 2012;12:287–287.
- Schonberg T, Daw ND, Joel D, O’Doherty JP. Reinforcement learning signals in the human striatum distinguish learners from nonlearners during reward-based decision making. *J Neurosci*. 2007;27:12860–7.
- Li J, Schiller D, Schoenbaum G, Phelps EA, Daw ND. Differential roles of human striatum and amygdala in associative learning. *Nat Neurosci*. 2011;14:1250–2.
- Hein TP, et al. Anterior cingulate and medial prefrontal cortex oscillations underlie learning alterations in trait anxiety in humans. *Commun Biol*. 2023;6:271.
- Kattner F, Cochrane A, Cox CR, Gorman TE, Green CS. Perceptual learning generalization from sequential perceptual training as a change in learning rate. *Curr Biol*. 2017;27:840–6.
- Green CS, Bavelier D. Action video game modifies visual selective attention. *Nature*. 2003;423:534–7.
- Gozli DG, Bavelier D, Pratt J. The effect of action video game playing on sensorimotor learning: Evidence from a movement tracking task. *Hum Mov Sci*. 2014;38:152–62.
- Nassar MR, et al. Age differences in learning emerge from an insufficient representation of uncertainty in older adults. *Nat Commun*. 2016;7:11609.
- Sohail A, Zhang L. Informing the treatment of social anxiety disorder with computational and neuroimaging data. *Psychoradiol*. 2024;kkae010.
- Li H, et al. Altered cortical morphology in major depression disorder patients with suicidality. *Psychoradiology*. 2021;1:13–22.
- Simoens J, Verguts T, Braem S. Learning environment-specific learning rates. *PLoS Comput Biol*. 2024;20:e1011978.
- Leimar, O., Quiñones, A. E. & Bshary, R. Flexibility of learning in complex worlds. *bioRxiv*, 2023.2006.2012.544544 (2023).
- Jiang J, Beck J, Heller K, Egner T. An insula-frontostriatal network mediates flexible cognitive control by adaptively predicting changing control demands. *Nat Commun*. 2015;6:8165.
- Jepma M, et al. Catecholaminergic Regulation of Learning Rate in a Dynamic Environment. *PLoS Comput Biol*. 2016;12:e1005171.
- Liu M, Dong W, Qin S, Verguts T, Chen Q. Electrophysiological Signatures of Hierarchical Learning. *Cereb Cortex*. 2022;32:626–39.
- Green CS, Sugarman MA, Medford K, Klobusicky E, Bavelier D. The effect of action video game experience on task-switching. *Comput Hum Behav*. 2012;28:984–94.
- Cain MS, Landau AN, Shimamura AP. Action video game experience reduces the cost of switching tasks. *Atten Percept Psychophys*. 2012;74:641–7.
- Karle JW, Watter S, Shedden JM. Task switching in video game players: Benefits of selective attention but not resistance to proactive interference. *Acta Physiol (Oxf)*. 2010;134:70–8.
- Bjorklund DF, Harnishfeger KK. The resources construct in cognitive development: Diverse sources of evidence and a theory of inefficient inhibition. *Dev Rev*. 1990;10:48–71.
- Bruder G, Lubos P, Steinicke F. Cognitive resource demands of redirected walking. *IEEE Trans Visual Comput Graphics*. 2015;21:539–44.

49. Green CS, Bavelier D. Learning, attentional control, and action video games. *Curr Biol*. 2012;22:R197–206.
50. Colzato LS, van den Wildenberg WP, Zmigrod S, Hommel B. Action video gaming and cognitive control: playing first person shooter games is associated with improvement in working memory but not action inhibition. *Psychol Res*. 2013;77:234–9.
51. Chopin A, Bediou B, Bavelier D. Altering perception: the case of action video gaming. *Curr Opin Psychol*. 2019;29:168–73.
52. Deterding S, Andersen MM, Kiverstein J, Miller M. Mastering uncertainty: A predictive processing account of enjoying uncertain success in video game play. *Front Psychol*. 2022;13: 924953.
53. Andersen MM, Kiverstein J, Miller M, Roepstorff A. Play in predictive minds: A cognitive theory of play. *Psychol Rev*. 2023;130:462.
54. Palmer CE, Auksztulewicz R, Ondobaka S, Kilner JM. Sensorimotor beta power reflects the precision-weighting afforded to sensory prediction errors. *Neuroimage*. 2019;200:59–71.
55. Salminen M, Ravaja N. Oscillatory brain responses evoked by video game events: The case of Super Monkey Ball 2. *Cyberpsychol Behav*. 2007;10:330–8.
56. McMahan T, Parberry I, Parsons TD. Modality specific assessment of video game player's experience using the Emotiv. *Entertainment Computing*. 2015;7:1–6.
57. Mishra J, Zinni M, Bavelier D, Hillyard SA. Neural basis of superior performance of action videogame players in an attention-demanding task. *J Neurosci*. 2011;31:992–8.
58. Iglesias S, et al. Cholinergic and dopaminergic effects on prediction error and uncertainty responses during sensory associative learning. *Neuroimage*. 2021;226: 117590.
59. Iglesias S, et al. Hierarchical Prediction Errors in Midbrain and Basal Forebrain during Sensory Learning. *Neuron*. 2019;101:1196–201.
60. Beier KT, et al. Circuit architecture of VTA dopamine neurons revealed by systematic input-output mapping. *Cell*. 2015;162:622–34.
61. Engelhard B, et al. Specialized coding of sensory, motor and cognitive variables in VTA dopamine neurons. *Nature*. 2019;570:509–13.
62. Wise RA. Roles for nigrostriatal—not just mesocorticolimbic—dopamine in reward and addiction. *Trends Neurosci*. 2009;32:517–24.
63. Klasen M, Weber R, Kircher TT, Mathiak KA, Mathiak K. Neural contributions to flow experience during video game playing. *Soc Cogn Affect Neurosci*. 2012;7:485–95.
64. Pujol J, et al. Video gaming in school children: How much is enough? *Ann. Neurol*. 2016;80:424–33.
65. Benady-Chorney J, et al. Action video game experience is associated with increased resting state functional connectivity in the caudate nucleus and decreased functional connectivity in the hippocampus. *Comput Hum Behav*. 2020;106: 106200.
66. Kühn S, Gleich T, Lorenz RC, Lindenberger U, Gallinat J. Playing Super Mario induces structural brain plasticity: gray matter changes resulting from training with a commercial video game. *Mol Psychiatry*. 2014;19:265–71.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.