Pupil-linked arousal modulates trial-by-trial motor adaptation in humans

Atsushi Yokoi[1,2,*], Jeffrey Weiler[3,4,5,6]

1. Center for Information and Neural Networks, National Institute for Information and Communications Technology, Suita, Japan
2. Graduate School of Frontier Biosciences, Osaka University, Suita, Japan
3. Schulich School of Medicine and Dentistry, Western University, London Ontario, Canada
4. The Gray Centre for Mobility and Activity, Parkwood Institute, London, Ontario, Canada.
5. Brain and Mind Institute, Western University, London, Ontario, Canada
6. Department of Physiology and Pharmacology, Western University, London, Ontario, Canada

∗ Correspondence: Atsushi Yokoi (ayokoi@nict.go.jp)

Highlights

- We assessed the dynamics of pupil-linked arousal system during short-term force-field adaptation paradigm in humans.
- Pupil dilated in both online and offline manner to movement error induced by force perturbation.
- The pupil responses showed a habituation-like reduction in the sensitivity to error size when faced with multiple reversals in perturbation.
- Arousal manipulation modulated single-trial motor learning in a baseline-dependent manner.

Acknowledgements

We thank S. Kitazawa, J. A. Pruszynski and J. Diedrichsen for granting the use of their facilities and comments on the manuscript. We also thank T. Ikegami, N. Hagura and M. Hirashima for helpful discussions, and S. Kato for assistance in data collection. The present work is supported by JSPS Fellowships for Young Scientists (#15J03233) and KAKENHI (#18K17916) to A.Y. J.W. was supported by post-doctoral fellowships from NSERC and the BrainsCAN program at Western University.

Conflict of Interests

The authors declare no competing financial interests.

Author contributions

Atsushi Yokoi: Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Resources, Software, Supervision, Validation, Visualization, Writing – Original Draft, Writing – Review & Editing

Jeffrey Weiler: Investigation, Writing – Review & Editing

Abstract

Research in reward-based decision-making showed that humans and animals dynamically modulate learning rate according to their belief about environmental change (volatility) and surprise about observation. Recent evidence also suggests that neuromodulator noradrenaline (NA) signals volatility and surprise. Despite the rich anatomical evidence suggesting the potential influence of NA on the motor system, it is still elusive how NA and volatility/surprise affect human motor learning. To address this issue, we ran a series of experiments in which we simultaneously tracked the pupil diameter, a non-invasive proxy for the central NA/arousal activity, during a short-term force-field reach adaptation paradigm. A sudden increase in error due to the force-field resulted in increased pupil dilation during movement followed by an elevated baseline diameter in the following trials. These online and offline pupil responses showed a consistent pattern with surprise and volatility simulated by a recent computational model which dynamically adjusts learning rate according to volatility estimated from experienced error (surprise). However, unlike the model's prediction, when participants experienced frequent reversals in force-field, the size of pupil responses rapidly diminished regardless of large errors induced by reversals. We further confirmed that the causal manipulation of participants' arousal by task-irrelevant auditory stimuli modulated the single-trial motor learning rate. Collectively, these results provide a compelling evidence that NA/arousal system acts as a common modulator of learning rate in both cognitive and motor domains. Rapid reduction in pupil responses at reversals suggests that error sensitivity for computing current environmental uncertainty and surprise is also highly dynamic.

Keywords

Reaching, motor learning, force field, pupil dilation, arousal

Introduction

Learning from an error to adjust one's action is an essential part of daily life. In the uncertain and dynamically changing world, it is likewise essential to know to what degree one should learn from a given error. For instance, a batter should be insensitive to an error caused by random fluctuation in wind but quickly adjust their swing based on an error that was caused by the pitcher switching from throwing a fastball instead of a changeup. In this sense, an optimal learner must take the multiple sources of uncertainty (e.g., random noise or true change in the environment) into account to dynamically adjust learning rate. It is still elusive how the central nervous system represents such multiple sources of uncertainty and adjusts learning rate based on such information.

One possible way to dissect uncertainty is to divide it into *expected* and *unexpected* source. Or said another way, uncertainty can be systemically attributed to random noise in a given environment (i.e., expected uncertainty) or to true environmental change (i.e., unexpected uncertainty) (Yu and Dayan, 2005; Soltani and Izquierdo, 2019). Under a dynamically changing environment, normative approaches postulate the learning algorithms that modulate the learning rate proportional to the unexpected uncertainty (Yu and Dayan, 2005; Behrens et al., 2007; Nassar et al., 2010; Mathys et al., 2011; Piray and Daw, 2020a). The plausibility of such schemes in human learning has been extensively tested mainly in cognitive learning tasks, such as reward-based decision-making paradigm, where the *volatility* (rate of environmental change) often represents the unexpected uncertainty (Behrens et al., 2007; Soltani and Izquierdo, 2019). Recent evidence suggests a tight link between subjective volatility (and thus unexpected uncertainty), learning rate, and the activity of central noradrenergic system (Yu and Dayan, 2005; Nassar et al., 2012; de Berker et al., 2016; Jepma et al., 2016).

Anatomical evidence shows the widespread projections from the locus coeruleus (LC), the central noradrenaline (NA) generating nucleus in the brainstem, to numerous brain areas and suggests that NA plays a critical role in various cognitive and memory functions (Doya, 2002; Galeotti et al., 2004; Sara, 2009; Szabadi, 2013). It has been suggested that one potential function of NA is to signal unexpected uncertainty and modulate learning rate (Yu and Dayan, 2005). Interestingly, the projection from LC also includes motor cortex and cerebellum (Doya, 2002; Sara, 2009; Szabadi, 2013), structures critical for error-based motor learning (Kitazawa et al., 1998; Imamizu et al., 2000; Diedrichsen et al., 2005; Inoue et al.,

2016). Despite the fact, little is known how NA (and the volatility/unexpected uncertainty) affect motor learning.

To address this issue, we ran a series of experiments in which we simultaneously tracked the pupil diameter, a non-invasive proxy for the central noradrenergic/arousal activity (Rajkowski et al., 1993; Aston-Jones and Cohen, 2005; Joshi et al., 2016; Joshi and Gold, 2020) during a short-term force-field reach adaptation paradigm (Lackner and Dizio, 1994; Shadmehr and Mussa-ivaldi, 1994). Our aim was to characterize (1) how pupil reacts and what sort of information is reflected in pupil size during this typically-used motor learning paradigm and (2) to test the possible link between NA/arousal system and motor learning.

We show that the pupil responded to force-field perturbations in both an online and offline manner; a sudden increase in error due to the introduction of force-field resulted in increase in pupil dilation during movement which was followed by an elevated baseline diameter in the following trials. These online and offline pupil responses gradually decreased as the participants adapted to the force-field. Furthermore, when there were frequent reversals in force-field direction, similar situation with the reversal learning in reward-based decision-making, the peak size of these responses gradually declined, regardless of the same large error experienced at every reversal. Finally, we also experimentally manipulated participants' arousal state and confirmed the modulation of motor learning rate which was baseline-dependent. These results provide a compelling evidence that NA/arousal system acts as a common modulator of learning rate in both cognitive and motor domains, highlighting the utility of pupillometry as a unique window into the motor system. Rapid habituation in pupil responses at reversals suggests that error sensitivity for computing current environmental uncertainty and surprise is also highly dynamic.

Results

In Experiment 1 we assessed the characteristics of the pupil response patterns during a simple Null-Perturbation-Null force-field protocol. In Experiment 2 we investigated how participants' pupil reacts to more complex perturbation schedules, which included unexpected reversals of force direction (Exps. 2A, 2B, and 2C). In Experiment 3 we tested whether and how pupil-linked arousal affects trial-by-trial learning rate by causally manipulating arousal and assessing its influence on single trial motor learning.

Pupillometry during motor adaptation paradigm.

To get the first insight about how pupil changes during a typical motor adaptation paradigm, we monitored 28 participants' eyes while they performed a reaching task with a force-field perturbation (Exp 1; Fig. 1A). To minimize measurement noise in pupillometry, we instructed the participants to keep their eyes fixated on the target and to refrain from blinking while they reached. The visual cursor feedback was occluded during reach. To minimize brightness-induced changes in pupil size, visual stimuli were iso-luminant with respect to the background (Fig. 1A). The participants showed a typical behavioral signature of force-field adaptation. A sudden introduction of the perturbation force disturbed the smooth, near straight hand trajectory (cycle 11, Fig. 1C) resulting in a large lateral deviation (cycle 14, Fig. 1C). The lateral deviations rapidly decreased with repeated reaches made under the perturbation (cycle 32, Fig. 1C; Fig. 1D), and the sudden removal of the force perturbation resulted in a large hand deviation towards the opposite direction (cycle 33, Fig. 1C; Fig. 1D), a signature of motor adaptation. With further trials the trajectory (and movement errors) returned to near-baseline level (cycle 52, Fig. 1C; Fig. 1D). The learning quantified in the force channel trials also showed the typical learning curve (Fig. 1E).

The participants' pupil dilated during movement (Fig. 1F). Analysis of the velocity of the pupil dilation revealed that when the perturbation was unexpectedly applied, the pupil exhibited additional dilation starting ~300 ms after movement onset and lasting until ~700 ms after movement onset (Fig. 1G). The trial-by-trial change in the pupil dilation velocity averaged over this window showed a sharp rise on the introduction of force field and gradual decline as the participants adapted to the force field and the movement error decreased (Fig.1H; Fig. 1D). Furthermore, the evoked pupil dilation (online pupil response) was also followed by increased baseline pupil diameter in the following trial (offline pupil response) (Fig. 1I, black arrows) (t_{47}=-2.855, p=0.006, two-sample t-test). This pupil response pattern is

consistent with recent views that task-induced pupil dilation and baseline pupil diameter may reflect surprise (e.g., unsigned prediction error) and subjective uncertainty about the environment (volatility), respectively (Preuschoff et al., 2011; Nassar et al., 2012). Visual inspection also suggested the increased subjective uncertainty about the task after the short rest periods as reflected in the slight increases in the baseline diameter at the beginning of a new block (Fig. 1I).

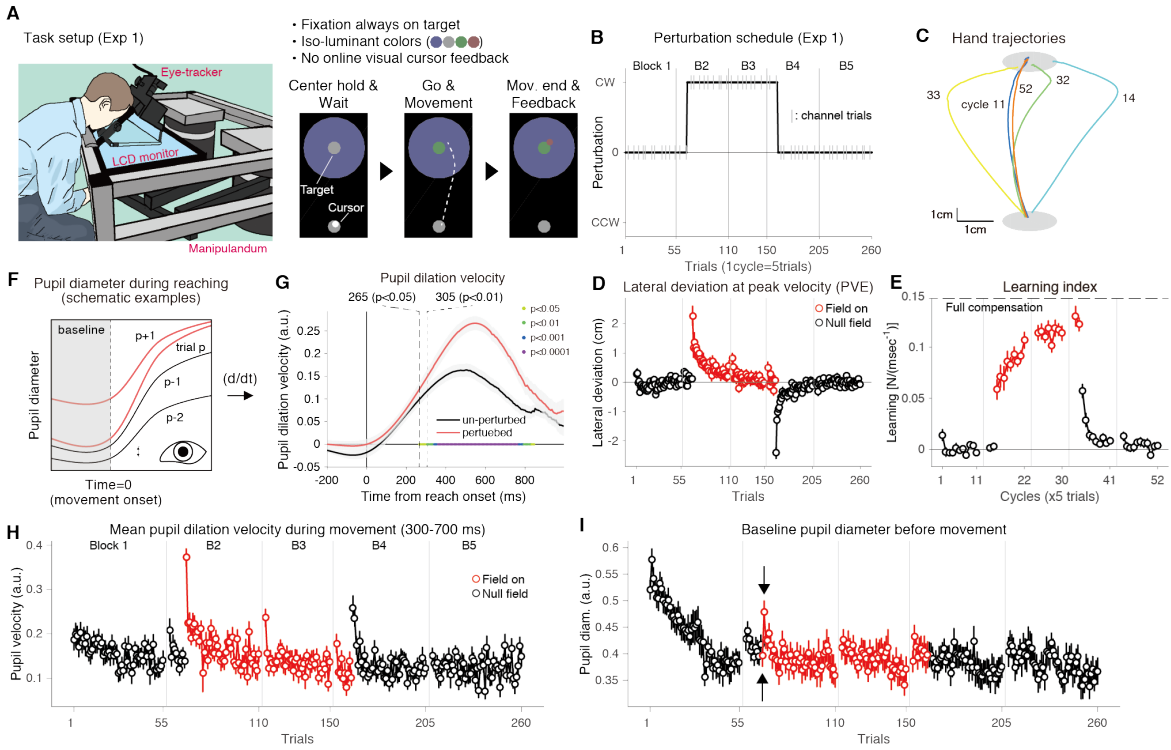

Figure 1. Pupillometry during force-field motor adaptation (Experiment 1). (A) *Left: Experimental setup used in Exp 1. Right: Schematics for a single trial. Participants were required to fixate on the center of the target.* ***(B)*** *The perturbation schedule for the Exp 1. The manipulandum applied the clockwise (CW) force field unexpectedly to the participants at the second block. The gray vertical lines indicate the block boundary. The short gray vertical lines represent the force channel trial where learning was quantified.* ***(C)*** *Hand trajectories of a representative participant. The cycles shown (11, 14, 32, 33, and 52) correspond to baseline, early perturbation, late perturbation, early washout, and late washout trials, respectively.* ***(D)*** *Mean lateral hand deviation at the peak tangential handle velocity. Positive values correspond to rightward deviation.* ***(E)*** *Mean learning index measured in the channel trials (once in a cycle). Learning index was the lateral force to the channel at the time of peak velocity divided by peak velocity (i.e., viscosity). The red dots represent values for perturbation trials (D, E).* ***(F)*** *A schematic example of pupil diameter timeseries during the current reaching paradigm aligned to the reach onsets. In general, the pupil shows different baseline values and dilates during reaches. On a sudden perturbation trial (trial=p), it strongly dilates. On the following trials (trial=p+1), it shows higher baseline value (see also, arrows in panel I).* ***(G)*** *The online pupil response to the force perturbation; the first time-derivative (=velocity) of pupil diameter timeseries (F) averaged over the baseline trials (un-perturbed; black trace) or the first 5 perturbation trials at the second block (perturbed; red trace). The gray area represents s.e.m. across the participants. The two vertical dashed lines indicate the time of significant difference by running paired-t tests. The dots on the y=0 line with different colors represents the significance levels (yellow: p<0.05, green: p<0.01, blue: p<0.001, purple: p<0.0001).* ***(H)*** *Trial-by-trial changes in the pupil dilation velocity averaged between 300 to 700 msec since the movement onset.* ***(I)*** *Trial-by-trial*

changes in the baseline pupil diameter. The red circles represent values for perturbation trials (D,E,H,I). Error bars correspond to s.e.m. across the participants (D,E,H,I). The arrows indicates the first two perturbation trials (upward: the first trial, downward: the second trial).

The pupil response features also showed the similar pattern with the surprise (unsigned prediction error) and volatility signals simulated by a recent computational model called Volatile Kalman Filter (VKF; Piray and Daw, 2020a) (Fig. S1). The VKF, representing the current notion about dynamic modulation of learning rate by volatility, dynamically adjusts learning rate according to the volatility estimated based on experienced prediction error on a trial-by-trial basis and is able to reproduce a wide range of animals' choice behavior. As the model updates the volatility estimate based on the squared prediction error, a sudden change in the environment (i.e., force-field introduction) resulting in a sudden increase in error size increased volatility estimate in the following trials, similar to the offline pupil response to the force-field (see Methods).

Rapid meta-habituation of task-evoked pupil responses after repeated exposure to change points.

The results from Exp 1 suggest that pupil diameter can be a good indicator of participants' subjective degree of surprise and belief about environmental uncertainty (volatility) also in motor learning paradigms. In Exp 2, to further get insights about how these variables are reflected in more complex situation, we monitored pupil responses of another 29 participants while they reach in the presence of force fields with unexpected reversals (Fig. 2A,B). The basic response property of pupil dilation to the unexpected perturbation (first introduction of force-field) was consistent between experiments 1 and 2 (Fig. 2C).

The participants were divided into three sub-groups (2A, B, and C) and experienced different force field schedules (Fig. 2B). For instance, Exp. 2A and 2B were different in frequency and timing of 'change point' in which the presence, absence, or direction of the force field changed unexpectedly to the participants (see Methods). At the change points the participants experienced sudden increases in the movement error (Fig. 2D) which imply the change in the environment (i.e., the direction of force) and would transiently increase the uncertainty about the current environmental identity, similarly to the case of probabilistic learning task with change points (Nassar et al., 2012; Muller et al., 2019; Piray and Daw, 2020a). Thus, we expected that 1) increases in perturbation-evoked pupil dilation and baseline pupil diameter following each change point (due to increases in error), and therefore, 2) differences in the change point statistics would lead to different profile of these online and

offline pupil responses. Simulation of the three force field schedules (Fig. 2B) with the VKF model (Piray and Daw, 2020a) predicted that the surprise (unsigned prediction error) and volatility estimate also follow above expectations (Fig. S2).

Intriguingly, however, the participants' pupil responded quite differently from the above expectation. Except for the first few change points, both the pupil dilation velocity and baseline pupil diameter showed rapid decline in response amplitude (Fig. 2E,F; Fig. S2C,D), despite the sharp increase in movement error at every change point (Fig. 2D). Such decrease in pupil response to errors may imply that the error sensitivity for updating subjective volatility and surprise may be more dynamic than currently assumed.

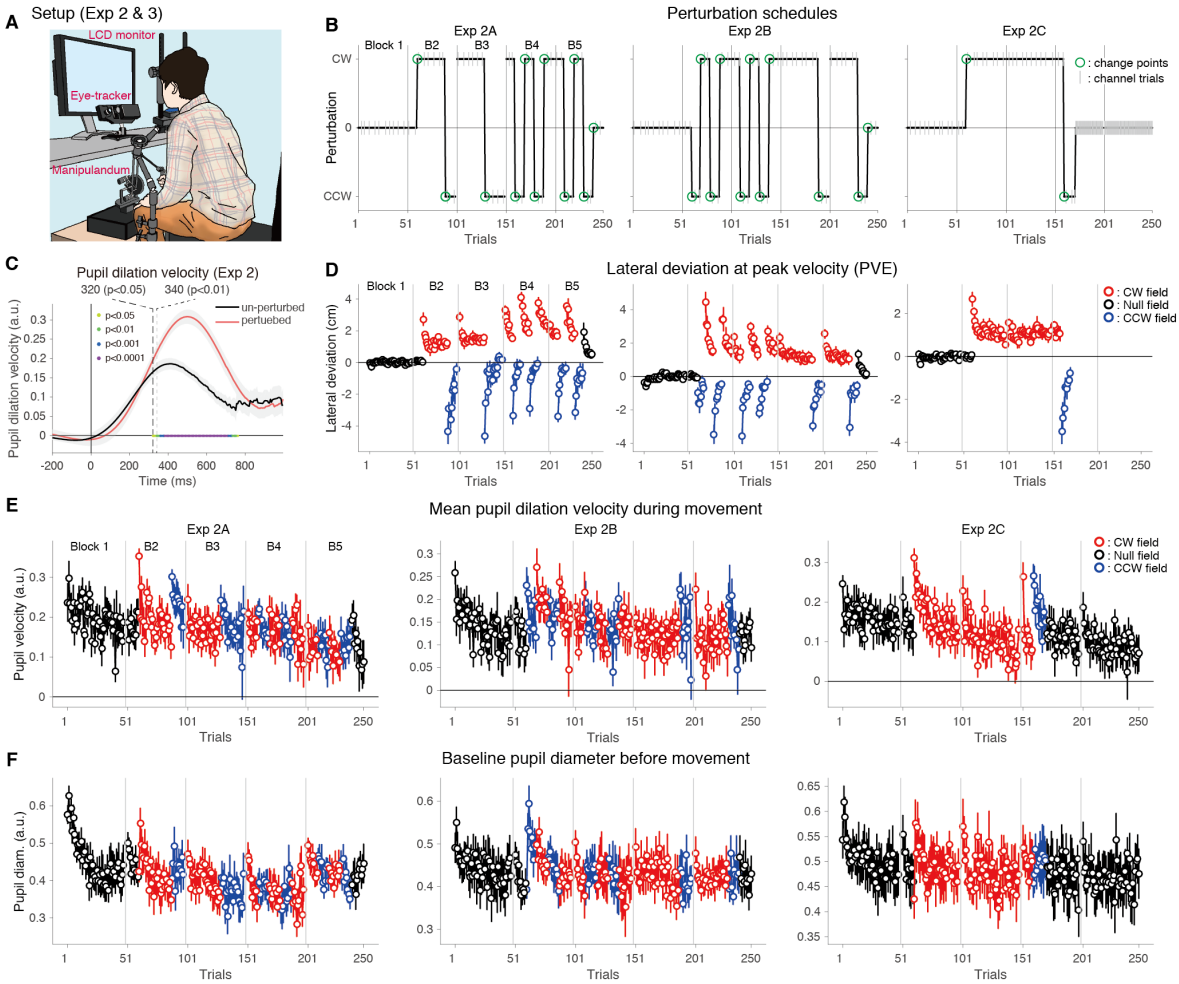

Figure 2. Pupil responses in schedules with multiple reversals in force field. ***(A)*** *Experimental setup for the Experiments 2 and 3.* ***(B)*** *Perturbation schedules for the Exp 2 (Exp 2A, 2B, and 2C). The green circles indicate the 'change point' trials where either magnitude (on/off) or direction (CW/CCW) of the force-field changes in the middle of the blocks (changes across the blocks are excluded). The gray vertical lines indicate the block boundary.* ***(C)*** *Time course of the pupil dilation velocity averaged over the baseline trials (un-perturbed; black trace) or the first 5 perturbation trials at the second block (perturbed; red trace), averaged over all sub-groups. The gray area represents s.e.m. across the participants. The two vertical dashed lines indicate the time of significant difference by running paired-t tests. The dots on the y=0 line with different colors represents the significance levels (yellow: p<0.05, green: p<0.01, blue: p<0.001, purple: p<0.0001).* ***(D, E, F)*** *Trial-by-trial*

change in the mean lateral hand deviation at the peak tangential handle velocity (D), the evoked pupil dilation velocity (E), and baseline pupil diameter (F) for the Exp 2A (first column), 2B (second column), and 2C (last column), respectively. The colors of circles (black, red, and blue) indicates the data for baseline, CW field, and CCW field, respectively. Positive values for the panel D correspond to rightward deviation. The gray vertical lines indicate the block boundary. Error bars represent s.e.m. across the participants. Trials at which the dot color changed indicate the change points.

To formally quantify the decline in pupil responses, we focused on how they behaved around each change point. In the early perturbation trials, the participants' pupil reacted consistently with our expectation. Around the first change point, both the online and offline pupil responses showed clear increase in response to the sudden introduction of the force field (Fig. 3A,C), with the online pupil response increasing immediately on the change point trial and the baseline diameter increasing in the next trial. However, as noted above, for the subsequent change points the response amplitude in the baseline pupil diameter and the pupil velocity showed a dramatic decline (Fig. 3B,D). A linear mixed-effect model on the baseline pupil diameter confirmed the significantly negative slope on the change point (t_{234}=-5.6, p=5.7×10^{-8}) after controlling for the random group/individual factors and the effect of kinematic and endpoint errors at each change point (see Methods). This was unexpected because the size of unsigned error increased almost two-fold at the second change point due to the reversal of force field direction (Fig. 2D). Such reduced sensitivity to error is unlikely due to fatigue or boredness alone, as the participants in Exp 2B already showed a large decline in baseline diameter change already at the second change point (Fig. 3B) which was still in the second block of the experiment (70th trial). Similarly, the pupil dilation velocity showed decreased sensitivity to errors. Although there was no significant monotonic decrease (t_{175}=0.04, p=0.93, the same linear mixed-effect model without the effect of endpoint error) as seen in the baseline pupil diameter, visual inspection suggested that this was due to the sudden increase in the online pupil response on the 9-th change point for the Exp 2B data (Fig. 3D). In fact, the slope up until the 8-th cp was significantly negative even after correcting for the error (t_{116}=-2.2, p=0.027, linear mixed-effect model), suggesting the diminished sensitivity to physical error in online pupil response, similar to the offline pupil response to the change points. The increase in response amplitude seen in the pupil dilation velocity (Fig. 3D; Exp 2B) may reflect the recovery from the habituation during the period of constant perturbation direction between the 8-th and the 9-th change point (Fig. 2B), although the interaction between change point (8 and 9) and group (Exp 2A and 2B) remained marginal (χ_1^2=3.4, p=0.064, likelihood ratio test between models with vs. without interaction). These results indicate that both online and offline pupil responses do not

faithfully reflect the physical size of errors but exhibited some form of "habituation" to the sudden increases in errors around the change points. Notably, as this is different from the simple habituation to the repeated exposure to the certain force perturbation (e.g., reduction in mean pupil dilation velocity in Fig. 1H), we have termed this "meta-habituation". Meta-habituation may possibly reflect the acquisition of the knowledge about task structure itself (e.g., the presence of change points), which is hierarchically higher than the task at hand (i.e., detail about the force-field).

Interestingly, such meta-habituation of surprise or subjective environmental uncertainty, though it sounds intuitive, has not previously been mentioned in decision-making literature where the relation between environmental uncertainty and pupil responses has been most rigorously studied (e.g., Preuschoff et al., 2011; Nassar et al., 2012; Muller et al., 2019). When we applied the same analysis to the data simulated by the VKF model (Piray and Daw, 2020a), the change in estimated volatility as well as the surprise (unsigned prediction error) showed very similar patterns with our pupil response data for only the first change point. (Fig 3E-H). While the simulated volatility/surprise showed clear increase in change point-induced response amplitude at the second change point and stayed large up until the 10-th change point (Fig. 3F, H), the observed pupil response exhibited markedly different pattern (Fig. 3B, D). This is due to the fact that the size of the prediction error has near-constant impact on the update of volatility estimate in VKF and other similar class of models (Mathys et al., 2011; Piray and Daw, 2020a). Our results thus demonstrate that in computing environmental surprise and uncertainty, the sensitivity to incoming information (i.e., prediction errors) is not fixed, but dynamically modulated.

We next asked whether/how the uncertainty/surprise signal that is reflected in pupil-linked arousal affects the motor learning processes. However, it was not easy to directly assess this question in the data from Experiment 1 and 2, as the design were not optimized for the accurate quantification of motor learning. To answer the question, we ran another series of experiments.

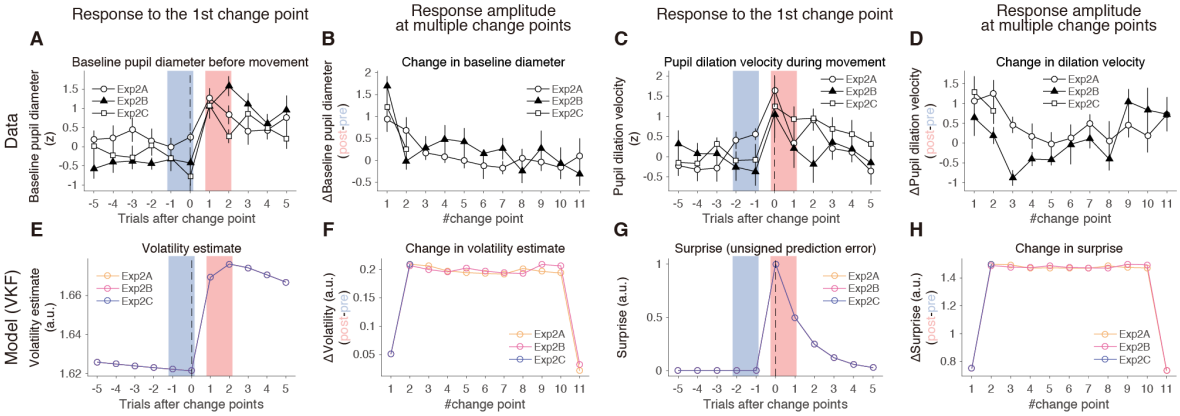

Figure 3. Pupil responses around multiple change points. (A) *Trial-by-trial change in the z-transformed baseline pupil diameter around the first change point. Pre- and post-change point values are defined by blue and red shades, respectively.* ***(B)*** *Amplitude of change point-induced baseline pupil diameter change (difference between post- vs. pre-change point values in A) as a function of the number of change points.* ***(C)*** *Trial-by-trial change in the z-transformed mean pupil dilation velocity around the first change points. Similarly, pre- and post-change point values are defined by blue and red shades, respectively.* ***(D)*** *Amplitude of cp-induced mean pupil dilation velocity change (difference between post- vs. pre-change point values in C) as a function of the number of change points.* ***(E-H)*** *The same analysis as A-D applied to the volatility estimate (E, F) and the surprise (G, H) simulated by the Volatile Kalman Filter (VKF) model using the same perturbation schedules (Exp 2A-C). The model parameters were estimated using the data from experiment 1. Note that, in F and H, the large increase (decrease) at the second (11th) change point reflects increase (decrease) in movement error due to the first force reversal (return to Null field) (see Fig. 2B, D). The error bars indicate s.e.m. across the participants.*

Manipulation of arousal state modulated trial-by-trial learning rate in a state-dependent manner.

To further investigate the link between pupil-linked arousal state and trial-by-trial motor learning, we designed another set of experiments that directly manipulated arousal state (Exps. 3A and 3B). To accurately estimate trial-by-trial learning rate, we employed a single-trial learning paradigm in which participants experienced a triplet of trials consisting of channel, perturbation, then channel trial, spaced by 2-6 null-perturbation trials (Fig. 4A, B). We quantified the single-trial learning by comparing the lateral force immediately before and after the perturbation trial measured by force-channel (e.g., Gonzalez Castro et al., 2014). As shown by Fig. 4D, the participants in Exp 3A and 3B exhibited similar amount of single-trial motor learning. We quantified the single-trial learning as the integral of force time series within the window from -200 to 0 msec relative to the time of peak movement speed (green shades, Fig. 4D). To test the direct link between arousal state and learning rate, we manipulated participants' arousal state through task-unrelated auditory stimuli (IADS2, Bradley et al., 2007) applied during the preparatory period for reaching during perturbation trials (Fig. 4B). As presented by Fig. 4C, the auditory stimuli induced reliable pupil dilation

according to its arousal scores (Low, High, or Null conditions) indicating that participants' arousal states were manipulated accordingly.

As there is a known baseline-dependent (inverted-U type) relationship between arousal and performance (Yerkes and Dodson, 1908), we analyzed the manipulation effect taking individuals' baseline arousal level, which was calculated as the normalized mean of baseline pupil diameter throughout the experiment (see Methods), into account. As expected, between-individual comparison of the learning data indicated the sign of the inverted-U relationship between the single-trial learning vs. mean arousal (Fig. 4E, first and second column). Interestingly, the comparison between High vs. Null/Low arousal conditions indicated that the direction of the effect of arousal manipulation flipped as the participants' mean arousal level increased (Fig. 4E, third and the last column), also consistent with the inverted-U type effect of arousal.

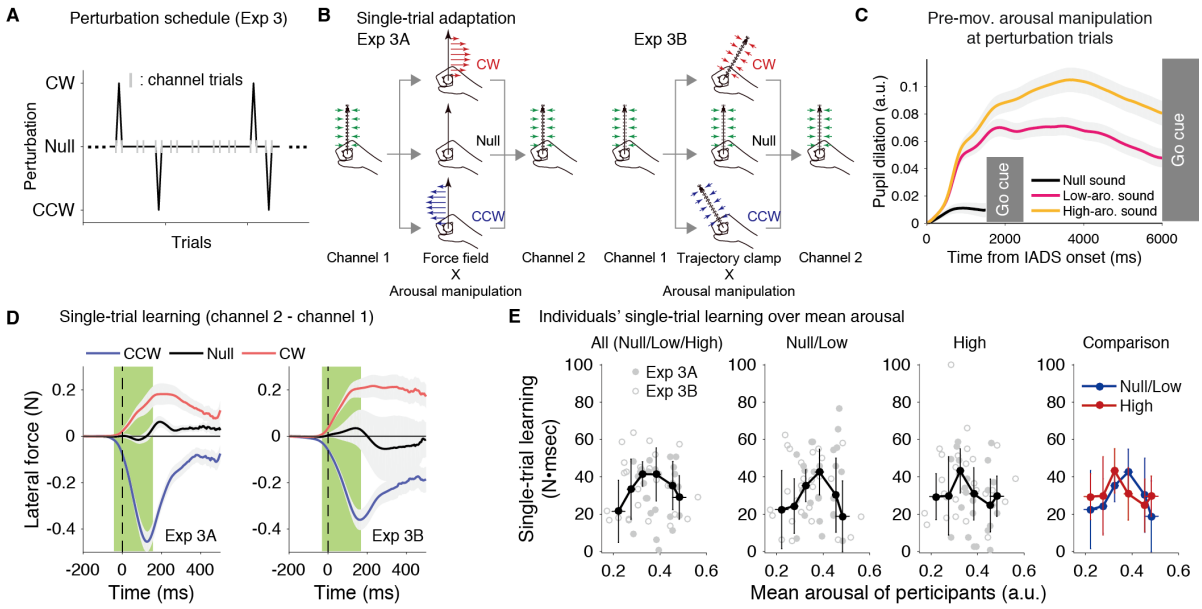

Figure 4. Arousal manipulation during single-trial motor learning paradigm. (A) *An exemplar part of the perturbation schedule for Exp 3A and 3B. Perturbation trials with channel trials immediately before and after (B,C) were separated with trials with Null perturbation.* ***(B)*** *The single-trial learning was quantified by measuring the lateral force output by participants immediately before and after the perturbation trial by using channel trials. For the Exp 3A, either null, CW, or CCW velocity-dependent force-field was applied, similarly to Exps 1 and 2 (left). For the Exp 3B, either CW (30 degree), CCW (-30 degree) or straight (0 degree) trajectory clamp was applied as perturbation to get better control of kinematic errors (right). Arousal manipulation was applied on the selected perturbation trials in randomized order.* ***(C)*** *Averaged time course of baseline pupil diameter change during pre-movement period for Null-sound (black), Low-arousal sound (magenta), and High-arousal sound (gold yellow). The presentation of IADS2 auditory stimuli (Bradley and Lang, 2007) started at time=0 and lasted for 6 seconds. For Null-sound condition, pre-movement period was set to 1500-2000 msec. (D) Averaged time series of single-trial learning (difference between force output at channel 2 vs. channel 1 trials, red: CW, black: Null, blue: CCW) for Exp 3A (left) and 3B (right). Trial-by-trial single-trial learning was assessed as the force impulse integrated over the time window*

of [-200 ~ 0 msec] relative to the time of peak velocity (the green shades). ***(E)*** *Individuals' single-trial learning values plotted over corresponding individuals' mean arousal (average of baseline pupil diameter for non-perturbation trials normalized by individuals' pupillary light reflex amplitude, see Fig. S3) averaged over all conditions (1st column), Null / Low conditions (2nd column), or High condition (3rd column). The gray dots represent individual participants' data. The black dots represent the median calculated over the 10, 25, 50, 75, and 90 percentiles of mean arousal. The error bars represent the median absolute deviation. Comparison between High vs. Null/Low averages is shown on the 4th column for visualization purpose.*

To more accurately assess the effect of arousal manipulation on the single-trial motor learning, we compared z-scored single-trial learning data between High vs. Null/Low conditions, controlling for the effect of error size and other confounding factors (see Methods). The resultant pairs of z-scored delta-learning values and mean arousal clearly indicated the presence of inverted-U type relationship between individuals' arousal and single-trial motor learning rate; increasing arousal by manipulation increased learning rate when the participants were in lower arousal level, but decreased it when the participants were in higher arousal level (Fig. 5). If there was no overall effect of arousal manipulation as well as no relationship between arousal manipulation effect and participants' mean arousal, the data should distribute evenly over the four quadrants defined by z=0 and the median arousal lines. In contrast, we found significantly higher density over the second and the fourth quadrants as evident in the density map (Fig. 5, likelihood ratio test, $G = 7.67$, $p = 0.0056$). Thus, our data shows that the pupil-linked arousal (and environmental uncertainty and surprise it reflects) has direct effect on trial-by-trial motor learning rate in a baseline-dependent manner.

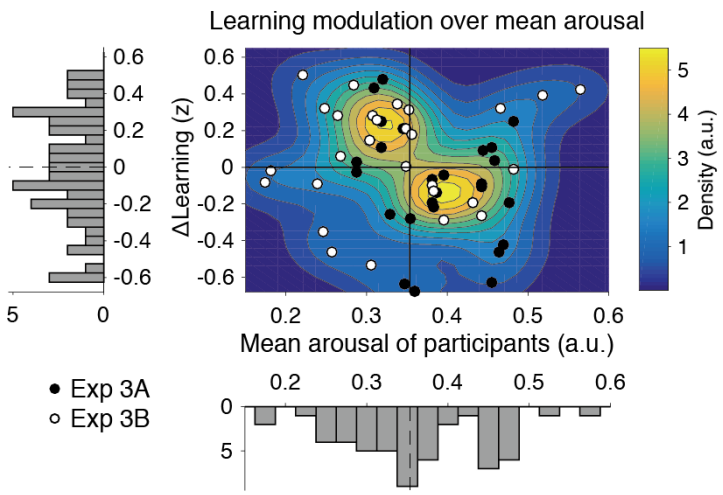

Figure 5. The baseline-dependent effect of arousal manipulation on single-trial motor learning. *The effect of arousal manipulation on single-trial learning was individually assessed by first z-scoring the single-trial learning (e.g., Fig. 4E) and then comparing between High vs. Null/Low conditions. The resultant z-scored delta-learning values of the individual participants were plotted over their mean arousal (black dots: participants for Exp 3A, white dots: participants for Exp 3B). The kernel*

density estimates of overall observations (x-y pairs) indicates that the arousal effects on single-trial motor learning are significantly more concentrated in the second and the fourth quadrants than chance (G = 7.67, p = 0.0056), indicating positive/negative effect for participants with lower/higher mean arousal level. The horizontal and vertical lines represent zero effect and median of mean arousal, respectively. The histograms for delta-learning (upper left) and mean arousal (lower right, the dashed line denotes the median arousal value) are also shown.

Discussion

In the present study we simultaneously monitored pupil diameter as a non-invasive proxy for the activity of central noradrenergic/arousal system while human participants perform typical motor learning tasks. Our results are on the one hand consistent with the recent view that the pupil-linked arousal reflect participants' subjective uncertainty about environment as well as surprise about observations (e.g., unsigned prediction errors) (Preuschoff et al., 2011; Nassar et al., 2012; Lavín et al., 2014; Murphy et al., 2016; Muller et al., 2019; Vincent et al., 2019) (Exp 1). On the other hand, our results further revealed (a) the novel meta-habituation of these subjective uncertainty/surprise to perturbations which was not solely explained by fatigue or boredom (Exp 2) and (b) the evidence of modulation of trial-by-trial motor learning rate by the pupil-linked arousal (Exp 3). These findings imply that the subjective uncertainty/surprise can modulate the trial-by-trial learning rate also in the motor system through pupil-linked arousal system.

Uncertainty and motor learning

Statistically optimal learning algorithms take multiple sources of "uncertainty" into account to dynamically modulate learning rate. One such algorithm is the Kalman filter (Kalman, 1960) which dynamically modulates the learning rate (the Kalman gain) by optimally balancing between the uncertainties in terms of observation noise and current estimate of the state, assuming the environment itself is stable. It has been repeatedly suggested that the humans also employ such Kalman filter-like adjustment of the learning rate both in motor (Wolpert et al., 1995; Baddeley et al., 2003; Burge et al., 2008; Wei and Körding, 2010) and other domain of learning (Rao, 1999; Shibata et al., 2009). When the environment itself can also change, learning algorithms need to additionally estimate how volatile the current environment is and take it into account for the update of learning rate (Yu and Dayan, 2005; Behrens et al., 2007; Nassar et al., 2010; Mathys et al., 2011; Iigaya, 2016; Piray and Daw, 2020a). While the plausibility of such volatility learning models in human learning has been extensively tested mainly in the context of reinforcement learning using decision-making tasks (Behrens et al., 2007; Nassar et al., 2010, 2012; Mathys et al., 2011, 2014; de Berker et al., 2016; Soltani and Izquierdo, 2019), less attempts have been made in the motor learning field, except for a few recent studies (Huang and Shadmehr, 2009; Gonzalez Castro et al., 2014; Herzfeld et al., 2014). Together with these previous results, our data provides

compelling evidence that motor learning rates are dynamically modulated by environmental uncertainty, possibly through the NA/arousal system.

One intriguing question is whether and to what degree each of the "explicit" and "implicit" components of motor learning (Taylor and Ivry, 2011; Taylor et al., 2014; Schween et al., 2020) is involved in the uncertainty/surprise-driven modulation of motor learning rate. As we did not employ the experimental design that explicitly quantified these components, this is still an open question. Notably, however, a very recent study (Albert et al., 2021) demonstrated that the learning rate for the "implicit" motor learning process is sensitive to environmental statistics (e.g., perturbation consistency), suggesting that the "implicit" motor learning process is likely modulated by environmental uncertainty/surprise.

Pupil diameter, central noradrenergic activity, and uncertainty

Under constant light, pupil diameter is known to reflect the central noradrenergic (NA) activity (Rajkowski et al., 1993; Aston-Jones and Cohen, 2005; Joshi and Gold, 2020), or at least the balance between sympathetic (mediated by NA) and parasympathetic (mediated by acetylcholine, ACh) tones (Steinhauer et al., 2004; Szabadi, 2013). Computationally, it has been hypothesized that NA signals the unexpected uncertainty, ACh signals the expected uncertainty, and the balance between these neuromodulators is crucial for optimally adjusting the learning rate (Yu and Dayan, 2005). Consistent with this theory, recent accumulation of evidence suggests that the task-evoked phasic change in pupil diameter may reflect some "unexpectedness", such as surprise (e.g., degree of violation from expectation) (Preuschoff et al., 2011; Kloosterman et al., 2015; Alamia et al., 2019; Joshi and Gold, 2020), as well as that baseline pupil diameter may reflect subjective uncertainty about current environmental identity (e.g., subjective estimate of environmental volatility) (Jepma and Nieuwenhuis, 2011; Nassar et al., 2012; Muller et al., 2019; Filipowicz et al., 2020). Our data is in line with this notion in that perturbation-evoked pupil dilation and baseline pupil diameter showed similar pattern with surprise (or unsigned prediction error) and model-estimated volatility, respectively, especially for the first few perturbation events. How, then, are NA and/or ACh activities related to the online and offline pupil responses we observed? Notably, a recent rodent study has reported that while slow pupil diameter change on the timescale of a few second is correlated with the cortical level of both the NA and ACh, more rapid change in pupil diameter and dilation velocity on shorter timescales is more strongly correlated with NA than ACh (Reimer et al., 2016). Thus, the perturbation-evoked pupil dilation during

reaches is more likely reflect the phasic burst of NA activity signalling unexpectedness, while the trial-by-trial change in baseline pupil diameter may reflect the mixed effect of how cortical NA and ACh signal expected and unexpected events. Interestingly, recent studies reported in the early period of human motor adaptation the transient increases in muscle co-activation (Franklin et al., 2008), long-latency stretch reflex gain (Coltman and Gribble, 2020), as well as muscle spindle firing (Dimitriou, 2016). These modulations may, thus, be associated with transient increase in the subjective uncertainty about the environment and resultant increase in baseline NA (and maybe ACh, as well) activity.

Whereas NA may play a key role in modulating muscle/spindle activation patterns and transcortical reflexes via respective LC projections to both to the spinal cord and the motor cortex (Szabadi, 2013), serotonin (5-HT) is also known to modulates the muscle reflex-gain in human spinal cord (Wei et al., 2014). A very recent study has reported that the activity of dorsal raphe (DR) 5-HT neurons in mice also tracks environmental uncertainties (Grossman et al., 2020). However, bidirectional connection between DR and LC (Szabadi, 2013) complicates tracking the origin of uncertainty signal reflected in these neuromodulators. It is also important to note that LC also receives inputs from the superior colliculus (SC) which also leads to pupil dilation (Corneil and Munoz, 2014). Although we carefully removed trials with saccades, some of observed online pupil dilation might reflect this more automatic orienting response to the perturbation. Nevertheless, the clear modulation of pupil dilation we observed suggests at least that such orienting response too is under the influence of the background LC activity which reflects environmental uncertainty. Overall, as the pupil diameter is an indirect and not a pure measure of central NA activity, more direct approach, such as invasive animal study or pharmacological manipulation, is required to further elucidate the roles of these neuromodulators (i.e., NA, ACh, and 5-HT) in the motor learning process.

Meta-habituation of pupil responses and acquisition of task knowledge

One novel finding that has not been well documented in the previous studies is the rapid reduction of surprise/uncertainty-like pupillary responses at force field change points across repeated exposures (experiment 2). We have called this meta-habituation, which may indicate the dynamic partitioning of uncertainty into expected vs. unexpected sources over the course of task experience, presumably through the interaction between bottom-up and top-down processes in assessing environmental uncertainty and surprise about observations (Filipowicz

et al., 2020; Joshi and Gold, 2020). In other words, acquisition of higher-level knowledge about the task structure (i.e., the presence of change points) might have made the sudden increase in motor errors no longer surprising, but somehow "expected". This idea fits well with the hypothesized roles of NA and ACh as indicators of unexpected and expected uncertainty (Yu and Dayan, 2005) and the fact that tonic, baseline pupil diameter likely reflects the balance between the two neuromodulators (Steinhauer et al., 2004; Szabadi, 2013; Reimer et al., 2016); The initial increase in unexpected uncertainty (NA) in earlier change points may have been gradually overridden by the slower increase in expected uncertainty (ACh) following the acquisition of the knowledge about the task structure. Consistent with this idea, a recent study has demonstrated that the pupil dilates in response to novel and unexpected objects, while it constricts to novel but expected objects (Kafkas, 2021). Similarly, the dynamic change in the sensitivity of online pupil dilation to movement errors observed in Experiment 2 resembles the scaling of reward prediction errors by the expected variability in rewards (Preuschoff and Bossaerts, 2007; Diederen and Schultz, 2015).

It is also conceivable that the acquisition of knowledge about the task structure may spontaneously proceed in a hierarchical manner (Collins and Frank, 2013; Sarafyazd and Jazayeri, 2019). In the current study, it would be reasonable to assume that the participants represented the task environment following a hierarchical structure shown in the Figure. 6A. Although highly speculative and needs direct test in future studies, the level of uncertainty/surprise accompanying the (estimated) state transition would be proportional to the vertical, across-hierarchy, distance it requires when moving from one state to another. For example, (estimated) transition within the same tree (e.g., between CW and CCW in Fig. 6A) accompanies lower uncertainty/surprise compared to that across the trees once the task structure is known (e.g., between Null and Perturbed, or between Experiment and Daily life in Fig. 6A). These assumptions are consistent with the common observation that baseline pupil diameter almost always shows highest value at the beginning of the whole experiment (i.e., transition from daily life state to experiment state), but the size of modulations by experimental manipulations rarely exceeds the initial baseline diameter and gradually diminishes with time, as the participants get more experience about the experiment. Recent evidence also suggests the (reword) prediction error (Zacks et al., 2011; Rouhani et al., 2020) as well as phasic pupil responses (Antony et al., 2020; Clewett et al., 2020) signals the subjective belief about environmental (contextual) change which helps to create event boundaries in memory structures (Clewett et al., 2020; Rouhani et al., 2020).

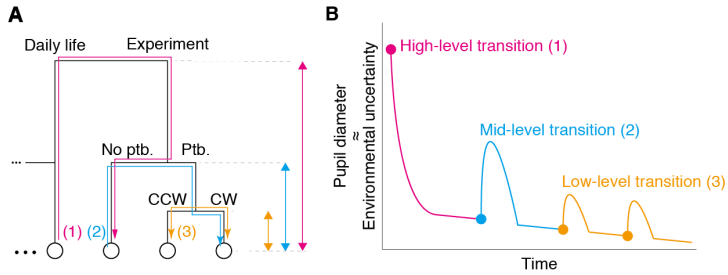

Figure 6. Hypothetical relationship between hierarchical state estimation/transition and environmental uncertainty/surprise. (A) *A likely hierarchical representation of task and daily life states for the current experiment.* ***(B)*** *A possible time-course of pupil diameter (and subjective environmental uncertainty). The degree of elicited uncertainty (pupil diameter) is proportional to (subjective) "across-hierarchy" distance between the two states (vertical double arrows on the right of panel A.*

Learning of higher-level task features (e.g., change points) certainly requires longer period of observation than that of lower-level task features (e.g., direction of force perturbation). We thus speculate the process of meta-habituation involves multiple, or at least two, timescales of memory updating with shorter timescales likely corresponding to the update of unexpected uncertainty and the longer timescales corresponding to the update of expected uncertainty. Although the exact underlying computation is still unclear, several recent computational studies might help to understand dynamics in updating subjective belief about unexpected and expected environmental uncertainty. First, Piray and Daw (Piray and Daw, 2020b) proposed the extended version of their original VKF (Piray and Daw, 2020a) that explicitly estimates volatility (= unexpected uncertainty) and unpredictability (≈ expected uncertainty), which was fixed in the original VKF, on the trial-by-trial basis. Next, and perhaps more related to our above idea of hierarchical task-feature learning, Heald et al. (Heald et al., 2020) proposed a (motor) learning algorithm that simultaneously estimates necessary (motor) output and current environmental state (context) based on hierarchical Dirichlet process mixture modelling. At each trial, the model emits the probability that the current state is novel, similar to the environmental uncertainty. While both Bayesian models rely on computationally more extensive sampling-based learning method, unlike the typically-used delta-rule type learning algorithms (Mathys et al., 2011; Piray and Daw, 2020a), due to their complexity, whether the brain actually employs the sampling-like learning is currently under debate (Fiser et al., 2010; Pouget et al., 2013).

Baseline-dependent modulation of motor learning rate by pupil-linked arousal system

As the LC's rich innervation throughout the brain includes areas crucial for motor learning, such as motor cortex and cerebellum (Sara, 2009; Szabadi, 2013), we expected that the background NA activity could affect the error-based motor learning process in humans. Several earlier rodent studies have suggested the critical role of adrenergic receptors in the cerebellum in acquisition of new motor skills (Watson and Mcelligott, 1984; van Neerven et al., 1990; Tan et al., 1991; Heron et al., 1996; Cartford et al., 2004). The current study demonstrated for the first time in humans that NA/arousal system can modulate motor learning rate in a baseline-dependent manner. Similar baseline-dependent effect of arousal manipulation has also been reported in the previous decision-making studies with auditory or pharmacological manipulation of NA/arousal (Nassar et al., 2012; Jepma et al., 2016, 2018). Together, our finding thus underscores the common role of the NA/arousal activity in modulation of the learning rate in both cognitive and motor domains in humans.

The baseline-dependent manipulation effect indicates the inverted-U relationship between NA/arousal and learning rate (Yerkes and Dodson, 1908). Such inverted-U relationship implies distinctive functions of NA at different concentrations, potentially through the different adrenergic receptor families: the alpha-1, the alpha-2, and the beta receptors. Whereas a large body of evidence, mainly in the context of PFC function, suggests that the recruitment order of NA receptors along the low to high NA concentrations is alpha-2, alpha-1, and then beta receptors (Ramos and Arnsten, 2007; Atzori et al., 2016), some studies have reported the opposite (i.e., beta at lower concentrations and alpha-2 at higher concentrations) in the cerebellum (Basile and Dunwiddie, 1984; Wakita et al., 2017). A recent study has shown that the NA can selectively reduce the release probability in the climbing fibre synapses on cerebellar Purkinje cells via the alpha-2 receptor which results in reduced plasticity at the parallel fiber-Purkinje cell synapses (Carey and Regehr, 2009). Therefore, although more work is needed, this at least explains the smaller learning rate at either lower or higher extreme of NA concentrations (arousal). The declined learning rate for very large movement errors (Wei and Körding, 2009) could be explained by the similar effect of too-high NA/arousal induced by experiencing very large error.

LC also receives input from various brain regions, such as the amygdala (Bouret et al., 2003; Szabadi, 2013). In fact, strong stress (e.g., fear) leads to higher NA concentrations (Atzori et al., 2016) which facilitates the formation of fear memory mediated by the beta adrenergic receptor in an amygdala-dependent manner (Strange et al., 2003). This leads to an

interesting possibility that not only environmental uncertainty, but also emotional/mood states could affect motor learning through NA/arousal system. In conclusion, our results indicate the importance of NA/arousal system in motor learning both in the context of sports and rehabilitation and highlight the utility of pupil diameter as a new interesting window into the motor system.

Materials and Methods

Participants:

We recruited 28, 30, 27, and 30 right-handed participants with no history of neurological disorders for Experiment 1, 2, 3, and 4 respectively (73 males, 38 females, age: 19-37). They participated in the current study after providing written informed consent. All the experimental protocols were approved by ethical committees of the University of Western Ontario (Experiment 1), Osaka University (Experiment 2), and Center for Information and Neural Networks (Experiment 3).

Experimental settings:

General settings:

The participants were instructed to make straight center-out reaching movement from a starting position to a goal target in a 2D plane holding the handle of a robotic manipulandum. The starting position (white circle, 1.6 cm diameter), the target (gray circle, 1.6 cm diameter), and the participants' current hand position (white dot cursor, 0.5 cm diameter) were displayed on LCD monitors. Throughout the task, we monitored the participants' eye gaze and pupil diameter using eye trackers (EyeLink 1000, SR Research Ltd., Ontario, Canada). To reduce the impact of unnecessary eye movements on pupillometry measurement, participants were required to always maintain fixation on the center of the goal target. Before the main experimental session, the participants went through a familiarization session with continuous visual feedback of their current hand position (i.e., the cursor). In the main experimental sessions, online visual feedback of the cursor during movement was removed and only terminal feedback was provided at movement offset. The colors used for the start position (gray), target (gray/green), background disk (pale blue), and endpoint feedback (magenta) were adjusted to ensure that they are approximately iso-luminant (details are provided in Table 1).

At the start of each trial, the robot's handle automatically moved the participants' hand into the starting position. During a trial, the participants maintained the cursor at the start location for 1 second while maintaining fixation to the center of target. Following this period, we measured the participants' pupil diameter for variable duration (3-11 sec). The start position would then change to green, informing the participants to initiate reaching. Movement was defined as the period where the hand movement velocity was above a

threshold (3.5cm/sec). At the completion of a reaching movement, endpoint feedback was provided by a magenta cursor (0.5 cm diameter) for 1000 msec.

We introduced a velocity-dependent curl force field (Shadmehr and Brashers-Krug, 1997) to establish the relationship between the pupil response and the motor adaptation. The force field was applied according to an equation

$$\begin{bmatrix} f_x \\ f_y \end{bmatrix} = \begin{bmatrix} 0 & B \\ -B & 0 \end{bmatrix} \begin{bmatrix} v_x \\ v_y \end{bmatrix}, \quad (1)$$

where f_x and f_y are the force applied to the handle (N), v_x, and v_y are the velocity of the handle (m/s) for *x*- and *y*-directions, respectively. For the clockwise (CW) force field, the viscosity coefficient B [N/(ms^{-1})] had positive values and for the counter-clockwise (CCW) field B had negative values. To quantify adaptation to the force field, we occasionally introduced "channel" trials, in which the handle motion was constrained on straight path between the home position and the target by simulated dumper and spring (Scheidt et al., 2000), to measure the force applied to the channel.

To make inter-participant comparison of pupil diameter interpretable, we additionally measured the physiological limits of pupil diameter within each participant by eliciting pupillary light reflex by changing the background color of the display from light blue to white (higher luminance) or black (lower luminance) (Table 1). These measured pupil limits were used for the within-individual normalization of pupil diameter data (Fig. S3).

Experiment 1:

Experiment 1 was conducted at the Brain and Mind Institute, University of Western Ontario (Ontario, Canada). The self-reported right-handed participants (n = 28; 13 males, 15 females; age: 24.3 ± 4.5) sat on a height-adjustable chair and held a handle of a robotic manipulundum (1000 Hz control rate) (Reichenbach et al., 2014). The position of the handle was represented as a cursor (white dot) by an LCD monitor (60 Hz update rate) placed directly above the handle to prevent the vision of the hand. The start position (white circle), goal target (gray circle), and a background (light blue circle, 15 cm diameter) to prevent sharp luminance change around the target were also presented on the display throughout the task (Fig. 1B). The eye tracker was mounted on the display to monitor participants' eye gaze and pupil diameter (Fig. 1A). The approximate distance between participants' eyes and the center of the monitor was 16 cm.

An experimental session consisted of 5 blocks of trials (59 trials per each, except that the fourth block consisted of 44 trials). There were short breaks (up to 1 min) inserted between the blocks. On each block, the first and the last 2 trials were used to measure the simple pupil light reflex of the participants. In each of these "light reflex" trials, the LCD screen was suddenly turned either black or white for 2000 msec. The pupil response was measured during and up to 3000 msec after the termination of the black/white color stimulus. The response strengths (i.e., trough for constriction, and peak for dilation) were averaged over the blocks and used to normalize individual pupil diameter during the main task. Following the initial "light reflex" trials participants performed 55 trials of center-out reaching task. On each trial, after confirming stable eye fixation (1000 msec of fixation without any blink) on the target and the cursor staying within the home position, the goal target turned green after variable delay (1000-1500 msec) cueing the reaching movement. To prevent possible predictive/reflexive eye movement and/or pupil dilation due to a moving cursor, visual feedback of the cursor was removed during movement and participants were instructed to maintain fixation on the center of the target. Terminal end-point feedback was provided with a near-isoluminant magenta cursor for 1000 msec when the cursor speed was less than 1 cm/sec w. After the feedback period, the robot handle automatically returned to the home position and then the next trial started.

Experiment 1 employed the typical Null-Force-Null paradigm. The force field was introduced at the 11th trial of the second block and removed at the 11th trial of the fourth block (Fig. 1B). The participants were not informed of on which trial the perturbation is introduced or removed in advance. Twenty participants were exposed with CW force field [B = 0.15 in Eq. (1)] and the remaining eight participants were exposed with CCW force field (B = -0.15). To quantify adaptation to the force field, we interleaved the channel trials in 20% of the trials. The channel stiffness and viscosity were 7000 (N/m) and 30 [N/(ms^{-1})], respectively.

Experiments 2 and 3:

Experiments 2 and 3 were conducted at the Osaka University/Center for Information and Neural Networks (Suita, Japan). The participants sat on a height-adjustable chair and held a handle of robotic manipulundum (Phantom Premium HF 1.5, 3D Systems, Inc., USA). The force field strength was set to B = 0.12 (CW) or -0.12 (CCW). The channel stiffness and viscosity were 2500 (N/m) and 25 [N/(ms^{-1})], respectively. The position of the handle was

visually feedbacked on a vertically placed LCD monitor (60 Hz update rate) as a cursor (white dot). A start position (gray circle), a goal target (gray disk), and a background disk (light blue disk, 21.6 cm diameter) to prevent sharp luminance change around the target were also presented on the display throughout the task (Fig. 2A). The colors used (gray, green, light blue, and magenta) were adjusted to be approximately iso-luminant (Table 1). A desktop eye tracker (Eyelink 1000, SR Research, Ontario, Canada) was placed under the LCD display to monitor participants' eye gaze and pupil diameter (Fig. 2A). The approximate distance between participants' eyes and the center of the monitor was 35 cm for Exps. 2 and 3A, and 44 cm for Exp. 3B. We assessed the participants' handedness by using a Japanese-translated version of the FLANDERS handedness questionnaire (Nicholls et al., 2013; Okubo et al., 2014) which ranges from +10 (perfect right-hander) to -10 (perfect left-hander).

Experiment 2: In Exp 2, thirty right-handed individuals participated for Exp 2. They were divided into three sub-groups (10 each). One participant assigned in Exp 2C did not complete the whole experimental sessions due to frequent blinks during an experimental trial. As a result, the data from 29 participants (24 males, 5 females; age: 22.3 ± 2.7; Handedness score: 9.9 ± 0.3) were analyzed (10 for Exp. 2A, 10 for Exp. 2B, and 9 for Exp. 2C). All three experiments consisted of 5 blocks of 50 reaching trials with short breaks (~1 min) between the blocks. Similar to Exp 1, in each block we added two 2 trials of light-reflex measurements before and after the reaching trials. The perturbation schedules used are characterized with sudden reversals in force field direction and summarized in Figure 2B. For Exp 2A, CW force was applied on trials = {61:89, 101:129, 151:159, 170:179, 190:209, 220:229}, and CCW force was applied on trials = {90:100, 130:150, 160:169, 180:189, 210:219, 230:239}. For Exp 2B, CW force was applied on trials = {70:79, 90:109, 120:129, 140:189, 201:229}, and CCW force was applied on trials = {61:69, 80:89, 110:119, 130:139, 190:200, 230:239}. For Exp 2C, CW force was applied on trials = {61:159}, and CCW force was applied on trials = {160:170}. For all groups (Exp 2A through 2C), channel trials were randomly interspersed in 20% of trials, except that for Exp 2C channel trials were repeated from the 171st trial to the last 250th trial.

Experiment 3: To accurately assess the effect of arousal on the trial-by-trial learning rate, we employed the "single trial" learning paradigm in which updated motor commands are directly estimated by comparing the force output measured at the force channel trial immediately before and after a trial with random perturbations (rightward, leftward, or null force perturbations) (Gonzalez Castro et al., 2014b; Herzfeld et al., 2014) (Fig. 4A). These

"triplets" of trials (i.e., channel-pre, perturbation, and channel-post trials) were interspersed with null trials of variable numbers (2-6 trials) (Fig. 4A). With this setting, we also externally manipulated participants' arousal state by using task-unrelated emotional auditory stimuli (International Affectional Digital Sounds 2: IADS2, Bradley et al., 2007). These stimuli consisted of 167 6-seconds sound clips each of which was rated by independent populations of at least 100 subjects in terms of valence, pleasure, and arousal. We chose the top 40 (High condition) and bottom 40 (Low condition) clips in terms of arousal scores and played them during the pre-movement hold period in the selected perturbation trials. The presenting order of the clips within High or Low conditions were randomized across the participants. Each clip was presented once.

For the Exp 3A, the total of 27 individuals were participated. One participant did not complete the whole experimental session, and the data from the participant was excluded from the analysis. Hence, we analyzed the data from the remaining 26 participants (17 males, 9 females; age: 23.7 ± 5.4; Handedness score: 9.7 ± 0.7). In this experiment, we employed either CCW, Null, or CW force fields as random perturbations (Fig. 4B). In these perturbation trials, we played either High, Low, or no IADS clips. The presenting order was randomized within each block, thus resulting in a 3x3 design. Each condition was repeated for 8 times and interspersed into 8 blocks of 60 trials. We added two 2 trials of light-reflex measurements (3 sec of white/black stimulus-on period followed by 3 sec of stimulus-off period) before and after the main experimental trials of each block.

For the Exp 3B, 30 individuals were participated. The data from the two participants were not analyzed due to the frequent head motion during the experiment which resulted in poor quality of pupil data. The remaining 28 participants' data was submitted for further analysis (19 males, 9 females; age: 22 ± 1.6; Handedness score: 9.6 ± 1.1). To gain better control of the kinematic error size, we employed the "trajectory clamp" perturbation in which the direction of the channel was angled either -30, 0, or 30 degree with respect to the straight line between the starting position and the target (Hayashi et al., 2020) (Fig. 4B). For this experiment, we played either High or Low IADS clips. Similarly, we randomized the presenting order within each block. To ensure the presence of perturbation unpredictable, the number of 0-degree perturbation trials was doubled so the probability of perturbation is 50 %. This yielded a 2x4 design. Each condition was repeated for 10 times (20 for 0-degree perturbation condition) interspersed into 10 blocks of 60 trials. After the completion of the blocks we measured the pupillary light-reflex responses (3 sec of white/black stimulus-on

period followed by 3 sec of stimulus-off period) for 8 times (4 times for each of white or black stimulus).

Data analysis:

The data was sampled at 200 Hz. All analyses were conducted using custom-written codes with MATLAB R2015b (Mathworks, Natick, MA, USA).

For all the experiments, the data from the manipulanda (x-y positions, x-y velocities, and x-y command forces for the manipulandum handle) was smoothed with Gaussian kernel of 35 msec full width half maximum (FWHM). Reach onset was defined by the time when the tangential hand velocity first exceeded the 10% of its peak value. Reach offset was defined the tangential hand velocity first dropped below the threshold determine for reach onset of that trial. For assess movement errors on each reach we computed peak velocity error (PVE) and endpoint error (EPE). PVE was defined as perpendicular hand displacement at the time of peak tangential hand velocity with respect to a straight line from the home position to the target, and the EPE was defined as the distance between the hand and the target positions at movement offset. As an index of motor learning, we used the x-force values at the time of peak tangential hand velocity divided the peak velocity during the channel trials (Exps. 1 and 2) (Yokoi et al., 2014). For a measure of single trial learning (Exp 3), we calculated the impulse (force integrated with respect to time) for the period between -200 to 0 msec relative to the time of peak tangential hand velocity, as the amount of change in force is smaller and more noisy in single trial learning paradigm (Hayashi et al., 2020). For Exp 1, the PVE and the learning index were sign-flipped for the participants who learned CCW force field before taking group-average.

For all the experiments, the eye data (x-y point-of-gaze position and pupil diameter) was analysed in the following way. First, we discarded 100ms of point-of-gaze position and pupil data prior to a blink event and 150ms of point-of-gaze position and pupil data after a blink event. We then interpolated over the discarded data using a piecewise cubic Hermite interpolating polynomial (a Matlab *interp1.m* function with 'pchip' option). After the interpolation, the eye position data was smoothed with Gaussian kernel with 35 msec FWHM. For the detection of saccade events, the unsmoothed eye position data was also smoothed with the second-order Savitzky-Golay filter with 55 msec frame length (*sgolayfilt.m* function). The eye velocities were then calculated by numerical derivative of the eye positions (*diff.m* function). We employed 30 degree/sec for a velocity threshold for

saccade detection. As the eye positions were defined in the display coordinate, we transformed them into visual angle coordinate before the saccade detection.

To remove high-frequency noise, the pupil data was smoothed with Gaussian kernel with 235 msec FWHM. Importantly, we individually normalized the pupil diameter data relative to the minima and maxima of the pupil diameter data measured during the light reflex trials. We also calculated the pupil dilation velocity through the numerical derivative (*diff.m* function) of the normalized pupil diameter data. The trial-by-trial summary of these pupil-related variables were defined as follows. The baseline pupil diameter was defined as the average pupil diameter during the waiting period before the onset of go cue. The mean pupil dilation velocity at each trial was defined as the average of pupil dilation velocity during the period from 375 ± 95 to 650 ± 100 msec from the movement onset (mean ± SD across experiments). These periods were defined in a post hoc manner to maximally reflect the effect of experimental manipulation (i.e., force perturbation) and roughly corresponded to the periods of $p < 0.0001$ (uncollected) for the comparison between the baseline vs. the first 3 perturbed trials (e.g., Fig. 1F,G). Pupil dilation velocity data for the trials where a saccade was detected during the movement period were excluded from the analysis. The percentage of excluded trials were 16.5 ± 12.9 (Exp 1), 15.5 ± 15.2 (Exp 2), 12.6 ± 17.3 (Exp 3A), and 10.2 ± 7.8 (Exp 3B).

Computational modeling:

To get a good reference to compare with our pupil response data, we employed the Volatile Kalman Filter (VKF) that dynamically adjusts learning rate according to the volatility estimated based on experienced prediction error on a trial-by-trial basis (Piray and Daw, 2020a). We chose this model (VKF) for its numerical stability over another similar Bayesian filter algorithm (Mathys et al., 2011) and its ability to accurately reproduce a wide range of animals' choice behavior data (Piray and Daw, 2020a). In brief, at every trial, the VKF model updates the volatility estimate (e.g., expected magnitude of force-field change, in our case) based on the squared prediction error (e.g., movement error) and current uncertainty about the estimated state (e.g., uncertainty about current force-field estimate). The learning rate is then adjusted by optimally balancing the current estimate about state uncertainty, the current volatility estimate, and the uncertainty about observations (fixed). Thus, when there is a sudden change in environment (e.g., force-field magnitude/direction, in our case) which leads to a sudden increase in error size, the volatility estimate shows a transient increase in the next

and following trials (Fig. S1D). Similarly, when there are multiple such changes in the environment (change points), the volatility estimate also increases and accumulates according to the change points' frequency (Fig. S2D).

We determined the parameters of the VKF model ($\lambda, \sigma^2, \nu_0, m_0$; learning rate for volatility, the variance of observation noise, the initial value for the volatility estimate, and initial value for the state estimate, respectively) by fitting the volatility estimates to the group-averaged baseline pupil diameter data in Exp 1 (*fminsearch.m*). Note that our purpose here is not to characterize the individual difference in pupil response pattern and motor learning behavior by individually estimating these model parameters but to characterize general features of pupil responses during a typical motor learning paradigm. We used similar parameter values for the simulation of Exp 2 (Fig. S2).

Statistical analysis:

Timeseries comparison: We assessed changes in pupil velocity timeseries between unperturbed (average of the 1st block data) vs. perturbed (average over the first 3 perturbed trials) trials by using group-wise two-sided paired *t*-tests applied at each time frame. Uncorrected p-values were reported.

Effect of repetitive change points for Exp 2: We assessed the effect of participants' repetitive experience of the change points by fitting the general linear mixed-effect model (*'fitlme.m'* function) to the individually z-scored response data around the change points (baseline pupil diameter, pupil dilation velocity). All models included fixed intercept and random intercepts for participants and groups (Exps 2A, 2B, and 2C), and fixed and random effects of the perturbation types (CW and CCW). As effects of interest, fixed and random effects for the number of the change points (#cps) were also included. To account for the trivial effect of error magnitude, we additionally included the absolute error term; for the baseline pupil data, trajectory error and endpoint error at the immediately previous trial, and for the pupil dilation velocity, trajectory error at the same trial, respectively. The importance of the random effect of #cp was assessed by the likelihood-ratio test (LRT) between the full model and an alternative model lacking the random slope term. The models were fit to the data with restricted maximum likelihood method (ReML) with random starting values. We evaluated the *p*-value of the estimated fixed-effect slope for the #cp term to test if the effect of change in #cp is statistically significant.

Tests for arousal manipulation (Exp 3): We evaluated the effect of the arousal manipulation on the single-trial motor learning as follows. First, the single trial learning and the trajectory error data were z-scored within each individual. Next, the effect of error and trial number were removed from the (z-scored) single-trial learning data by fitting the general linear mixed-effect model considering the group- and participant-level random effects similarly to above analysis. On the resultant residuals, we calculated the contrast for the manipulation effect (Null-sound vs. High-arousal for Exp 3A, and Low-arousal vs. High-arousal for Exp 3B). This yielded the z-valued estimate of the effect of arousal manipulation on the single trial learning for each participant. To test the baseline-dependent effect, we individually calculated the average baseline pupil diameter over the whole experimental session, excluding the perturbation trials which included the arousal manipulation, as a proxy for individuals' mean arousal level. If there was no relationship between the manipulation effect and the mean arousal level, the data should equally distribute across the four quadrants defined by z=0 and median arousal level (i.e., expected # of participants = 54/4). To test the deviation from this null hypothesis, we conducted the likelihood-ratio *G*-test. A standard chi-squared test gave similar result.

Data and code availability:

The data and the custom-written Matlab codes used for the analysis will be uploaded to the publicly available server upon publication. Until then, requests should be addressed to the corresponding author (ayokoi@nict.go.jp).

Tables

Table 1. Luminance value measurements for display colors.

	Exp 1	Exp 2, 3
Black	0.6	0.4
White	178	249
Gray	80	107
Green	80	106
Magenta	80	108
Pale blue	80	105

Measured by CA-100 Plus Color Analyzer (Konica Minolta Sensing Americas, Inc., NJ, USA) for Exp 1 and by SR-3AR Spectroradiometer (Topcon Technohouse, Corp., Tokyo, Japan) for Exp 2, 3. Units are in (cd/m^2).

Figure legends

Figure 1. Pupillometry during force-field motor adaptation (Experiment 1).
(A) Left: Experimental setup used in Exp 1. Right: Schematics for a single trial. Participants were required to fixate on the center of the target. **(B)** The perturbation schedule for the Exp 1. The manipulandum applied the clockwise (CW) force field unexpectedly to the participants at the second block. The gray vertical lines indicate the block boundary. The short gray vertical lines represent the force channel trial where learning was quantified. **(C)** Hand trajectories of a representative participant. The cycles shown (11, 14, 32, 33, and 52) correspond to baseline, early perturbation, late perturbation, early washout, and late washout trials, respectively. **(D)** Mean lateral hand deviation at the peak tangential handle velocity. Positive values correspond to rightward deviation. **(E)** Mean learning index measured in the channel trials (once in a cycle). Learning index was the lateral force to the channel at the time of peak velocity divided by peak velocity (i.e., viscosity). The red dots represent values for perturbation trials (D, E). **(F)** A schematic example of pupil diameter timeseries during the current reaching paradigm aligned to the reach onsets. In general, the pupil shows different baseline values and dilates during reaches. On a sudden perturbation trial (trial=p), it strongly dilates. On the following trials (trial=p+1), it shows higher baseline value (see also, arrows in panel I). **(G)** The online pupil response to the force perturbation; the first time-derivative (=velocity) of pupil diameter timeseries (F) averaged over the baseline trials (un-perturbed; black trace) or the first 5 perturbation trials at the second block (perturbed; red trace). The gray area represents s.e.m. across the participants. The two vertical dashed lines indicate the time of significant difference by running paired-t tests. The dots on the y=0 line with different colors represents the significance levels (yellow: $p<0.05$, green: $p<0.01$, blue: $p<0.001$, purple: $p<0.0001$). **(H)** Trial-by-trial changes in the pupil dilation velocity averaged between 300 to 700 msec since the movement onset. **(I)** Trial-by-trial changes in the baseline pupil diameter. The red circles represent values for perturbation trials (D,E,H,I). The arrows indicates the first two perturbation trials (upward: the first trial, downward: the second trial). Error bars correspond to s.e.m. across the participants (D,E,H,I).

Figure 2. Specific response pattern in the pupil diameter facing with multiple change points.

(A) Experimental setup for the Experiments 2 and 3. **(B)** Perturbation schedules for the Exp 2 (Exp 2A, 2B, and 2C). The green circles indicate the 'change point' trials where either magnitude (on/off) or direction (CW/CCW) of the force-field changes in the middle of the blocks (changes across the blocks are excluded). The gray vertical lines indicate the block boundary. **(C)** Time course of the pupil dilation velocity averaged over the baseline trials (un-perturbed; black trace) or the first 5 perturbation trials at the second block (perturbed; red trace), averaged over all sub-groups. The gray area represents s.e.m. across the participants. The two vertical dashed lines indicate the time of significant difference by running paired-t tests. The dots on the y=0 line with different colors represents the significance levels (yellow: $p<0.05$, green: $p<0.01$, blue: $p<0.001$, purple: $p<0.0001$). **(D, E, F)** Trial-by-trial change in the mean lateral hand deviation at the peak tangential handle velocity (D), the evoked pupil dilation velocity (E), and baseline pupil diameter (F) for the Exp 2A (first column), 2B (second column), and 2C (last column), respectively. The colors of circles (black, red, and blue) indicates the data for baseline, CW field, and CCW field, respectively. Positive values for the panel D correspond to rightward deviation. The gray vertical lines indicate the block boundary. Error bars represent s.e.m. across the participants. Trials at which the dot color changed indicate the change points.

Figure 3. Pupil responses around multiple change points.

(A) Trial-by-trial change in the z-transformed baseline pupil diameter around the first change point. Pre- and post-change point values are defined by blue and red shades, respectively. **(B)** Amplitude of change point-induced baseline pupil diameter change (difference between post- vs. pre-change point values in A) as a function of the number of change points. **(C)** Trial-by-trial change in the z-transformed mean pupil dilation velocity around the first change points. Similarly, pre- and post-change point values are defined by blue and red shades, respectively. **(D)** Amplitude of cp-induced mean pupil dilation velocity change (difference between post- vs. pre-change point values in C) as a function of the number of change points. **(E-H)** The same analysis as A-D applied to the volatility estimate (E, F) and the surprise (G, H) simulated by the Volatile Kalman Filter (VKF) model using the same perturbation schedules (Exp 2A-C). The model parameters were estimated using the data from experiment 1. Note that, in F and H, the large increase (decrease) at the second (11th) change point reflects

increase (decrease) in movement error due to the first force reversal (return to Null field) (see Fig. 2B, D). The error bars indicate s.e.m. across the participants.

Figure 4. Arousal manipulation during single-trial motor learning paradigm.
(A) An exemplar part of the perturbation schedule for Exp 3A and 3B. Perturbation trials with channel trials immediately before and after (B,C) were separated with trials with Null perturbation. **(B)** The single-trial learning was quantified by measuring the lateral force output by participants immediately before and after the perturbation trial by using channel trials. For the Exp 3A, either null, CW, or CCW velocity-dependent force-field was applied, similarly to Exps 1 and 2 (left). For the Exp 3B, either CW (30 degree), CCW (-30 degree) or straight (0 degree) trajectory clamp was applied as perturbation to get better control of kinematic errors (right). Arousal manipulation was applied on the selected perturbation trials in randomized order. **(C)** Averaged time course of baseline pupil diameter change during pre-movement period for Null-sound (black), Low-arousal sound (magenta), and High-arousal sound (gold yellow). The presentation of IADS2 auditory stimuli (Bradley and Lang, 2007) started at time=0 and lasted for 6 seconds. For Null-sound condition, pre-movement period was set to 1500-2000 msec. **(D)** Averaged time series of single-trial learning (difference between force output at channel 2 vs. channel 1 trials, red: CW, black: Null, blue: CCW) for Exp 3A (left) and 3B (right). Trial-by-trial single-trial learning was assessed as the force impulse integrated over the time window of [-200 ~ 0 msec] relative to the time of peak velocity (the green shades). **(E)** Individuals' single-trial learning values plotted over corresponding individuals' mean arousal (average of baseline pupil diameter for non-perturbation trials normalized by individuals' pupillary light reflex amplitude, see Fig. S3) averaged over all conditions (1st column), Null / Low conditions (2nd column), or High condition (3rd column). The gray dots represent individual participants' data. The black dots represent the median calculated over the 10, 25, 50, 75, and 90 percentiles of mean arousal. The error bars represent the median absolute deviation. Comparison between High vs. Null/Low averages is shown on the 4th column for visualization purpose.

Figure 5. The baseline-dependent effect of arousal manipulation on single-trial motor learning.

The effect of arousal manipulation on single-trial learning was individually assessed by first z-scoring the single-trial learning (e.g., Fig. 4E) and then comparing between High vs. Null/Low conditions. The resultant z-scored delta-learning values of the individual participants were plotted over their mean arousal (black dots: participants for Exp 3A, white dots: participants for Exp 3B). The kernel density estimates of overall observations (x-y pairs) indicates that the arousal effects on single-trial motor learning are significantly more concentrated in the second and the fourth quadrants than chance ($G = 7.67$, $p = 0.0056$), indicating positive/negative effect for participants with lower/higher mean arousal level. The horizontal and vertical lines represent zero effect and median of mean arousal, respectively. The histograms for delta-learning (upper left) and mean arousal (lower right, the dashed line denotes the median arousal value) are also shown.

Figure 6. Hypothetical relationship between hierarchical state estimation/transition and environmental uncertainty/surprise reflected in pupil diameter.

(A) A likely hierarchical representation of task and daily life states for the current experiment. **(B)** A possible time-course of pupil diameter (and subjective environmental uncertainty). The degree of elicited uncertainty (pupil diameter) is proportional to (subjective) "across-hierarchy" distance between the two states (vertical double arrows on the right of panel A.

References

Alamia A, VanRullen R, Pasqualotto E, Mouraux A, Zenon A (2019) Pupil-linked arousal responds to unconscious surprisal. J Neurosci 39:5369–5376.

Albert ST, Jang J, Sheahan HR, Teunissen L, Vandevoorde K, Herzfeld DJ, Shadmehr R (2021) An implicit memory of errors limits human sensorimotor adaptation. Nat Hum Behav Available at: https://doi.org/10.1038/s41562-020-01036-x.

Antony JW, Hartshorne TH, Pomeroy K, Gureckis TM, Hasson U, McDougle SD, Norman KA (2020) Behavioral, Physiological, and Neural Signatures of Surprise during Naturalistic Sports Viewing. Neuron:1–14 Available at: https://doi.org/10.1016/j.neuron.2020.10.029.

Aston-Jones G, Cohen JD (2005) An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. Annu Rev Neurosci 28:403–450.

Atzori M, Cuevas-Olguin R, Esquivel-Rendon E, Garcia-Oscos F, Salgado-Delgado RC, Saderi N, Miranda-Morales M, Treviño M, Pineda JC, Salgado H (2016) Locus ceruleus norepinephrine release: A central regulator of cns spatio-temporal activation? Front Synaptic Neurosci 8.

Baddeley RJ, Ingram H a, Miall RC (2003) System identification applied to a visuomotor task: near-optimal human performance in a noisy changing task. J Neurosci 23:3066–3075 Available at: http://www.ncbi.nlm.nih.gov/pubmed/12684493.

Basile AS, Dunwiddie T V. (1984) Norepinephrine elicits both excitatory and responses from Purkinje cells in the in vitro rat cerebellar slice. Brain Res 296:15–25.

Behrens TEJ, Woolrich MW, Walton ME, Rushworth MFS (2007) Learning the value of information in an uncertain world. Nat Neurosci 10:1214–1221.

Bouret S, Duvel A, Onat S, Sara SJ, Curie M (2003) Phasic Activation of Locus Ceruleus Neurons by the Central Nucleus of the Amygdala. J Neurosci 23:3491–3497 Available at: http://www.jneurosci.org/content/23/8/3491.short.

Bradley MM, Lang PJ, Margaret M, Peter J (2007) The International Affective Digitized Sounds Affective Ratings of Sounds and Instruction Manual.

Burge J, Ernst MO, Banks MS (2008) The statistical determinants of adaptation rate in human reaching. J Vis 8:1–19.

Carey MR, Regehr WG (2009) Noradrenergic Control of Associative Synaptic Plasticity by Selective Modulation of Instructive Signals. Neuron 62:112–122.

Cartford MC, Samec A, Fister M, Bickford PC (2004) Cerebellar norepinephrine modulates learning of delay classical eyeblink conditioning : Evidence for post-synaptic signaling via PKA. 6:732–737.

Clewett D, Gasser C, Davachi L (2020) Pupil-linked arousal signals track the temporal organization of events in memory. Nat Commun 11:1–14 Available at: http://dx.doi.org/10.1038/s41467-020-17851-9.

Collins AGE, Frank MJ (2013) Cognitive control over learning: Creating, clustering, and generalizing task-set structure. Psychol Rev 120:190–229.

Coltman XSK, Gribble PL (2020) Time course of changes in the long-latency feedback response parallels the fast process of short-term motor adaptation. J Neurophysiol 124:388–399.

Corneil BD, Munoz DP (2014) Review Overt Responses during Covert Orienting. Neuron 82:1230–1243 Available at: http://dx.doi.org/10.1016/j.neuron.2014.05.040.

de Berker AO, Rutledge RB, Mathys C, Marshall L, Cross GF, Dolan RJ, Bestmann S (2016) Computations of uncertainty mediate acute stress responses in humans. Nat Commun 7:10996 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4820542&tool=pmcentrez&rendertype=abstract.

Diederen KMJ, Schultz W (2015) Scaling prediction errors to reward variability benefits error-driven learning in humans. J Neurophysiol 114:1628–1640.

Diedrichsen J, Hashambhoy Y, Rane T, Shadmehr R (2005) Neural correlates of reach errors. J Neurosci 25:9919–9931 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1479774&tool=pmcentrez&rendertype=abstract [Accessed November 11, 2013].

Dimitriou M (2016) Enhanced Muscle Afferent Signals during Motor Learning in Humans. Curr Biol 26:1062–1068 Available at: http://dx.doi.org/10.1016/j.cub.2016.02.030.

Doya K (2002) Metalearning and neuromodulation. Neural Networks 15:495–506.

Filipowicz ALS, Glaze CM, Kable JW, Gold J (2020) Pupil diameter encodes the idiosyncratic, cognitive complexity of belief updating. Elife 9:1–23.

Fiser J, Berkes P, Orbán G, Lengyel M (2010) Statistically optimal perception and learning: from behavior to neural representations. Trends Cogn Sci 14:119–130.

Franklin DW, Burdet E, Tee KP, Osu R, Chew C, Milner TE, Kawato M (2008) CNS Learns Stable , Accurate , and Efficient Movements Using a Simple Algorithm. 28:11165–

11173.

Galeotti N, Bartolini A, Ghelardini C (2004) Alpha-2 agonist-induced memory impairment is mediated by the alpha-2A-adrenoceptor subtype. Behav Brain Res 153:409–417.

Gonzalez Castro LN, Hadjiosif AM, Hemphill M a, Smith M a (2014a) Environmental consistency determines the rate of motor adaptation. Curr Biol 24:1050–1061 Available at: http://www.ncbi.nlm.nih.gov/pubmed/24794296 [Accessed August 31, 2014].

Gonzalez Castro LN, Hadjiosif AM, Hemphill MA, Smith MA (2014b) Environmental consistency determines the rate of motor adaptation. Curr Biol 24:1050–1061 Available at: http://dx.doi.org/10.1016/j.cub.2014.03.049.

Grossman CD, Bari BA, Cohen JY (2020) Serotonin neurons modulate learning rate through uncertainty. bioRxiv:1 Available at: https://doi.org/10.1101/2020.10.24.353508.

Hayashi T, Kato Y, Nozaki D (2020) Divisively normalized integration of multisensory error information develops motor memories specific to vision and proprioception. J Neurosci 40:1560–1570.

Heald JB, Lengyel M, Wolpert DM (2020) Contextual inference underlies the learning of sensorimotor repertoires. Mlmc 2:3–4.

Heron C, Gould TJ, Bickford P (1996) Acquisition of a runway motor learning task is impaired by a beta adrenergic antagonist in F344 rats. Behav Brain Res 78:235–241.

Herzfeld DJ, Vaswani P a, Marko MK, Shadmehr R (2014) A memory of errors in sensorimotor learning. Science (80-) 345:1349–1353 Available at: http://www.ncbi.nlm.nih.gov/pubmed/25123484.

Huang VS, Shadmehr R (2009) Persistence of motor memories reflects statistics of the learning event. J Neurophysiol 102:931–940.

Iigaya K (2016) Adaptive learning and decision-making under uncertainty by metaplastic synapses guided by a surprise detection system. Elife 5:1–22.

Imamizu H, Miyauchi S, Tamada T, Sasaki Y, Takino R, Pütz B, Yoshioka T, Kawato M (2000) Human cerebellar activity reflecting an acquired internal model of a new tool. Nature 403:192–195.

Inoue M, Uchimura M, Kitazawa S (2016) Error Signals in Motor Cortices Drive Adaptation in Article Error Signals in Motor Cortices Drive Adaptation in Reaching. Neuron 90:1–13 Available at: http://dx.doi.org/10.1016/j.neuron.2016.04.029.

Jepma M, Brown SBRE, Murphy PR, Koelewijn SC, de Vries B, van den Maagdenberg AM, Nieuwenhuis S (2018) Noradrenergic and Cholinergic Modulation of Belief Updating. J

Cogn Neurosci 30:1803–1820.

Jepma M, Murphy PR, Nassar MR, Rangel-Gomez M, Meeter M, Nieuwenhuis S (2016) Catecholaminergic Regulation of Learning Rate in a Dynamic Environment. PLoS Comput Biol 12:1–24 Available at: http://dx.doi.org/10.1371/journal.pcbi.1005171.

Jepma M, Nieuwenhuis S (2011) Pupil Diameter Predicts Changes in the Exploration – Exploitation Trade-off : Evidence for the Adaptive Gain Theory. :1587–1596.

Joshi S, Gold JI (2020) Pupil Size as a Window on Neural Substrates of Cognition. Trends Cogn Sci 24:466–480 Available at: https://doi.org/10.1016/j.tics.2020.03.005.

Joshi S, Li Y, Kalwani RM, Gold JI (2016) Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus , Colliculi , and Cingulate Cortex Article Relationships between Pupil Diameter and Neuronal Activity in the Locus Coeruleus , Colliculi , and Cingulate Cortex. Neuron 89:221–234 Available at: http://dx.doi.org/10.1016/j.neuron.2015.11.028.

Kafkas A (2021) Encoding-linked pupil response is modulated by expected and unexpected novelty : Implications for memory formation and neurotransmission. 44:1–23.

Kalman RE (1960) A new approach to linear filtering and prediction problems. J Fluids Eng Trans ASME 82.

Kitazawa S, Kimura T, Yin PB (1998) Cerebellar complex spikes encode both destinations and errors in arm movements. Nature 392:494–497.

Kloosterman NA, Meindertsma T, van Loon AM, Lamme VAF, Bonneh YS, Donner TH (2015) Pupil size tracks perceptual content and surprise. Eur J Neurosci 41:1068–1078.

Lackner JR, Dizio P (1994) Rapid Adaptation to Coriolos Force Perturbations of Arm Trajectory. J Neurophysiol 72:299–313.

Lavín C, San Martín R, Rosales Jubal E (2014) Pupil dilation signals uncertainty and surprise in a learning gambling task. Front Behav Neurosci 7:218 Available at: http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=3879532&tool=pmcentrez&rendertype=abstract [Accessed September 24, 2014].

Mathys C, Daunizeau J, Friston KJ, Stephan KE (2011) A Bayesian foundation for individual learning under uncertainty. Front Hum Neurosci 5:9.

Mathys CD, Lomakina EI, Daunizeau J, Iglesias S, Brodersen KH, Friston KJ, Stephan KE (2014) Uncertainty in perception and the Hierarchical Gaussian filter. Front Hum Neurosci 8:1–24.

Muller TH, Mars RB, Behrens TE, O'Reilly JX (2019) Control of entropy in neural models

of environmental state. Elife 8:1–30.

Murphy PR, Van Moort ML, Nieuwenhuis S (2016) The pupillary orienting response predicts adaptive behavioral adjustment after errors. PLoS One 11:1–11 Available at: http://dx.doi.org/10.1371/journal.pone.0151763.

Nassar MR, Rumsey KM, Wilson RC, Parikh K, Heasly B, Gold JI (2012) Rational regulation of learning dynamics by pupil-linked arousal systems. Nat Neurosci 15:1040–1046 Available at: http://dx.doi.org/10.1038/nn.3130.

Nassar MR, Wilson RC, Heasly B, Gold JI (2010) An approximately Bayesian delta-rule model explains the dynamics of belief updating in a changing environment. J Neurosci 30:12366–12378.

Nicholls MER, Thomas NA, Loetscher T, Grimshaw GM (2013) The flinders handedness survey (FLANDERS): A brief measure of skilled hand preference. Cortex 49:2914–2926 Available at: http://dx.doi.org/10.1016/j.cortex.2013.02.002.

Okubo M, Suzuki H, Nicholls MER (2014) A Japanese version of the FLANDERS handedness questionnaire. Japanese J Psychol 85:474–481.

Piray P, Daw ND (2020a) A simple model for learning in volatile environments. PLoS Comput Biol 16:1–26 Available at: http://dx.doi.org/10.1371/journal.pcbi.1007963.

Piray P, Daw ND (2020b) Unpredictability vs . volatility and the control of learning. :1–29.

Pouget A, Beck JM, Ma WJ, Latham PE (2013) Probabilistic brains: knowns and unknowns. Nat Neurosci 16:1170–1178 Available at: http://www.ncbi.nlm.nih.gov/pubmed/23955561.

Preuschoff K, Bossaerts P (2007) Adding Prediction Risk to the Theory. 146:135–146.

Preuschoff K, Marius B, Einhäuser W (2011) Pupil dilation signals surprise : evidence for noradrenaline ' s role in decision making. 5:1–12.

Rajkowski J, Kubiak P, Aston-Jones G (1993) Correlations between locus coeruleus (LC) neural activity, pupil diameter and behavior in monkey support a role of LC in attention. Soc Neurosci Abstr 19.

Ramos BP, Arnsten AFT (2007) Adrenergic pharmacology and cognition : Focus on the prefrontal cortex. 113:523–536.

Rao RPN (1999) An optimal estimation approach to visual perception and learning. 39:1963–1989.

Reichenbach A, Franklin DW, Zatka-Haas P, Diedrichsen J (2014) A dedicated binding mechanism for the visual control of movement. Curr Biol 24:780–785.

Reimer J, Mcginley MJ, Liu Y, Rodenkirch C, Wang Q, Mccormick DA, Tolias AS (2016) Pupil fluctuations track rapid changes in adrenergic and cholinergic activity in cortex. Nat Commun 7 Available at: http://dx.doi.org/10.1038/ncomms13289.

Rouhani N, Norman KA, Niv Y, Bornstein AM (2020) Reward prediction errors create event boundaries in memory. Cognition 203:104269 Available at: https://doi.org/10.1016/j.cognition.2020.104269.

Sara SJ (2009) The locus coeruleus and noradrenergic modulation of cognition. Nat Rev Neurosci 10:211–223 Available at: http://www.ncbi.nlm.nih.gov/pubmed/19190638 [Accessed July 9, 2014].

Sarafyazd M, Jazayeri M (2019) Hierarchical reasoning by neural circuits in the frontal cortex. Science (80-) 364.

Scheidt RA, Reinkensmeyer DJ, Conditt MA, Rymer WZE V, Mussa-ivaldi FA (2000) Persistence of Motor Adaptation During Constrained , Multi-Joint , Arm Movements. :853–862.

Schween R, McDougle SD, Hegele M, Taylor JA (2020) Assessing explicit strategies in force field adaptation. J Neurophysiol 123:1552–1565.

Shadmehr R, Brashers-Krug T (1997) Functional stages in the formation of human long-term motor memory. J Neurosci 17:409–419.

Shadmehr R, Mussa-ivaldi FA (1994) Adaptive Representation of Dynamics during Learning of a Motor Task. J Neurosci 14:3208–3224.

Shibata K, Yamagishi N, Ishii S, Kawato M (2009) Boosting perceptual learning by fake feedback. Vision Res 49.

Soltani A, Izquierdo A (2019) Adaptive learning under expected and unexpected uncertainty. Nat Rev Neurosci 20 Available at: http://dx.doi.org/10.1038/s41583-019-0180-y.

Steinhauer SR, Siegle GJ, Condray R, Pless M (2004) Sympathetic and parasympathetic innervation of pupillary dilation during sustained processing. 52:77–86.

Strange BA, Hurlemann R, Dolan RJ (2003) An emotion-induced retrograde amnesia in humans is amygdala- and β-adrenergic-dependent. Proc Natl Acad Sci U S A 100:13626–13631.

Szabadi E (2013) Functional neuroanatomy of the central noradrenergic system.

Tan HS, Van JN, Collewijn H, P O (1991) Effects of a-noradrenergic substances on the optokinetic and vestibulo-ocular responses in the rabbit : a study with systemic and intrafloccular injections. 562:207–215.

Taylor JA, Ivry RB (2011) Flexible cognitive strategies during motor learning. PLoS Comput Biol 7.

Taylor JA, Krakauer JW, Ivry RB (2014) Explicit and implicit contributions to learning in a sensorimotor adaptation task. J Neurosci 34.

van Neerven J, Pompeiano O, Collewijn H, van der Steen J (1990) Injections of β-noradrenergic substances in the flocculus of rabbits affect adaptation of the VOR gain. Exp Brain Res 79:249–260.

Vincent P, Parr T, Benrimoh D, Friston KJ (2019) With an eye on uncertainty : Modelling pupillary responses to environmental volatility. :1–22.

Wakita R, Tanabe S, Tabei K, Funaki A, Inoshita T, Hirano T (2017) Differential regulations of vestibulo-ocular reflex and optokinetic response by β - and α 2-adrenergic receptors in the cerebellar flocculus. :1–12.

Watson M, Mcelligott JG (1984) Cerebellar Norepinephrine Depletion and Impaired Acquisition of Specific Locomotor Tasks in Rats. 296:129–138.

Wei K, Glaser JI, Deng L, Thompson CK, Stevenson IH, Wang Q, Hornby TG, Heckman CJ, Kording KP (2014) Serotonin affects movement gain control in the spinal cord. J Neurosci 34:12690–12700.

Wei K, Körding K (2009) Relevance of error: What drives motor adaptation? J Neurophysiol 101:655–664.

Wei K, Körding K (2010) Uncertainty of feedback and state estimation determines the speed of motor adaptation. 4:1–9.

Wolpert DM, Ghahramani Z, Jordan M (1995) An Internal Model for Sensorimotor Integration. Science (80-) 269:1880–1882.

Yerkes RM, Dodson JD (1908) The relation of strength of stimulus to rapidity of habit-formation. J Comp Neurol Psychol 18.

Yokoi A, Hirashima M, Nozaki D (2014) Lateralized sensitivity of motor memories to the kinematics of the opposite arm reveals functional specialization during bimanual actions. J Neurosci 34:9141–9151.

Yu AJ, Dayan P (2005) Uncertainty , Neuromodulation , and Attention. 46:681–692.

Zacks JM, Kurby CA, Eisenberg ML, Haroutunian N (2011) Prediction error associated with the perceptual segmentation of naturalistic events. J Cogn Neurosci 23:4057–4066.

Supplementary Figures

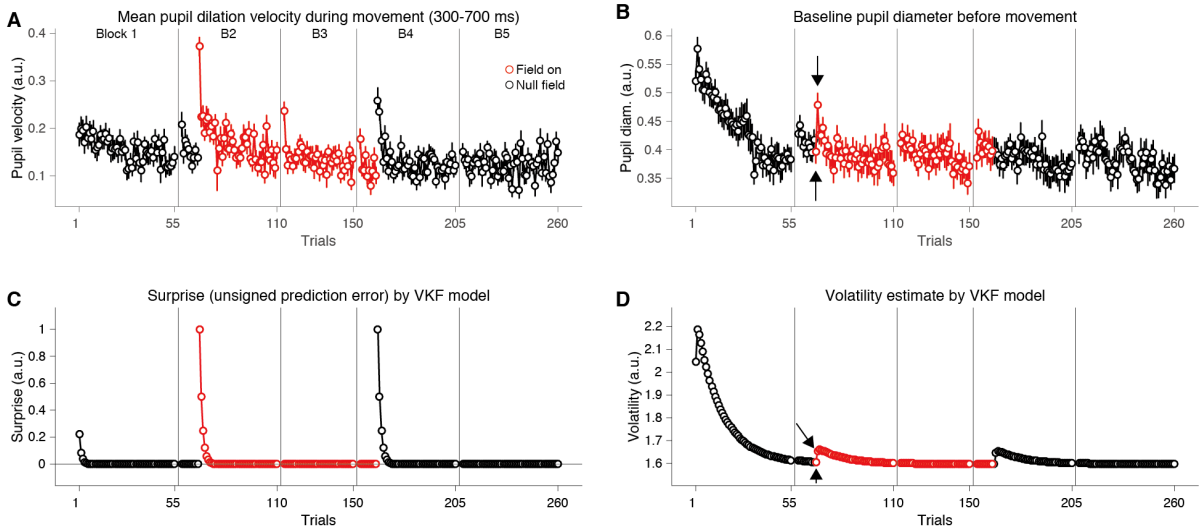

Figure S1. Volatile Kalman Filter (VKF) model captured qualitative features of online and offline pupil responses to force-field perturbation (related to Fig. 1). (A) *Online pupil response (mean pupil dilation velocity during movement).* ***(B)*** *Offline pupil response (pre-movement baseline pupil diameter).* ***(C)*** *Surprise (unsigned prediction error) and* ***(D)*** *volatility signal simulated by VKF model* (Piray and Daw, 2020a). *Figure format is same as Fig. 1H and I. Pupil dilation velocity (A) and surprise (C) shared a similar trial-by-trial change pattern; a large increase followed by gradual decrease to baseline level both at the introduction and removal of force-field. While baseline pupil diameter (B) and volatility estimate (D) showed similar pattern at the introduction of force-field (indicated by arrows), baseline pupil diameter showed little or no increase at the removal of the force-field (161-th trial). Note that our purpose here is not to characterize individual difference in pupil response pattern and motor learning behavior by individually estimating model parameters, but to characterize general features of pupil responses during a typical motor learning paradigm. Also note that the baseline pupil diameter increased at the beginning of each block, implying increased subjective uncertainty about environment after the short rest period, yet this was not modelled in the original VKF model. Model parameters;* $\lambda = 0.1933, \sigma^2 = 3.1942, \nu_0 = 2.0459, m_0 = -0.2198$.

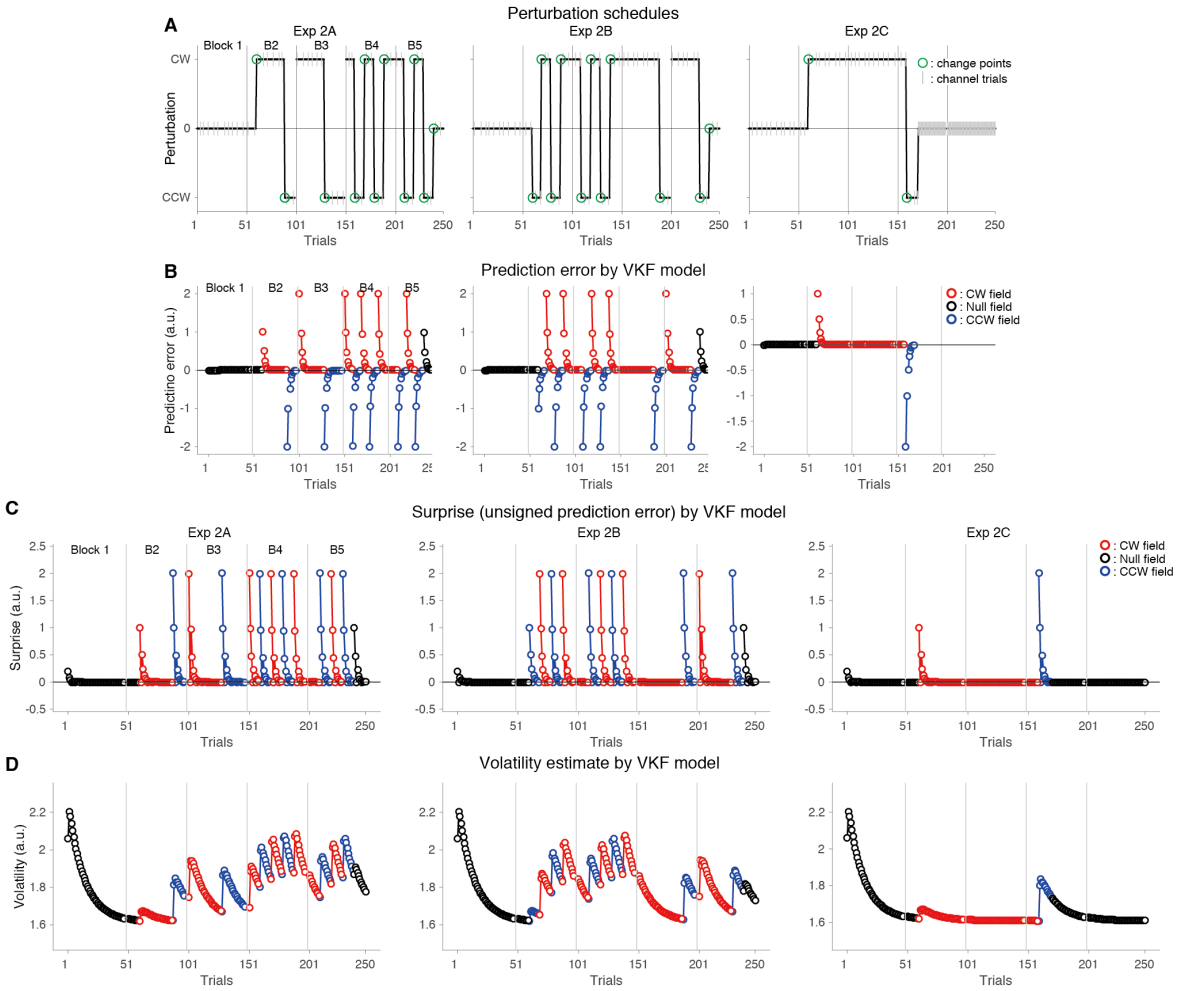

Figure S2. Simulation of force-field scheduled for Exp 2 using VKF model (related to Fig. 2). (A) *Force-field schedules for sub-groups (Exp 2A, 2B, and 2C).* ***(B)*** *Prediction error about the mean.* ***(C)*** *Surprise (unsigned prediction error).* ***(D)*** *Volatility estimates. Figure format is the same as Fig. 2. The reversal of force-field direction (second and later change points) induced larger prediction error (B) and hence surprise (unsigned prediction error; panel C) which resulted in transient increase in volatility (D). In other word, the size of surprise and volatility change closely follows the size of experienced prediction error. More frequent change points (later trials in Exp 2A or earlier trials in Exp 2B) let to accumulation of volatility. Model parameters;* $\lambda = 0.1934, \sigma^2 = 3.2192, \nu_0 = 2.0589, m_0 = -0.20$.

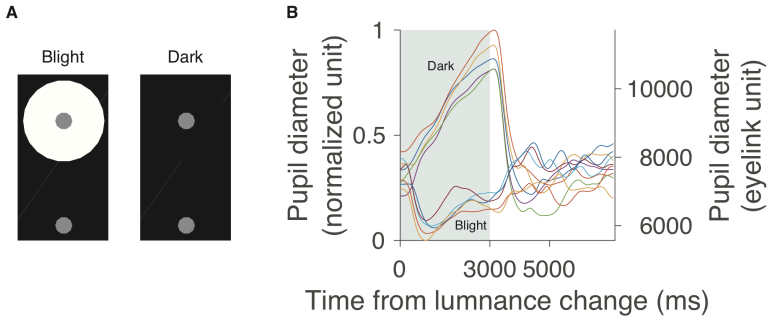

Figure S3. Within individual normalization of pupil diameter using pupillary light reflex (PLR). (A) *Luminance changes were induced by changing background color (white or black).* ***(B)*** *An exemplar pupillary response from one participant. The gray shaded area indicate the duration of luminance change (3000 msec). The arbitrary unit in the eyetracker (right y axis) was normalized into the range of individuals' pupil modulation size (left y axis). Traces represent the responses in the single light reflex trial. For more detail, see Methods.*