Learning task rule updating strategies requires extensive practice

Shengjie Xu (Shengjie.Xu@Ugent.be)
Department of Experimental Psychology, Ghent University, Belgium

Tanya Wen (2tanyawen@gmail.com)
Meta Reality Labs, Redmond, U.S.A
Center for Cognitive Neuroscience, Duke University, U.S.A

Tobias Egner (tobias.egner@duke.edu)
Center for Cognitive Neuroscience, Duke University, U.S.A
Department of Psychology and Neuroscience, Duke University, U.S.A

Tom Verguts (Tom.Verguts@UGent.be)
Department of Experimental Psychology, Ghent University, Belgium

Senne Braem (Senne.Braem@UGent.be)
Department of Experimental Psychology, Ghent University, Belgium

Abstract

People can adjust how fast they update task rules, depending on the volatility of their environment. We investigated whether this adaptivity is primarily driven by recently experienced volatility in task demands, or can also be shaped by learned, environment-specific associations with expected levels of volatility. To this end, we trained participants on a Wisconsin Card Sorting Task where different environments required different speeds of task rule updating. We demonstrate that, initially, participants updated strategies depending on the most recent experienced levels of volatility and feedback (Experiment 1). However, after extensive (four days) training (Experiment 2), participants also developed environment-specific associations. Our findings provide important insights in how people learn to regulate cognitive flexibility.

Keywords: cognitive control; cognitive flexibility; context; reinforcement learning; multi-day training

Introduction

Responding to stimuli based on a certain task rule generally comes easy to humans. However, we are often faced with environments with changing task rules and the speed at which task rules change may differ, so we need to learn not only when but also how quickly we should decide to change tasks. In a more volatile environment, task rules change more frequently, requiring a higher learning rate to update task rules. In contrast, in a less volatile environment, fast task rule updating may be suboptimal, and it can be better to await more feedback before changing task rules, favoring a lower learning rate. The study of updating task rules in response to environment-specific needs has been the topic of research on cognitive control (Badre, 2025; Egner, 2023; Monsell, 2003; Musslick & Cohen, 2021). However, it remains unclear how much of these behavioral adaptations reflect adapting to recently experienced volatility, or may also rely on knowing about environment-specific levels of task volatility (e.g., Braem et al., 2024).

Contemporary theories on cognitive control emphasize that the regulation of cognitive control requires learning (Abrahamse et al., 2016; Braem & Egner, 2018; Egner, 2014; 2023; Verguts & Notebaert, 2009), specifically meta-learning (Griffiths et al., 2019; Wang, 2021). Here, humans can learn how to set up control parameters, sometimes referred to as meta-control (Eppinger et al., 2021), by associating environment-specific features to different task control processes (Braem & Egner, 2018; Chiu & Egner, 2019; Xu et al., 2024). In this way, associated control settings can be evoked when humans later revisit these same environments, allowing for faster, adaptive changes in task updating.

One recent study found that people indeed show such environment-specific regulations of control in a task switching paradigm, but only after four days of training (Xu et al., 2024). This suggests that extensive task experience might be needed before people both learn and rely on such environment-specific strategies (Braem et al., 2024). However, in traditional task switching paradigms like the one by Xu et al. (2024), participants are always cued about which task to perform next, leaving little uncertainty about the current need for task rule updating, which could demotivate people to learn environment-specific statistics.

Instead, more uncertain environments with noisy feedback might be more encouraging for people to both learn and rely on environment-specific task updating strategies. Contrasting a high versus low volatility environment in a Wisconsin Card Sorting Task with probabilistic feedback, Wen and colleagues (2023) recently tested whether people can both learn and transfer the learning rate to a subsequent testing phase with medium levels of volatility, in the same or even different versions of the task. They found that participants indeed showed different learning rates for task updating based on high versus low environmental volatility, as also observed in the testing phases. However, this finding may reflect a carry-over effect of recently experienced volatility, rather than a learning and memorizing of environment-specific control settings. To test this, a within-subject design is required to investigate if humans can strategically regulate cognitive flexibility in response to different environments.

Here, we developed a customized Wisconsin Card Sorting

In D. Barner, N.R. Bramley, A. Ruggeri and C.M. Walker (Eds.), *Proceedings of the 47th Annual Conference of the Cognitive Science Society*

Task, closely inspired by the design by Wen and colleagues (2023), where participants experienced both high and low volatilities in two different environments, and then were examined in a testing phase. We conducted two experiments: one-day training in Experiment 1 and four-day training in Experiment 2. Similarly, we fitted choice behavior using a standard Rescorla–Wagner model and a dual-rates model, which estimates learning rates for positive and negative feedback separately. We hypothesized that learned environment-specific task rule updating strategies can only be observed after extensive training, typically not achievable within a single experimental session. Concretely, we expected participants to rely more on locally experienced differences in demands for task rule changes in a first session, without relying on environment-specific cues. However, through learning over successive days, participants can gradually learn (and start relying on) environment-specific task rule updating strategies.

Methods

Participants

Participants were students from Ghent University who received course credits for their participation in this study. We recruited 65 participants in Experiment 1, and 62 participants in Experiment 2. Participants with accuracy below 65% were excluded from both experiments. In Experiment 1, we excluded nine participants, leaving a final sample size of 55 (43 females, M_{age}= 18.87 years, SD_{age}= 1.29, range = 17-23) for the model fitting and statistical analyses. In Experiment 2, we excluded seven participants, resulting in a final sample size of 55 (49 females, M_{age}= 18.40 years, SD_{age}= 1.57, range = 17-28). Both experiments were approved by the Ethical Committee of the Faculty of Psychology and Educational Sciences of Ghent University.

Wisconsin Card Sorting Task

On each trial, participants needed to choose between two choice cards to match the reference card on top (Figure 1A). Each choice card only shared one feature with the top reference card, e.g., blue color on the left choice card and item number three on the right card. There were four possible feature dimensions on these cards: color (blue, green, red or purple), filling (checkered, dots, wave, or grid), item number (one, two, three or four) and item shape (circle, triangle, plus, or star), resulting in 256 possible combinations. For each participant, only two feature dimensions were randomly selected as relevant for the task, counterbalanced across participants. On each trial, only one feature dimension was relevant to the correct card matching rule. Participants were asked to make as many correct choices as possible throughout the experiment.

Procedure

Each trial began with a 1000 ms fixation period, after which the cards were presented for 2000 ms, or until a choice was made. Next, reward feedback was presented for 500 ms. Reward feedback was probabilistic: On 80% of trials the feedback was correct, but on 20% of trials, participants received positive feedback for incorrect choices or negative feedback for correct choices. The probabilistic feedback was provided to prevent participants from relying on trial counting to anticipate task switches. When participants failed to respond in 2000 ms, they would receive a "Too slow" message instead.

We studied task choice behavior in response to the different (associated) volatility contexts in each block. Volatility was defined as the frequency of rule changes, where low indicated that the matching rule changed every 30 trials, and high indicated that the matching rule changed every 10 trials. Each volatility context was associated to two different table backgrounds: wood table or stone table (Figure 1B). In the training phase (Figure 1C), each table picture was always presented in either the low or high volatility context, indicating to participants which volatility context they were in. Table-volatility associations were counterbalanced among participants and were not instructed, so participants needed to learn them over trials. To test whether participants successfully learned and used these environment-volatility associations to influence task rule updating, we inserted probe blocks (Figure 1C) where the different tables were again presented in the background, but the matching rule changed every 20 trials.

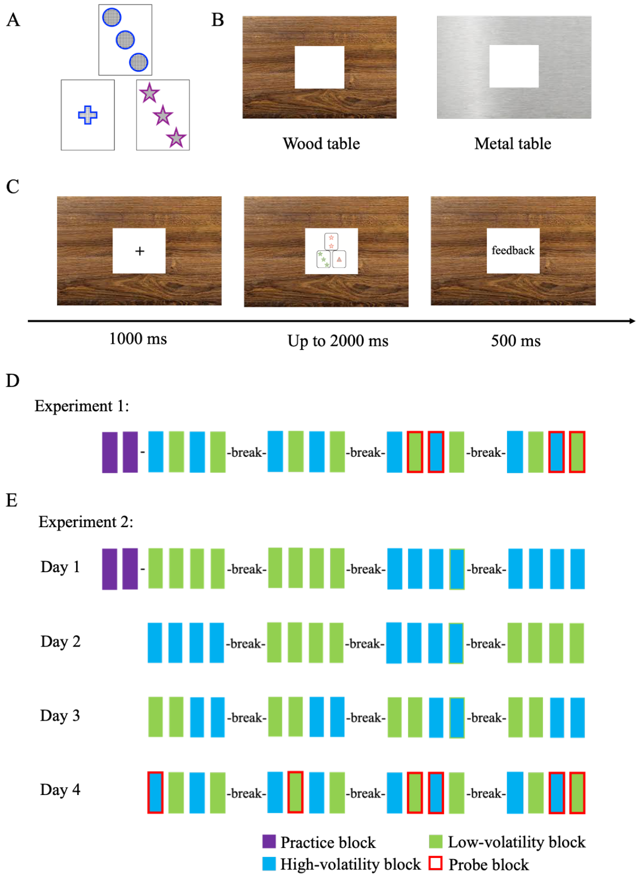

Figure 1: Wisconsin card sorting task paradigm. (A) Example

cards presented on each trial. (B) Table pictures. (C) Trial overview. (D) Overview of Experiment 1. (E) Overview of Experiment 2.

Experiment 1 In Experiment 1 (Figure 1D), participants began with two practice blocks, where the actual matching rule was presented on each trial to help them get familiar with the task. Next, participants entered the learning phase, where they experienced either high or low volatility block by block, in an interleaved manner. Each block consisted of 60 trials. In the second half of the experiment, probe blocks were inserted. There was at least one learning block between two probe blocks with the same level of the volatility. Participants randomly started with a high or low volatility block.

Experiment 2 In Experiment 2 (Figure 1E), participants also started with practice blocks. Differently, the learning blocks gradually changed from a chunked order to an interleaved order over days (see also, Xu et al., 2024), as humans are better at learning context-specific knowledge in a blocked manner compared to an interleaved manner (Flesch et al., 2018). The probe blocks only appeared on the last day: one probe block of each context in the first half and two probe blocks of each context in the second half. The first volatility context was randomly chosen and shifted across learning days. Participants were asked to start the experiment each day approximately at the time they started on the first day, allowing for a variation of up to three hours (24 ± 3 hours).

Reinforcement Learning Models

We fitted participants' choice behaviors to two reinforcement learning models: a standard Rescorla-Wagner (RW) model and a dual-rates (DR) model, where we fitted individual learning rates for positive and negative feedback separately. We constructed each model using a hierarchical framework, implementing mixed-effects linear models to account for both fixed and random effects within the data structure for parameter estimates.

At the top level, the conditional learning rate α and inverse temperature β in each experimental condition were derived from their group-level average:

$$\alpha_C[p, v] = \alpha_G + \alpha_R[p, v]$$

$$\beta_C[p, v] = \beta_G + \beta_R[p, v]$$

Here, p stands for the factor of experiment phase (learning versus probe), and v stands for the factor of volatility level (high versus low). $\alpha_C[p, v]$ and $\beta_C[p, v]$ represent the learning rates and inverse temperature parameters in each condition, which are generated by a linear combination of group-level averages across all conditions and participants, α_G and β_G, and conditional random effects, $\alpha_R[p, v]$ and $\beta_R[p, v]$.

Participants' individual parameters were modelled at the next hierarchical level:

$$\alpha[i, p, v] = \alpha_C[p, v] + \alpha'_R[i, p, v]$$

$$\beta[i, p, v] = \beta_C[p, v] + \beta'_R[i, p, v]$$

Similarly, individual parameters, $\alpha[i, p, v]$ and $\beta[i, p, v]$ in each condition are products of a linear combination of their conditional values, $\alpha_C[p, v]$ and $\beta_C[p, v]$ and random effects, $\alpha'_R[i, p, v]$ and $\beta'_R[i, p, v]$, at the individual level where i represents each participant.

At the lowest level of the hierarchy, the probability of choosing a card was computed by a SoftMax function, and the values of choosing each card were updated after feedback:

$$P_{i,c,t} = \frac{e^{\beta[i,p,v]*Q_{i,c,t}}}{\sum_{n=1}^{2} e^{\beta[i,p,v]*Q_{i,n,t}}}$$

$$Q_{i,c,t+1} = Q_{i,c,t} + \alpha[i, p, v] * (R_t - Q_{i,c,t})$$

Namely, the Q values of the two choice cards were sent to a SoftMax function together with the conditional inverse temperature parameter to compute the probability of choosing a certain card on trial t. Next, the Q value of the chosen card was updated by the prediction error, which is the difference between the received reward feedback R_t and the Q value of this card $Q_{i,c,t}$ on trial t. This update was weighted by the individual conditional learning rate $\alpha[p, v]$.

The DR model was constructed in the same hierarchical way, but the Q value update was implemented separately according to the received reward feedback:

$$\text{If } R_t = 1: \quad Q_{i,c,t+1} = Q_{i,c,t} + \alpha^{+}[i, p, v] * (R_t - Q_{i,c,t})$$

$$\text{If } R_t = 0: \quad Q_{i,c,t+1} = Q_{i,c,t} + \alpha^{-}[i, p, v] * (R_t - Q_{i,c,t})$$

where $\alpha^{+}[p, v]$ and $\alpha^{-}[p, v]$ represent learning rates for positive reward feedback (positive learning rate), $R_t = 1$ or negative reward feedback (negative learning rate), $R_t = 0$, respectively.

The same models were also used in Experiment 2. To account for the random effect of experiment days, we used another factor *day* in the hierarchical structure.

Model Fitting and Model Comparison

Parameters in the two hierarchical reinforcement models were estimated using Hamiltonian Monte Carlo sampling, implemented in Stan. The learning rates at the top hierarchy were sampled from a normal distribution $\mathcal{N}(0, 1)$, while inverse temperature parameters were sampled from a normal distribution $\mathcal{N}(5, 5)$. This prior selection was determined according to a previous study with a similar design (Wen et al., 2023). We also applied non-centered parameterization to reduce parameter dependence across hierarchies for a more reliable estimation. Hierarchical Bayesian modeling estimates the posterior distributions, not point estimates, of parameters in each condition. Therefore, we computed posterior probabilities to compare parameters in each condition. The model comparison was conducted by using the leave-one-out information criterion (LOOIC) method, with a higher expected log pointwise predictive density (ELPD) value indicating better fitting performance.

Results

Switch Frequency

We first analyzed the switch frequency in both experiments to see whether our volatility settings evoked different task updating strategies using a two (volatility: high versus low) by two (phase: learning blocks versus probe blocks) repeated measure ANOVA (rmANOVA). We applied a Greenhouse-Geisser sphericity correction to within-subject factors that violated the sphericity assumption and reported corrected statics, and the same procedure was taken for all following rmANOVAs.

Switch frequency is defined as the proportion of trials in which participants chose to switch the matching rule within each block. In Experiment 1, we observed a significant main effect of volatility, $F(1, 54) = 39.99$, $p < .001$, $\eta_p{}^2 = 0.43$, indicating that participants switched more frequently in the high-volatility context compared to the low-volatility context (Fig. 2A, 95% confidence interval (*CI*) = [0.87, 5.03]). Moreover, the switch frequency was also higher in the learning blocks, $F(1,54) = 17.65$, $p < .001$, $\eta_p{}^2 = 0.25$, 95% *CI* = [0.16, 4.35]. However, the significant interaction of volatility and phase, $F(1, 54) = 79.62$, $p < .001$, $\eta_p{}^2 = 0.60$, indicated that although participants showed a higher switch frequency in learning blocks, $t(54) = 11.32$, $p < .001$, Cohen's $d = 1.53$, 95% *CI* = [3.95, 9.37], there was no difference between the two volatility contexts in the probe blocks, $t(54) = -1.16$, $p = .249$, Cohen's $d = 0.16$, 95% *CI* = [-3.76, 2.22]. Similarly, after four days' training in Experiment 2, participants robustly showed a higher switch frequency in the high volatility context (Fig. 2B), indicated by a main effect of volatility, $F(1, 54) = 38.94$, $p < .001$, $\eta_p{}^2 = 0.42$, 95% CI = [1.00, 4.41]. We also observed a higher switch frequency in the learning blocks, $F(1, 54) = 16.44$, $p < .001$, $\eta_p{}^2 = 0.23$, 95% *CI* = [-0.33, 3.14]. There was again a significant interaction between volatility and phase in Experiment 2, $F(1, 54) = 57.00$, $p < .001$, $\eta_p{}^2 = 0.51$. The post-hoc test results revealed that, while participants kept presenting a higher switch frequency in the learning blocks, $t(54) = 12.30$, $p < .001$, Cohen's $d = 1.65$, 95% *CI* = [3.66, 7.98], this effect was again missing in the probe blocks, $t(54) = -0.59$, $p = .557$, Cohen's $d = 0.08$, 95% *CI* = [-2.94, 2.11]. Taken together, we found that participants switched between task rules more frequently in the high volatility context, but this effect did not extend to the probe blocks, even after extensive training.

We further checked how switch frequency changed over training (Fig. 3C), using a two (volatility: high versus low) by three (training days 1-3) rmANOVA. First, we observed a significant main effect of training days, $F(1.70, 91.88) = 20.85$, $p < .001$, $\eta_p{}^2 = 0.28$. Post-hoc analyses revealed that, participants switched match rules less frequently after day 1, $t(54) = 5.12$, $p < .001$, Cohen's $d = 0.69$, 95% *CI* = [-5.47, -1.39], and their switching frequency remained stable from Day 2 onward, $t(54) = 0.46$, $p = .648$, Cohen's $d = 0.06$, 95% *CI* = [-2.27, 1.82]. In addition, across three training days, participants consistently switched their matching rules more frequently in a high-volatility environment, $F(1, 54) = 262.78$, $p < .001$, $\eta_p{}^2 = 0.83$, 95% *CI* = [4.85, 8.54]. However, the difference in switch frequency between two volatility environments was stable over the course of learning, indicated by an insignificant interaction of training days and volatility, $F(1.55, 83.78) = 0.23$, $p = .740$, $\eta_p{}^2 < 0.01$.

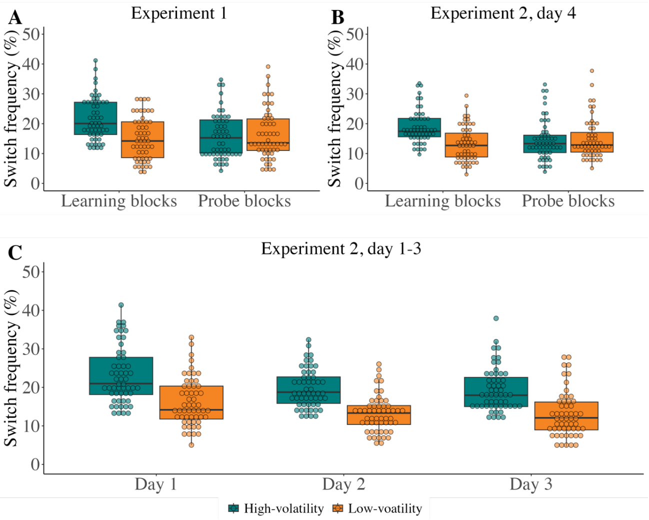

Figure 2: Switch frequency in Experiment 1 (A), Experiment 2, day 4 (B) and Experiment 2 day 1-3 (C). Individual dots represent each participant's values.2

Model Comparison

We computed relative ELPD values (0 for the best model) and the standard error (SE) of the ELPD difference between a model and the best model. The DR model outperformed the RW model in both experiments (Table 1). Therefore, our remaining analyses focused on the DR model parameters.

Table 1: Model comparison results

	Model type	Δ ELPD	Δ SE
Exp 1	DR model	0	0
	RW model	-84.0	12.8
Exp 2	DR model	0	0
	RW model	-1509.9	49.6

Experiment 1

Figure 3 illustrates parameter estimates in Experiment 1 from the DR model.

Learning Phase We first examined whether our volatility settings evoked different learning rates. We observed that positive learning rates were significantly higher in the high-volatility context in the learning phase (Figure 3A), posterior probability (*p*_post) = 0.019, but not the negative learning rates (Figure 3B), *p*_post = 0.908. The inverse temperature parameters (Figure 3C) were also lower in the high-volatility context, *p*_post = 0.004, indicating behavior generation was more random.

Probe Phase Critical to our hypothesis, we also wanted to investigate whether the observed parameter patterns in the learning phase would extend to the probe phase. However, there was no difference between the two volatility contexts in positive learning rates, *p*_post = 0.176, negative learning rates, *p*_post = 0.691, or inverse temperature parameters, *p*_post = 0.342. These null results suggest that participants did not seem to use a learned environment-specific strategy but rather showed a more direct adjustment to the local needs of the probe phase.

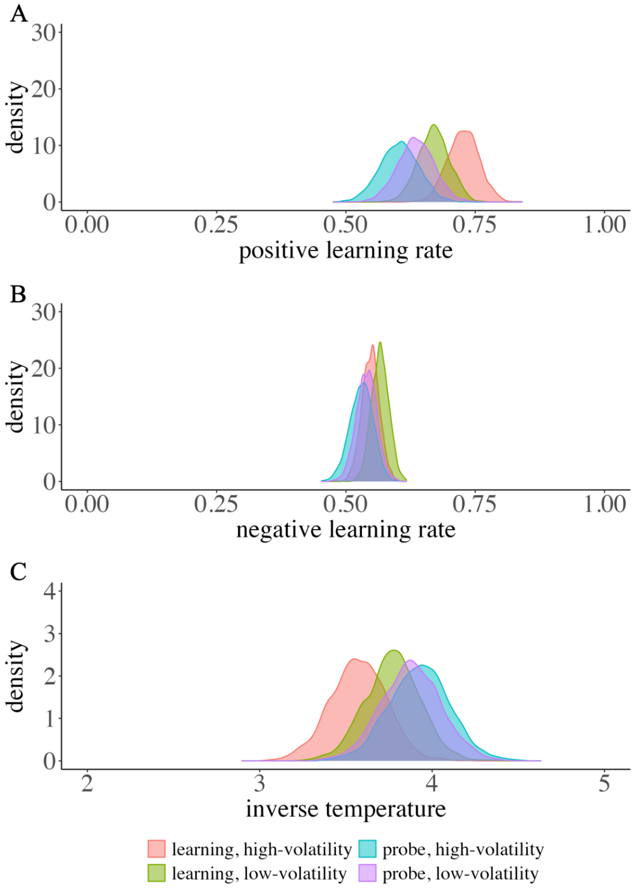

Figure 3: DR model's parameter estimates in Experiment 1.

Experiment 2

Figure 4 illustrates parameter estimates in Experiment 2 from the DR model over learning days. In Experiment 2, our key hypothesis was that participants would show context-specific learning rates in the probe phase, evoked by the context picture, after four days' training. Therefore, we first analyzed parameters from the DR model on the last day.

Learning Phase, Day 4 After four days' training, participants consistently applied higher positive learning rates in response to high volatility (Figure 4A), *p*_post < 0.001. However, there was no difference between both volatility contexts in negative learning rates (Figure 4B), *p*_post = 0.261. Similarly, the inverse temperature parameters were again lower in the high-volatility context (Figure 4C), *p*_post < 0.001.

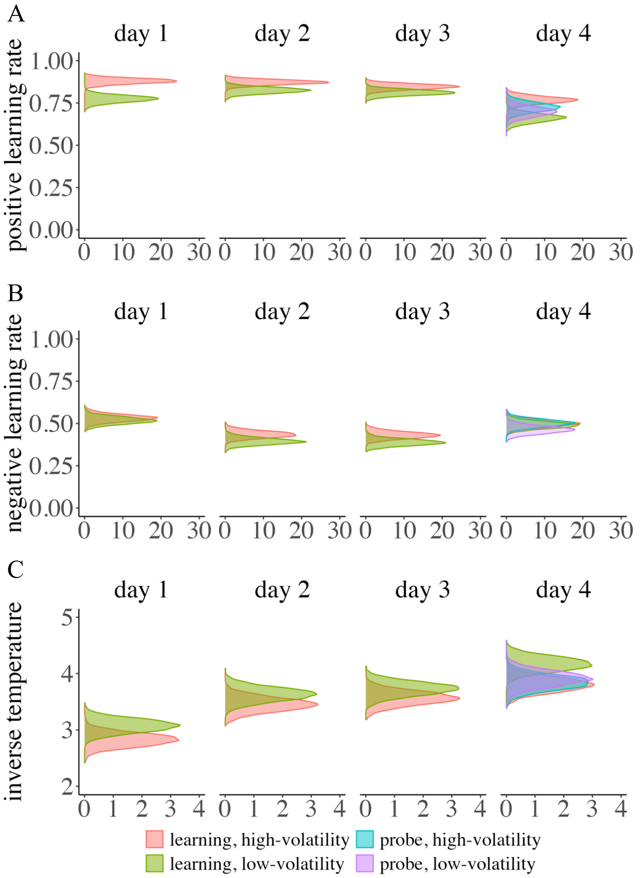

Figure 4: DR model's parameter estimates in Experiment 2.

Probe Phase, Day 4 Our main research question was whether we would observe a parameter difference between both contexts in the probe phase, after the more extensive training in Experiment 2. There was no difference in positive learning rates, *p*_post = 0.199. However, negative learning rates were higher in the high volatility context, *p*_post = 0.011. The inverse temperature parameters showed no difference between two contexts, *p*_post = 0.755. These results provide important empirical evidence of the necessity of extensive training in establishing environment-specific strategy to regulate cognitive flexibility.

Parameter Evolution over Learning To study how participants improved their task knowledge and developed a more environment-specific control strategy, we also analyzed the parameters from the first three learning days. We first checked the overall trend of parameter estimates across the learning days, by comparing parameter estimates between two successive learning days. The results revealed that there was no difference in positive learning rates between Day 1 and Day 2, *p*_post = 0.663 or Day 2 and Day 3, *p*_post = 0.133. However, the negative learning rates decreased on Day 2, *p*_post < 0.001 and became stable as of Day 3 *p*_post = 0.343. Similarly, the inverse temperature parameters increased on Day 2, *p*_post < 0.001. Although the inverse temperature parameters were larger on Day 3, this difference did not reach significance, *p*_post = 0.083.

Next, to compare the parameter difference between the two volatility contexts over learning, we computed a distribution-free overlapping measure, $\hat{\eta}$, to quantify the similarity of posterior distributions in each volatility context, with a lower value meaning less overlap between the two posterior distributions (Pastore & Calcagnì, 2019).

We first observed increasing $\hat{\eta}$ values of positive learning rates over days (Table 2), indicating a reducing difference. As for negative learning rates, they did not show a difference on Day 1, *p*_post = 0.182. However, from Day 2 on, there was a larger negative learning rate in the high-volatility context, *p*_post < 0.001, and such posterior difference further increased on Day 3, *p*_post < 0.001. In contrast to the positive learning rates, differences between negative learning rates increased over days. Last, the inverse temperature parameters were consistently larger in the low-volatility context, with all *p*_post > 0.989. and the similarity between their distributions also increased over learning days.

Table 2: Posterior distribution similarity over learning

	day 1	day 2	day 3
positive learning rate, $\hat{\eta}$	0.007	0.154	0.304
negative learning rate, $\hat{\eta}$	0.736	0.321	0.288
inverse temperature, $\hat{\eta}$	0.291	0.471	0.506

Discussion

The current study aimed to examine the environment-specific regulation of task rule updating strategies using a probabilistic Wisconsin Card Sorting Task. Across two experiments, we observed that participants indeed adopted different task rule updating strategies in the learning phase, indicated by our switch frequency results. They switched more frequently in the high volatility environment compared to in a low volatility environment. Importantly, they were also able to learn associations between these different task rule updating strategies and co-occurring environmental features, i.e., table pictures, which enabled them to use environmental features to regulate task rule updating strategies. Participants seemed to gradually evolve from a more local experience-based regulation strategy to a more learned knowledge-based regulation strategy across a multi-day training schema.

Interestingly, we observed larger positive learning rates compared to negative learning rates, as well as a differential evolution of positive versus negative learning rates over days. Over the course of learning, the difference in positive learning rates between the two volatility contexts diminished, whereas the difference in negative learning rates increased. These findings suggest that participants might be generally more sensitive to positive feedback (rewards) at first, when needing to optimize their behavior in response to different volatilities. However, they gradually switched their attention to negative feedback to adjust task rule updating strategies. Indeed, negative feedback is a critical signal in our task, indicating where a shift in task rules is needed in high volatile environments, while more often reflecting the noisiness of our feedback in low volatile environments. Previous studies also revealed that using different negative learning rates is more adaptative when people experience different levels of volatility (Simoens et al., 2024), or encounter new task rules or stimuli (Wen et al., 2023). This might help explain why, with more experience on the task, participants developed their targeted strategy mostly on negative feedback.

More broadly, our findings indicate that although people show immediate adaptivity in learning rates in response to different levels of volatility, they still fine-tuned these learning parameters to further optimize their behaviors over days. This fine-tuning process requires exploring environment structures to find the best solutions, and therefore is time-consuming (Botvinick et al., 2019). In a similar vein, cognitive control has been argued to be implemented hierarchically, analogous to a decision tree process from higher-level abstract features to lower-level concrete features (Badre, 2008). We believe that to fully explore and learn these hierarchical structures for adaptive control, humans probably start learning from more concrete lower-level features first, such as the arbitrary stimulus-response mappings, followed by learning more abstract higher-level features, like environment-control associations.

Taken together, our study investigated behavioral mechanisms underlying the environment-specific regulation of task updating strategies. We provide empirical evidence to support a critical contribution of multi-day training schemas to form environment-control associations. Our results provide new insights into the gradual learning of control, highlighting how people eventually learn to adaptively regulate task rule updating strategies with the help of environmental features.

References

Abrahamse, E., Braem, S., Notebaert, W., & Verguts, T. (2016). Grounding cognitive control in associative learning. Psychological Bulletin, 142(7), 693–728.

Badre, D. (2008). Cognitive control, hierarchy, and the rostro–caudal organization of the frontal lobes. Trends in Cognitive Sciences, 12(5), 193–200.

Badre, D. (2025). Cognitive Control. Annual Review of Psychology, 76(Volume 76, 2025), 167–195.

Botvinick, M., Ritter, S., Wang, J. X., Kurth-Nelson, Z., Blundell, C., & Hassabis, D. (2019). Reinforcement Learning, Fast and Slow. Trends in Cognitive Sciences, 23(5), 408–422.

Braem, S., Chai, M., Held, L. K., & Xu, S. (2024). One cannot simply 'be flexible': Regulating control parameters requires learning. Current Opinion in Behavioral Sciences, 55, 101347.

Braem, S., & Egner, T. (2018). Getting a Grip on Cognitive Flexibility. Current Directions in Psychological Science, 27(6), 470–476.

Egner, T. (2023). Principles of cognitive control over task focus and task switching. Nature Reviews Psychology, 1–13.

Flesch, T., Balaguer, J., Dekker, R., Nili, H., & Summerfield, C. (2018). Comparing continual task learning in minds and machines. Proceedings of the National Academy of Sciences, 115(44), E10313–E10322.

Monsell, S. (2003). Task switching. Trends in Cognitive Sciences, 7(3), 134–140.

Musslick, S., & Cohen, J. D. (2021). Rationalizing constraints on the capacity for cognitive control. Trends in Cognitive Sciences, 25(9), 757–775.

Pastore, M., & Calcagnì, A. (2019). Measuring Distribution Similarities Between Samples: A Distribution-Free Overlapping Index. Frontiers in Psychology, 10.

Simoens, J., Verguts, T., & Braem, S. (2024). Learning environment-specific learning rates. PLOS Computational Biology, 20(3), e1011978.

Verguts, T., & Notebaert, W. (2009). Adaptation by binding: A learning account of cognitive control. Trends in Cognitive Sciences, 13(6), 252–257.

Wen, T., Geddert, R. M., Madlon-Kay, S., & Egner, T. (2023). Transfer of learned cognitive flexibility to novel stimuli and task sets. Psychological Science, 34(4), 435-454.

Xu, S., Simoens, J., Verguts, T., & Braem, S. (2024). Learning where to be flexible: Using environmental cues to regulate cognitive control. Journal of Experimental Psychology: General, 153(2), 328–338.

Xu, S., Verguts, T., & Braem, S. (2024). Cognitive flexibility versus stability via activation-based and weight-based adaptations. OSF.