

<https://doi.org/10.1038/s44271-024-00123-3>

# Model-free decision-making underlies motor errors in rapid sequential movements under threat

Check for updates

Pranav Sankhe <sup>1,2</sup> & Masahiko Haruno <sup>1,3</sup>

Our movements, especially sequential ones, are usually goal-directed, i.e., coupled with task-level goals. Consequently, cognitive strategies for decision-making and motor performance are likely to influence each other. However, evidence linking decision-making strategies and motor performance remains elusive. Here, we designed a modified version of the two-step task, named the two-step sequential movement task, where participants had to conduct rapid sequential finger movements to obtain rewards ( $n = 40$ ). In the shock session, participants received an electrical shock if they made an erroneous or slow movement, while in the no-shock session, they only received zero reward. We found that participants who prioritised model-free decision-making committed more motor errors in the presence of the shock stimulus (shock sessions) than those who prioritised model-based decision-making. Using a mediation analysis, we also revealed a strong link between the balance of the model-based and the model-free learning strategies and sequential movement performances. These results suggested that model-free decision-making produces more motor errors than model-based decision-making in rapid sequential movements under the threat of stressful stimuli.

In real life, the actions we carry out are often oriented towards achieving a certain goal. This relationship indicates that the decisions we make and the actions we execute to carry out those decisions are tightly coupled<sup>1,2</sup>. This coupling is particularly important when one has to perform under pressure or stress, such as in sport and music competitions, where an individual must make several decisions and perform precise actions during a short period of time. In addition, motor tasks in such real-life situations involve a sequence of actions where individual motor movements are assembled in a structure. However, research in motor control has mainly focused on single-shot actions, such as reaching and generating grip force<sup>3,4</sup>, and has not considered decision-making strategies<sup>5–8</sup>.

Previous work in computational neuroscience has suggested two distinct mechanisms are employed to learn the value of actions from outcomes<sup>9–11</sup>. The ‘model-based’ system builds an internal model of the environment based on state-action transitions to prospectively compute the best course of actions<sup>12–14</sup>. In contrast, the ‘model-free’ system relies solely on accumulating past experiences for learning and utilises estimated action values to make decisions<sup>15–17</sup>. Model-based learning is flexible to changes in the environment, as it can update its ‘world model’, but it is also computationally costly owing to the prospective computation of all possible action courses. On the other hand, model-free learning is inflexible to changes in

outcomes and state transitions but is computationally easy, as one can simply choose the action with the highest value.

Multiple studies have shown evidence for the coexistence of both model-based and model-free learning mechanisms in human and animal behaviour<sup>18–25</sup>. This naturally leads to the question: How does the brain manage the trade-off between these two mechanisms? Past research has suggested that an arbitration system in the brain allocates control to these two systems<sup>9,26–29</sup>. Lee and colleagues reported that the brain arbitrates between the two systems and uses them in conjunction based on the reliability of their predictions<sup>30</sup>. A study by Kim et al. showed that individuals tend to prioritise model-based learning strategies with increasing task complexity but resort to model-free learning when task uncertainty and complexity are both high<sup>31</sup>. Similarly, Lockwood and team found that individuals relied more on model-free learning strategies when the task involved avoiding harm to others<sup>32</sup>. Otto and colleagues showed that cold-pressure stress can decrease model-based decisions with working memory capacity, contributing to the detrimental impact of stress on decision-making<sup>33</sup>.

We are interested in whether the balance between model-based and model-free decision-making influences the motor performance of actions carried out to execute decisions, in particular when participants are under the threat of stressful stimuli<sup>34–38</sup>. For instance, imagine driving a car on an

<sup>1</sup>Center for Information and Neural Networks, NICT, 1-4 Yamadaoka, Suita, Osaka, 565-0871, Japan. <sup>2</sup>Institute of Cognitive Neuroscience, University College London, 17-19 Queen Square, London, WC1N 3AZ, UK. <sup>3</sup>Graduate School of Frontier Biosciences, Osaka University, 1-3 Yamadaoka, Suita, Osaka, 565-0871, Japan. e-mail: [pranav.sankhe.22@ucl.ac.uk](mailto:pranav.sankhe.22@ucl.ac.uk); [mharuno@nict.go.jp](mailto:mharuno@nict.go.jp)

icy road. One must constantly make decisions about what information to use to decide when to take a turn and when to accelerate or apply brakes so that the destination is reached without incident. Some drivers may actively keep track of state transitions and predict future states, i.e. the model-based strategy. Alternatively, other drivers may rely more on their previous experience of driving in similar conditions, i.e. the model-free strategy. Importantly, each decision strategy can be coupled with different sequences of motor actions.

Our hypothesis in this study is that model-free/model-based decision-making strategies have crucial influence on sequential motor performances, particularly under the threat of stressful stimuli. To test this hypothesis, we modified the two-step task, which was originally developed to dissociate model-based and model-free decision-making<sup>14</sup>. Our two-step sequential movement task requires participants to make choices using rapid sequential movements rather than a single movement under time pressure. Participants perform the task in two types of sessions. In the no-shock session, a motor error results only in zero reward, while in the shock session, a motor error is penalised by an electric shock and zero reward. Participants experienced two no-shock sessions and a subsequent shock session. We compared the second no-shock session and the shock session in our main analysis.

**Methods**

**Participants**

In total, 40 participants (24 male participants, 16 female participants, mean age = 21.8 years, s.d. 2.09) were recruited from the Osaka University community. Participants were asked to self-report their sex. We did not collect the participants' race/ethnicity data. The participants gave informed consent for participating in the experiments, and the experiments were approved by the ethics committee at the National Institute of Information and Communications Technology (NICT), Japan. They were paid ¥3000 on top of the money they earned while performing the task. All experimental procedures for each participant were finished on a single day. This study was not preregistered.

**Learning of two motor sequences**

In our study, both Sequence-A and Sequence-B comprised five button presses, which were executed using the right index, middle, and ring fingers

on a 3-button keypad (Fig. 1B). The buttons on the keypad were labelled as '1', '2', and '3' from left to right. In Sequence-A, participants were instructed to press the buttons in the order 1-3-2-1-3 (which corresponded to the respective button positions on the keypad). Similarly, in Sequence-B, the button presses followed the order 2-3-1-3-2. Prior to the main experiment, participants underwent extensive training on these two motor sequences. During training, participants performed five blocks, with each block consisting of 40 trials. Initially, participants approached the sequence execution cautiously, pressing each button individually. However, as the training progressed, their actions became more fluid, and by the end of the training period, the learning of the motor sequences seemed to have plateaued (Supplementary Fig. 1A, B).

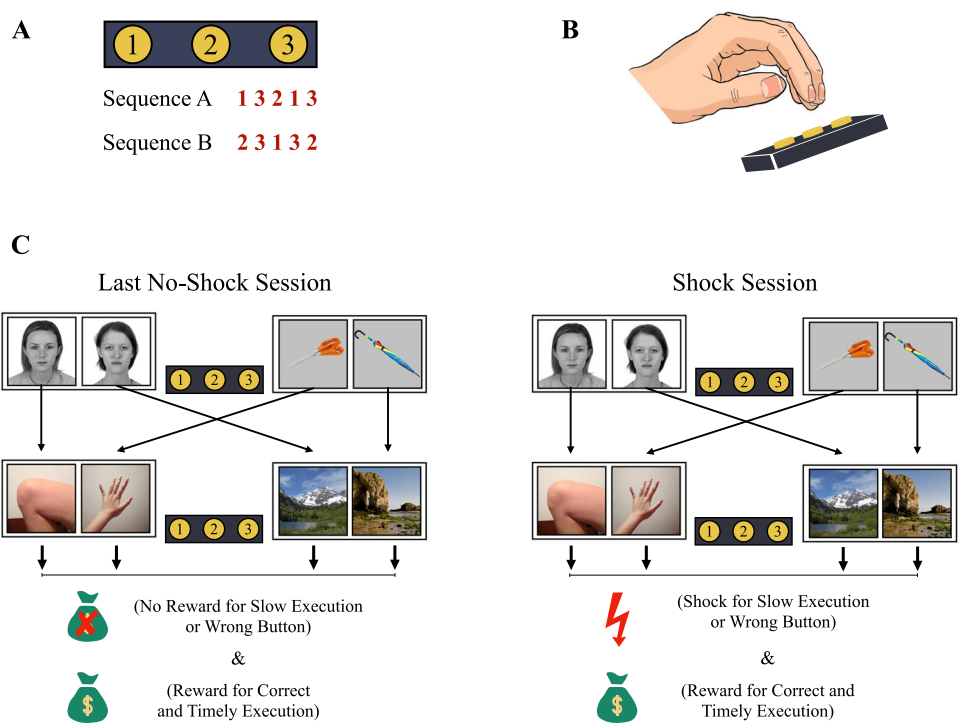
**Electric shock calibration and GSR setup**

Electrodes from an electrical stimulator were placed on the participants' left forearms for the experiment. We administered a brief Gaussian burst of electrical current (25 ms) as a shock stimulus. Participants were instructed to close their eyes, and, starting from a minimal level, we gradually increased the shock's intensity until the participants reported feeling it on their skin. Once the perceptible threshold was determined, participants were asked to indicate the shock level at which they would feel anxious and scared. To converge on this threshold, they were instructed to imagine a level that they would not mind experiencing the shock once but would find distressing if subjected to it three times in quick succession. Participants were encouraged to explore higher shock levels and then decrease the level if they felt uncomfortable, ensuring they identified a level that genuinely induced pressure. Following the calibration of the stimulator, electrodes for measuring the galvanic skin response (GSR) were attached to the participant's right index and middle fingers using BIOPAC Systems, Inc., equipment.

**Two-step sequential movement task**

Figure 1 illustrates the two-step sequential-movement task. The design of the task is based on past studies<sup>14,26</sup> and consists of two stages. In stage 1, participants encounter one of two states: 'Faces' or 'Objects'. The choices made at stage 1 deterministically decide the subsequent state at stage 2, which could be either 'Body Parts' or 'Scenes'. Notably, the available choices in the two states (Faces or Objects) in stage 1 are similar: selecting one of the

**Fig. 1 | Task design of the two-step sequential movement task.** **A** Motor sequences used to select the images at both stages in the two-step sequential movement task. Sequence-A and Sequence-B were used to select the image on the left and right side of the screen, respectively. **B** A keypad was used to enter the sequence. **C** The state transition structure from stage 1 to stage 2 was deterministic such that the two states (i.e. Faces and Objects) at stage 1 were essentially equivalent. The face and tool depicted on the left always resulted in the body part category, while the face and tool shown on the right always led to the scene category in stage 2. In the two no-shock sessions, if there were any errors in the sequence execution or if the participants timed out, no reward was awarded. By contrast, in the subsequent shock session, in such cases, along with no reward, an electric shock was delivered. We compared the second no-shock session and the shock session. The faces used in the figure are taken from the Chicago Face Database<sup>51</sup>.



tools (depicted in Fig. 1C on the right) or one of the faces (depicted in Fig. 1C on the left) always results in the same set of Scenes, while choosing the other tool or face leads to Body Parts. This equivalent structure helped us distinguish between model-based and model-free decision-making strategies, as only model-based learners can generalise their experiences across equivalent options at stage 1. Model-based learners utilise estimations of the expected outcomes for each option in stage 2 to determine their respective values in stage 1. Consequently, the impact of each second-stage outcome on stage 1 preference in subsequent trials remains the same, irrespective of whether the new trial begins with the same state as the previous one (e.g., faces followed by faces) or a different state (e.g., faces followed by objects). In contrast, a model-free learner evaluates options based solely on their past outcomes: the outcomes obtained from one starting state does not influence subsequent choices from the other starting state.

To make their choices, participants had sequence-A to select the image on their left and sequence-B to select the image on their right at both stages. The position of the images within each category varied randomly, but the mapping of sequences with left and right positions remained fixed throughout the task. We adopted this design to couple sequential motor execution with decision-making, and the two sequences (A and B) were selected so that their entropies were the same and the button positions were sufficiently distinguishable. All participants ( $N = 40$ ) were extensively trained in pressing the two sequences before the main experiment. Each option in stage 2 was rewarded with a monetary reward. The reward distribution was randomly initialised with either 30 points or 70 points for the images of both states in stage 2. In order to incentivise learning throughout the experiment, the reward values change slowly and independently according to a Gaussian random walk. The reward values underwent slow changes following a Gaussian random walk process that reflected the boundaries at 1 and 100. The random walk had a mean of 0 and a standard deviation of 20.

Participants completed a total of three sessions continually on the same day, each consisting of 120 trials, in the two-step sequential movement task. In the first session, participants had enough time to make their choice (3 seconds). In the second session, the time available to make the choice was reduced from 3 seconds to 2 seconds in a linear manner so that the participants became accustomed to the time pressure of the task. In the third and last session, the time available for a decision at both stages was 2 seconds. Notably, to induce strong pressure, participants were told that if they made an error in executing the sequence (pressing an incorrect button in a sequence) or if they were too slow in making a choice and executing it (time threshold set to 2 seconds), they would receive an electric shock on their arm. The intensity of the electrical shock was individually calibrated for each participant. The first session was intended for the participants to familiarise themselves with the task and allow us to compare the last two sessions. For the rest of the paper, we will call the second session the ‘no-shock session’ and the third session the ‘shock session’.

Participants were provided with feedback regarding the type of error made during each trial. If their trial timed out, they were notified of a ‘Late’ error at the end of the trial. If they failed to enter either sequence correctly, they were informed of a ‘Button Miss’ error. If an error occurred at stage 1 of the sequential two-step task, participants were unable to progress to stage 2. They were shown their error type, and they experienced the associated outcome, including zero points and, during the shock session, an electric shock.

All three sessions were conducted on the same day, separated by a 5-minute break. The shock and no-shock conditions were not counter-balanced because the participants were made to adjust for the time pressure that steadily decreased in the second session (the last no-shock session) and then kept fixed in the shock session to maximise stress.

We noticed that reaction times for executing sequences decreased over sessions. This might be due to both learning and the progressively stricter time limits set as sessions advanced. Additionally, the number of errors, including late and incorrect button presses, increased from the first to the second no-shock session, likely because of the stricter time limit in the latter session. However, the number of errors significantly decreased from the second no-shock session to the shock session, an effect that can be attributed

both to learning and the threat of an electrical shock. Overall, these findings suggest that sequence learning persisted across experimental sessions. Supplementary Fig. 2A and 2B illustrate the number of errors and sequence completion times across sessions.

### Computational model

We utilised a hybrid reinforcement learning model that was adapted to our task design. Choice data were fitted to this computational model for each participant; the model learns the values of actions by using a combination of model-based and model-free approaches. At stage 2 of the two-step task, learning from the outcomes is solely model-free, as there are no state transitions following the choice that could be exploited. For each of the states  $s_2$  (body parts, scenes) at stage 2, state action values  $Q_2$  ( $Q$  values) are learnt for both actions  $a_2 \in \{a_{2X}, a_{2Y}\}$  ( $a_{2X}, a_{2Y}$  refer to the two available choices at stage 2).  $Q_2$  are updated at each trial as per the following:

$$Q_2(s_{2,t}, a_{2,t}) = Q_2(s_{2,t}, a_{2,t}) + \alpha \delta_{2,t} \tag{1}$$

where  $\delta_{2,t}$  is the reward prediction error at stage 2. Since there are no subsequent stages after stage 2 in our task, the reward prediction error at stage 2 is driven by the reward as follows:

$$\delta_{2,t} = r_t - Q_2(s_{2,t}, a_{2,t}) \tag{2}$$

Here,  $\alpha$  is a free parameter representing the learning rate that modulates the effect of the prediction error in outcomes on action values.

At stage 2, the estimation of the  $Q$  values is purely model-free, because the  $Q$  values are computed based on the immediate reward. At stage 1, however, both model-based and model-free strategies contribute to the estimation of the  $Q$  values, since state transitions from stage 1 to stage 2 can also be considered in the learning. According to model-free learning (SARSA( $\lambda$ )), the action values  $Q_{MF}$  are learnt for each action  $a_1$  for each state  $s_1$  (faces, objects) at stage 1 as follows:

$$Q_{MF}(s_{1,t}, a_{1,t}) = Q_{MF}(s_{1,t}, a_{1,t}) + \alpha \delta_{1,t} + \lambda \alpha \delta_{2,t} \tag{3}$$

The reward prediction error at stage 2 ( $\delta_{2,t}$ ) is used to update the  $Q$  value at stage 1, and the size of the effect is controlled by the free parameter  $\lambda$ , which is also known as the eligibility trace parameter.

The reward prediction error at stage 1 differs from that at stage 2, as rewards are available only after a choice is made at stage 2. The prediction error at stage 1 is calculated as follows:

$$\delta_{1,t} = Q_2(s_{2,t}, a_{2,t}) - Q_{MF}(s_{1,t}, a_{1,t}) \tag{4}$$

As can be seen in the above equations, model-free learning does not consider the fact that choices are equivalent regardless of the state at stage 1. The model-free system separately learns the value of choosing actions in the face state and the tool state according to their respective outcomes. On the other hand, model-based learning uses this equivalence to compute the value of actions in a prospective manner. For the current two-step task, this means calculating for each action at stage 1 can estimate the rewards available at stage 2 based on the choice made. The model-based  $Q$  value,  $Q_{MB}$ , for each state  $s_1$  and action  $a_1$  is calculated as follows:

$$Q_{MB}(s_{1,t}, a_{1,t}) = \max_{a \in \{a_{2X}, a_{2Y}\}} Q_2(S(s_{1,t}, a_{1,t}), a) \tag{5}$$

Here,  $S(s_{1,t}, a_{1,t})$  is the state at stage 2 that results in choosing action  $a_1$  in state  $s_1$ . Given that model-based learning uses information about the state transitions in the task,  $S(s_{1,t}, a_{1,t})$  generalises between the two states (Faces, Objects) at stage 1. Thus the model-based  $Q$  value allows us to access action  $a_2$  at stage 2 in stage 1.

Finally, to compute the value of actions at stage 1, the  $Q$  values computed by the model-based approach and model-free approach are combined by a free model-based weight  $w$  ( $w = 1$  for a purely model-

based agent and  $w = 0$  for a pure model-free agent). The resultant  $Q$  value at stage 1 ( $Q1_{net}$ ) is:

$$Q1_{net}(s1_t, a1_t) = wQ_{MB}(s1_t, a1_t) + (1 - w)Q_{MF}(s1_t, a1_t) \quad (6)$$

To select the action using the  $Q$  values, we used the softmax decision rule as follows:

$$p(a_{i,t} = a|s_{i,t}) = \frac{\exp(\beta[Q1_{net}(s_{i,t}, a) + \pi.rep(a) + \rho.resp(a)])}{\sum_{\hat{a}} \exp(\beta[Q1_{net}(s_{i,t}, \hat{a}) + \pi.rep(\hat{a}) + \rho.resp(\hat{a})])} \quad (7)$$

Following the computational model used in a previous study<sup>26</sup>, in addition to the standard inverse temperature parameter  $\beta$ , we added two free parameters:  $\pi$  (choice stickiness parameter) and  $\rho$  (response stickiness). The choice stickiness parameter is multiplied by  $rep(a)$  (an indicator variable, which is 1 if stage 1 action is the same as the one of the previous trial, and zero otherwise) to capture the tendency to stick with the same choice or to switch it. Since the position of the images is not fixed across trials but randomly varied, participants may exhibit a tendency to repeat/alternate the sequences (A and B). To account for this possibility, we added the term that is a product of response stickiness ( $\rho$ ) and  $resp(a)$  (is coded 1 if the same sequence was entered as the previous trial and 0 otherwise).

We used the mfit toolbox<sup>39</sup> to fit the choice data to our reinforcement learning models and estimate the free parameters. We used a hierarchical maximum *a posteriori* estimation using group priors based on previous work<sup>39</sup>. Using both the Akaike Information Criterion (AIC) and Bayesian

Information Criterion (BIC), we checked whether including either of the stickiness parameters ( $\pi$  and  $\rho$ ) enabled better modelling of the observed behaviour. Additionally, we used AIC and BIC to compare model fits of the hybrid model (including the  $w$  parameter) as well as a pure model-based and a pure model-free model. We found that both AIC and BIC favoured the hybrid model with both stickiness parameters ( $\pi$  and  $\rho$ ) (Supplementary Tables 1 and 2). We also ran the optimisation algorithm fifteen times for each participant to avoid local optimum solutions and randomly selected initialisations for every parameter of the model.

In our reported results, we calculated parameters separately for the no-shock and shock sessions. To measure a shift in strategy, we computed the difference between the model-based weights of the shock and no-shock sessions. Finally, we have used the standard inferential statistical tests (such as t-test and correlations) with an assumption that the data distribution of parameters is normal, but this isn't formally tested.

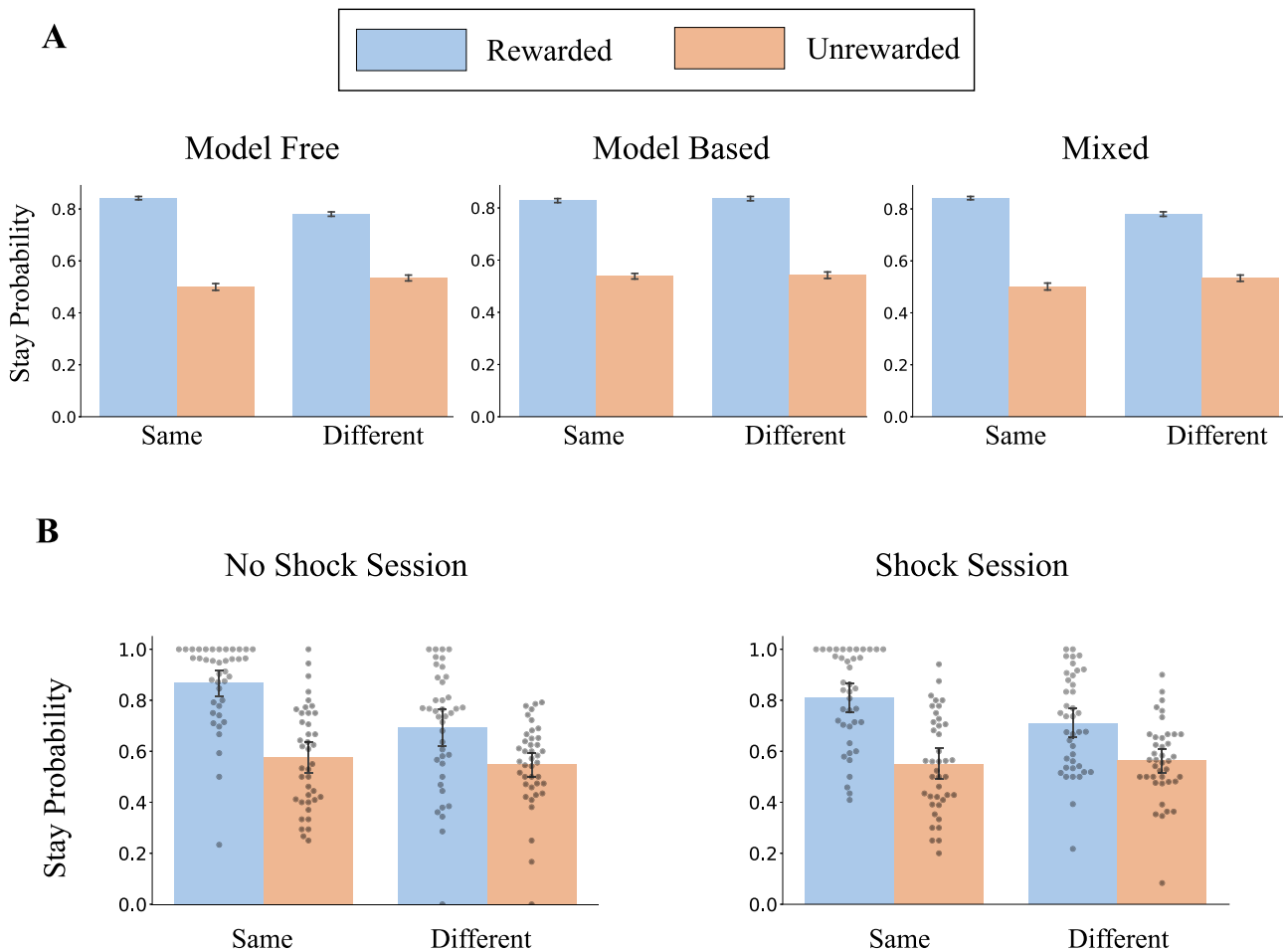
**Reporting summary**

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

**Results**

**Participants showed both model-based and model-free behaviour**

We first examined how the combination of previous rewards and the starting state affect the stay probability (choosing the same action as in the previous trial) to differentiate the model-based and model-free strategies



**Fig. 2 | Model-predicted and real behavioural data.** A Behaviours of different models conditioned on whether the previous trial led to a reward and whether the state at stage 1 was the same as the previous trial. The behaviour depicted in the plots is predicted for our task by generative reinforcement learning models.

**B** Participants' ( $n = 40$  participants) Stay vs. Switch behaviour plotted for no-shock and shock sessions shows that the participants used both model-based and model-free behaviour to make their choices. Error bars represent 95% confidence intervals.



(Fig. 2). Simulating generative reinforcement learning agents for our task, we observed that differences in model-based, model-free and mixture reinforcement learning agents were characterised by an interaction between the similarity in the start state and past reward, particularly in a decrease in the stay probability in the rewarded and different conditions (Fig. 2A). We can see that participants in our task qualitatively show a mix of model-based and model-free learning in the no-shock and shock sessions (Fig. 2B). Figure 2B also indicates that the contribution of the model-free strategy is larger in the shock session than in the no-shock session. However, a precise differentiation in arbitration between model-based and model-free strategies should be conducted by estimating computational model parameters<sup>26</sup>.

We estimated such parameters for computational models of a pure-model based agent, pure-model free agent and a hybrid learner using our data. We performed model comparisons to check which model best captured the behaviour for both the no-shock and shock sessions. Such a trial-by-trial analysis that captures individual choice preferences allowed us to test whether either model or their combination best produced the behaviours. The Akaike information criteria (AIC) across participants revealed that the hybrid model provided a significantly more accurate explanation of behaviour than a purely model-free or purely model-based learner (Supplementary Table 1). In other words, we quantitatively showed not only that a reward generalisation between equivalent starting states was significant (model-based learning) but also that a larger effect of the rewards was observed when the starting state remained the same (model-free learning).

The best-performing hybrid model included five parameters: learning rate  $\alpha$ , inverse temperature  $\beta$ , which controls the randomness of the choice selection; model-based weight  $w$ , which represents the relative balance between the model-based and model-free strategies; choice-stickiness parameter  $\pi$ , which captures the degree of stay versus switch at stage 1; and

response-stickiness parameter  $\rho$ , which represents the tendency of a participant to repeat or alternate the motor sequences (see Methods for details).

### Stressful stimuli Increased GSR

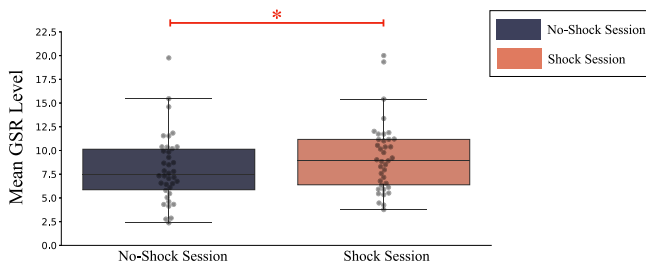
The threat of receiving an electric shock is likely to increase the stress level in participants. We confirmed that the galvanic skin response (GSR) was higher in the shock session than in the no-shock session (two-sample  $t$ -test  $t(39) = 4.17, p < 0.001, d = 0.67, 95\%CI = [0.62, 1.80]$ ) (Fig. 3). This indicates that participants were significantly more stressed due to the threat of punishment in the shock session as compared to the no-shock session.

### Model-free decision-making is associated with motor errors under threat of stressful stimuli

We next examined whether a stressful stimulus affects motor performance. We measured motor performance by calculating the number of errors in the sequence execution, i.e., incorrect or late button presses.

We first tested whether the relative contribution of the model-based decision (vs. model-free decision) was associated with the number of motor errors in both sessions. In the shock session, we found that the model-based weight and number of errors (shocks) were significantly correlated ( $r(38) = -0.51, p < 0.001, 95\%CI = [-0.74, -0.23], BF_{10} = 46.48$ ) (Fig. 4A). In the no shock session, the correlation between the number of errors and the model-based weights did not reach statistical significance, nor did the BF provide at least moderate evidence for or against a relationship ( $r(38) = -0.26, p = 0.10, 95\%CI = [-0.54, 0.05], BF_{10} = 0.73$ ) (Fig. 4B). This finding suggests that participants with a higher contribution of the model-free decision-making strategy made more errors, particularly under the threat of stressful stimuli. We also observed that participants performed better in the shock session than in the no-shock session as they made less errors in the shock session than in the no-shock session (two sample  $t$ -test  $T(39) = 2.79, p = 0.008, d = 0.44, 95\%CI = [-6.63, -1.06]$ ). This performance improvement could be attributed to the learning effect alongside increased attention and risk aversion prompted by the threat of shocks.

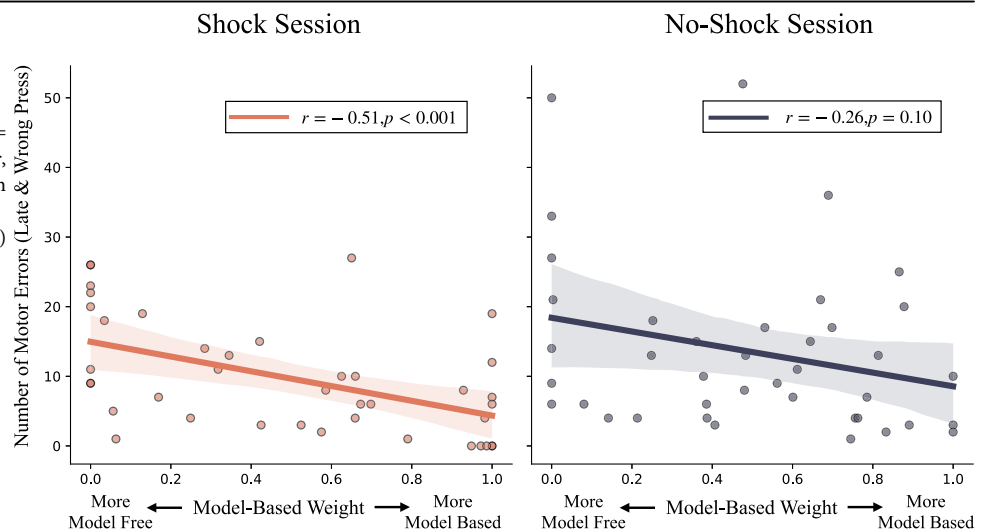
A potential alternative explanation of the correlation between model-free decision-making and motor errors is that the participants whose behaviour was mainly model-free were less engaged with the experiment overall, making more errors and receiving more shocks. However, we observed that the model-based weight of participants changed from the no-shock session to the shock session (Fig. 5). In other words, participants who were mostly model-free in the no-shock sessions were not necessarily the same participants in the shock session. We also checked whether more model-free participants showed reduced learning of the motor sequences than model-based participants. To test this, we first checked for correlations between the model-based weight and the accuracy of the motor sequence

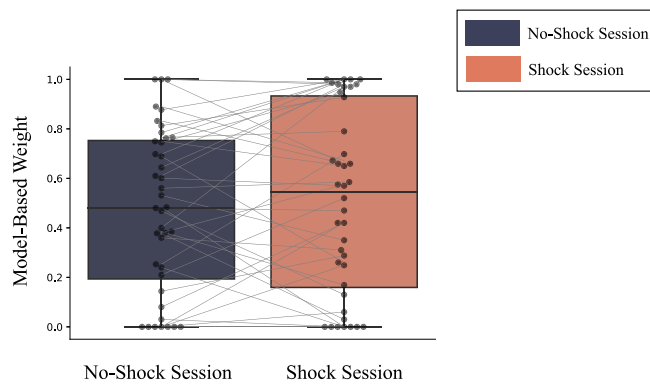


**Fig. 3 | Threat of shock affected stress.** The Galvanic Skin Response (GSR) was higher in the shock session than in the no-shock session (two sample  $t$ -test  $T(39) = 4.17, p < 0.001$ ) ( $n = 38$  participants). Error bars represent 95% confidence intervals.

### Fig. 4 | Model-free decision-making underlies motor errors.

**A** The model-based weight in the shock session was correlated ( $r(38) = -0.51, p < 0.001, 95\%CI = [-0.74, -0.23], BF_{10} = 46.48$ ) with the number of motor errors in the shock session ( $n = 40$  participants). **B** In the no-shock session, however, we did not observe a significant correlation between motor errors and the model-based weight ( $r(38) = -0.26, p = 0.10, 95\%CI = [-0.54, 0.05], BF_{10} = 0.73$ ) ( $n = 40$  participants).





**Fig. 5 | Model-based weights in no-shock and shock sessions.** Upon the transition from the no-shock session to the shock session, we found that participants ( $n = 40$  participants) changed their model-based weight. Therefore, not all participants who were mostly model-based (or model-free) in the no-shock sessions necessarily remained that way in the shock session. Error bars represent 95% confidence intervals.

learning. The accuracy of the sequence learning was calculated based on the number of times a sequence was correctly executed in the last block of training. We found no significant correlation between the sequence learning accuracy and model-based weight (Sequence-A: No-Shock Session,  $r(38) = 0.19$ ,  $p = 0.25$ , 95%CI =  $[-0.14, 0.50]$ ,  $BF_{10} = 0.37$ ; Shock Session,  $r(38) = -0.05$ ,  $p = 0.77$ , 95%CI =  $[-0.30, 0.25]$ ,  $BF_{10} = 0.2$ ; Sequence-B: No-Shock Session,  $r(38) = 0.07$ ,  $p = 0.65$ , 95%CI =  $[-0.29, 0.40]$ ,  $BF_{10} = 0.22$ ; Shock Session,  $r(38) = -0.06$ ,  $p = 0.7$ , 95%CI =  $[-0.36, 0.24]$ ,  $BF_{10} = 0.21$ ). We also tested for associations between the model-based weight and the reaction time of sequences towards the end of the sequence learning but found no significant correlations (Sequence-A: No-Shock Session,  $r(38) = -0.09$ ,  $p = 0.59$ , 95%CI =  $[-0.35, 0.20]$ ,  $BF_{10} = 0.23$ ; Shock Session,  $r(38) = -0.24$ ,  $p = 0.14$ , 95%CI =  $[-0.57, 0.01]$ ,  $BF_{10} = 0.57$ ; Sequence-B: No-Shock Session,  $r(38) = -0.09$ ,  $p = 0.6$ , 95%CI =  $[-0.39, 0.23]$ ,  $BF_{10} = 0.22$ ; Shock Session:  $r(38) = -0.2$ ,  $p = 0.21$ , 95%CI =  $[-0.54, 0.14]$ ,  $BF_{10} = 0.42$ ). The correlations which yielded Bayes Factor less than 0.3 provide at least moderate support for the absence of associations between model-based weight and sequence learning measures.

We observed that participants changed the relative weight of their model-based behaviour upon transitioning from the no-shock to the shock session. We, therefore, examined whether this shift in decision-making strategy can predict the vulnerability to making errors in motor execution under pressure. We found that the shift in the model-based weight from the no-shock session to the shock session towards model-free learning is positively correlated to the number of shocks ( $r(38) = -0.45$ ,  $p = 0.003$ , 95%CI =  $[-0.64, -0.23]$ ,  $BF_{10} = 12.1$ ) received in the shock session. We also checked for associations between the shift in the model-based weight and the difference in the number of motor errors between the two sessions. There was no statistically significant linear correlation ( $r(38) = -0.01$ ,  $p = 0.94$ , 95%CI =  $[-0.41, 0.32]$ ,  $BF_{10} = 0.19$ ), but this result is possibly because the number of shocks is not a linear measure of performance. When the absolute number of errors is high, each additional error may not correspond to a proportionate decrease in performance, which led us to hypothesise that the relationship is logarithmic. Indeed, the correlation between the shift in the model-based weight and the difference in the logarithmic number of motor execution errors was significant ( $r(38) = -0.33$ ,  $p = 0.04$ , 95%CI =  $[-0.61, -0.01]$ ,  $BF_{10} = 1.54$ ). This suggests that shift in strategy towards model-based learning is associated with less motor errors, but the relationship is not linear.

To examine the reliability of the estimation of the model-based weight, we also calculated the split-half reliability of model-based weight in our task by estimating separately for even and odd trials. We confirmed the overall high internal consistency of the model-based weight with the split-half reliability of 0.864.

## Model-free decision-making is associated with slower movement under threat of stressful stimuli

Our analysis so far revealed that the adoption of model-free decision-making is associated with a decline in motor performance, resulting in a higher occurrence of shocks in the shock session. Next, we sought to understand the reason for errors in the motor sequence, which were the result of wrong button presses and late button presses. (Note, slowly executed sequences were accompanied by incorrect button presses (Supplementary Fig. 3).)

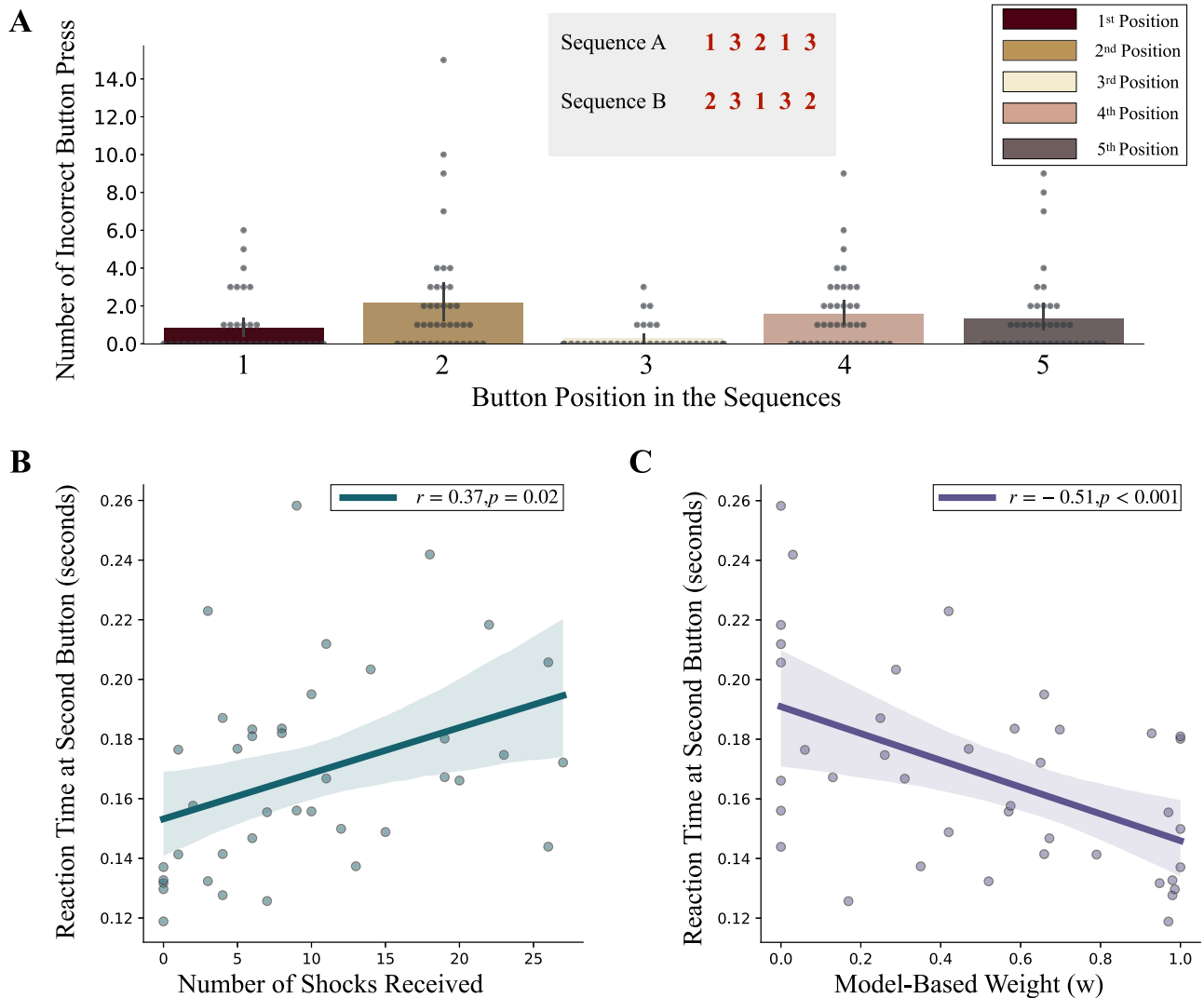
We plotted the number of wrong button presses across the positions in the five-element sequence (Fig. 6A). Notably, participants were most prone to incorrectly press the button at the second and fourth positions. This could be because both those presses were executed using the ring finger, which has less dexterity than the other fingers. To examine this possibility, we looked at the time taken to press the button at positions 2 and 4 in the sequence (button-2 reaction time and button-4 reaction time) and the number of shocks received. We found that the participant's button-2 reaction time was correlated with the number of shocks ( $r(38) = 0.37$ ,  $p = 0.02$ , 95%CI =  $[0.11, 0.61]$ ,  $BF_{10} = 2.93$ ) (Fig. 6B). (Note that the Bayes Factor value of 2.93 indicates a moderate evidence of the association). This effect was observed for button-2 reaction times at both stages in the two-step sequential movement task (Supplementary Fig. 4). We did not see a similar effect for button-4 reaction times ( $r(38) = 0.2$ ,  $p = 0.21$ , 95%CI =  $[-0.12, 0.57]$ ,  $BF_{10} = 0.42$ ). We also tested for other button times but did not see a significant correlation. We reason that the significant correlation for only button-2 may arise because, in sequential motor tasks, only the first action is preplanned<sup>40</sup>.

The correlation between button time and number of shocks (errors) seems to contradict the speed-accuracy trade-off at first glance, which would predict an inverse relationship between the reaction time and number of errors. On the other hand, it could also mean that people who are faster at executing the motor sequences do not make many errors because they have learnt the sequences better. To understand this correlation further, we examined whether it arose from the specific cognitive strategy (model-based vs. model-free learning) adopted by the participants. We observed that in the shock session, model-based participants had faster button-2 reaction times than model-free participants by observing that the model-based weight was negatively correlated with the button-2 reaction time ( $r(38) = -0.5$ ,  $p = 0.001$ , 95%CI =  $[-0.68, -0.25]$ ,  $BF_{10} = 35.59$ ) (Fig. 6C). We did not observe a similar correlation in the no-shock session ( $r(38) = -0.23$ ,  $p = 0.15$ , 95%CI =  $[-0.53, 0.04]$ ,  $BF_{10} = 0.53$ ). This suggests that some participants, in particular, model-free participants found it difficult to execute the early components of the motor sequences (at button-2).

To explain that result, we speculated that model-based participants had less time to execute their decision due to the 2-second threshold to make a decision and execute the sequence and that the model-based inference might require more time. However, we found no correlation between the model-based weight and the time taken to make the decision in the shock session ( $r(38) = 0.19$ ,  $p = 0.24$ , 95%CI =  $[-0.15, 0.48]$ ,  $BF_{10} = 0.38$ ).

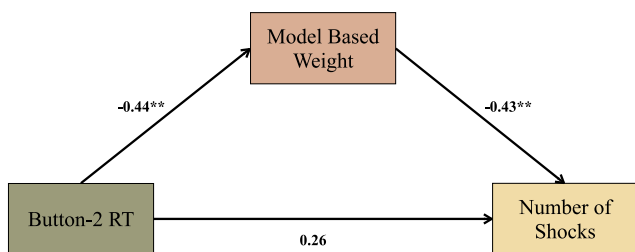
## Arbitration between strategies impacts motor performance

To further examine the directional link between the cognitive strategy (model-based and model-free learning) and the performance of the sequential motor performance, we conducted a mediation analysis. Specifically, we investigated whether the model-based weight mediates the relationship between the button-2 reaction time and the number of errors made during the sequence execution. The results revealed that the model-based weight significantly mediates the relationship between the sequence completion time and the number of errors made during the task (Fig. 7, Supplementary Table 3). The directionality of the association between the model-based weight and the number of shocks was further strengthened by treating the button-2 reaction time as the outcome, number of shocks as the primary predictor, and the model-based weight as the mediator. However, we did not observe a significant mediation effect. This finding suggests that individuals who exhibit a greater tendency towards model-based behaviour have a more efficient strategy for completing the task, resulting in fewer



**Fig. 6 | Model-free decision-making is associated with slower movement.** A Histogram of the wrong button presses across buttons in the five-element sequences ( $n = 40$  participants) Error bars represent 95% confidence intervals. B The button-2 reaction time was correlated with the number of shocks ( $r(38) = 0.37$ ,

$p = 0.02$ , 95%CI = [0.11, 0.61],  $BF_{10} = 2.93$ ) ( $n = 40$  participants). C The negative correlation with the model-based weight shows that model-based participants were faster in their button-2 reaction time than model-free participants ( $r(38) = -0.5$ ,  $p = 0.001$ , 95%CI = [-0.68, -0.25],  $BF_{10} = 35.59$ ) ( $n = 40$  participants).



**Fig. 7 | Mediation Analysis.** A mediation model examining the mediation effects of the model-based weight on the association between the button-2 reaction time and number of shocks ( $n = 40$  participants). Numbers indicate standardised regression coefficients. \*\* $p < 0.01$ . The model-based weight significantly mediated the relationship between the button-2 reaction time and the number of errors made during the task. RT reaction time.

errors. Conversely, individuals who exhibit a greater tendency towards model-free behaviour are more susceptible to making errors due to a less efficient strategy for completing the task. Altogether, the present study demonstrated that the balance between the model-based and model-free

strategies at the individual level underlies sequential motor performances under pressure.

**Discussion**

Actions in our daily lives are rarely performed in isolation and are often intertwined with decision-making that defines the goals of our actions. This relationship leads to an intriguing question: To what extent does the decision-making strategy influence the way the actions are executed and vice versa, particularly when we are under pressure? In the present study, we demonstrated that the detrimental effect of the threat of electrical shocks on rapid sequential movements is tightly coupled with the decision-making strategies that produce the actions. More specifically, we obtained compelling evidence showing that the extent to which people employ model-based or model-free strategies impacts the sequential movement performance produced from the decision.

When we make decisions and motor actions under the threat of stressful stimuli or with an increasing complexity of environments, we could use either a more deliberative, planned approach (model-based) or a more computationally easy habitual approach (model-free). The balance between the two strategies encompasses the inherent trade-off between meticulous planning to avert aversive outcomes and the conservation of cognitive

resources. Our results revealed that under the threat of stressful stimuli, people are more susceptible to a poorer motor performance when they adopt or shift towards a model-free strategy compared to a model-based strategy.

As an explanation for the performance deterioration under pressure<sup>41–44</sup>, the self-focus theory proposed in the context of choking posits that individuals focus too much attention on actions under pressure<sup>45–47</sup>. For example, in domains such as sports and music, performers are extensively and explicitly trained on a sequence of actions, with the assembly of actions becoming increasingly implicit (automatic) with experience. At the beginning of learning, performers are consciously aware of every step of a task. As they learn and improve their performance of the task, their actions gradually become automatic, and they become less conscious about each step of the task. Once under the condition of pressure, however, they become anxious about making errors and pay explicit attention to each step of the task again. In other words, in the presence of pressure, explicit processes interfere with the automatic control, which leads to poorer performance. The present study suggested that the explicit and stepwise attentional shift in self-focus theory is related to the shift towards model-free decision-making, which does not consider long-term state transitions and is likely to make the extent of attention narrower and treat sequential movements in a more fragmented manner.

Consistent with this view, we observed that model-free participants who suffered more shocks in the shock session used relatively more time to execute the sequence after making a decision (i.e., after pressing the first button of the sequence) than model-based participants. The slower execution of the sequence for model-free participants seems to contradict the traditional concept of a model-free strategy at first glance, which is habitual when one deploys a relatively inflexible but fast experience-based controller. However, a recent study reported that humans can reduce mental effort by using a model-based strategy alone<sup>48</sup>, suggesting that the model-based strategy is not necessarily slow, and model-free decision-making can be associated with slower motor execution. Thus, the introduction of decision-making into studies of motor control can deepen the understanding of motor performances.

In addition, we observed slower button presses by model-free participants for the time taken to press button-2 but not the other buttons. This observation may be connected to a sequential moving task experiment showing preplanning for the first movement and not subsequent movements<sup>49</sup>. In our task, after making the decision and pressing the first button in the sequence, the next action is to press button-2, possibly explaining why errors made at button-2 are associated with the overall motor performance in the shock session. However, we performed the same analysis for the no-shock session and did not observe a significant correlation between the button reaction time and the number of errors for any buttons. Future research should examine why model-free decision-making takes longer to upload motor sequence memory.

### Limitations

It is known that well-designed verbal instructions can significantly shift model-free decision-makers to model-based ones<sup>49</sup>. This effect suggests that verbal instructions can also change decision-making strategies in goal-oriented sequential movements and improve performance under pressure, such as playing sports and music in competition. Investigating the effect of verbal instructions on goal-oriented sequential movement performances would be an interesting topic for future research. It may be useful, for example, to emphasise longer structures or states in sequential movements.

Given that psychological pressure influenced the performance of model-free participants more in the shock session, certain cognitive biases and personality traits might also have influenced the tendency to use a model-free strategy and be more prone to the detrimental effect of pressure. Personality scores such as the Big Five Personality Traits, and trait anxiety scores should help contextualise and deepen the insights from our findings<sup>50</sup>.

In our task, we calibrated the value of shock individually for participants but did not collect data of the absolute value of shock delivered. Future studies can collect this data, and the value of shock could be introduced as a covariate in the analysis.

To measure the shift in strategy from the no-shock session to the shock session, the present study calculated the raw difference of model-based weights between the two sessions for each participant. However, the estimation of this measure would be done more reliably and directly by using hierarchical modelling.

### Conclusion

Overall, the present study demonstrated that a balance between model-based and model-free decision-making strategies impacts the performance of rapid motor sequences for achieving a decision goal. Unlike previous studies focusing either on motor performance or on decision-making strategies, the present study sheds light on the interplay between the two and exemplifies an interdisciplinary integration where an understanding of decision-making strategies specifies the mechanism underlying motor errors.

### Data availability

All data for this study have been made publicly available in an anonymized form at <https://osf.io/6x2qt/>.

### Code availability

All the code used for task, model, and analysis has been made publicly available at <https://osf.io/6x2qt/>.

Received: 11 November 2023; Accepted: 30 July 2024;

Published online: 27 August 2024

### References

1. Fermin, A., Yoshida, T., Ito, M., Yoshimoto, J. & Doya, K. Evidence for model-based action planning in a sequential finger movement task. *J. Mot. Behav.* **42**, 371–379 (2010).
2. Wolpert, D. M. & Landy, M. S. Motor control is decision-making. *Curr. Opin. Neurobiol.* **22**, 996–1003 (2012).
3. Yu, R. Choking under pressure: the neuropsychological mechanisms of incentive-induced performance decrements. *Front. Behav. Neurosci.* **9**, 19 (2015).
4. Kurniawan, I. T. et al. Choosing to make an effort: the role of striatum in signaling physical effort of a chosen action. *J. Neurophysiol.* **104**, 313–321 (2010).
5. Shadmehr, R., Huang, H. J. & Ahmed, A. A. A representation of effort in decision-making and motor control. *Curr. Biol.* **26**, 1929–1934 (2016).
6. Pessiglione, M. et al. An effect of dopamine depletion on decision-making: the temporal coupling of deliberation and execution. *J. Cogn. Neurosci.* **17**, 1886–1896 (2005).
7. Ramakrishnan, A. & Murthy, A. Brain mechanisms controlling decision making and motor planning. *Prog. Brain Res.* **202**, 321–345 (2013).
8. Ganesh, G., Minamoto, T. & Haruno, M. Activity in the dorsal ACC causes deterioration of sequential motor performance due to anxiety. *Nat. Commun.* **10**, 4287 (2019).
9. Daw, N. D., Niv, Y. & Dayan, P. Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* **8**, 1704–1711 (2005).
10. Doya, Kenji, Kazuyuki Samejima, Kenichi Katagiri & Kawato, Mitsuo Multiple model-based reinforcement learning. *Neural Comput.* **14** **6**, 1347–1369 (2002). no.
11. Dolan, R. J. & Dayan, P. Goals and habits in the brain. *Neuron* **80**, 312–25 (2013). pmid:24139036.
12. Kuvayev, L., & Sutton, R. S. (1996). Model-based reinforcement learning with an approximate, learned model. In Proceedings of the



- ninth Yale workshop on adaptive and learning systems (pp. 101–105). Yale University New Haven, CT.
13. Miller, K. J., Botvinick, M. M. & Brody, C. D. Dorsal hippocampus contributes to model-based planning. *Nat. Neurosci.* **20**, 1269–1276 (2017).
  14. Doll, B. B., Duncan, K. D., Simon, D. A., Shohamy, D. & Daw, N. D. Model-based choices involve prospective neural activity. *Nat. Neurosci.* **18**, 767–772 (2015).
  15. Dickinson, A. Actions and habits: the development of behavioural autonomy. *Philosophical Transactions of the Royal Society of London. B. Biol. Sci.* **308**, 67–78 (1985).
  16. Sutton, R. S., & Barto, A. G. (2018). *Reinforcement Learning: An Introduction*. MIT Press.
  17. Schultz, W., Dayan, P. & Montague, P. R. A neural substrate of prediction and reward. *Science* **275**, 1593–1599 (1997).
  18. Daw, N. D., Gershman, S. J., Seymour, B., Dayan, P. & Dolan, R. J. Model-based influences on humans' choices and striatal prediction errors. *Neuron* **69**, 1204–1215 (2011).
  19. Gremel, C. M. & Costa, R. M. Orbitofrontal and striatal circuits dynamically encode the shift between goal-directed and habitual actions. *Nat. Commun.* **4**, 2264 (2013).
  20. Akam, T., Costa, R. & Dayan, P. Simple plans or sophisticated habits? State, transition and learning interactions in the two-step task. *PLoS Comput. Biol.* **11**, e1004648 (2015).
  21. Gruner, P., Anticevic, A., Lee, D. & Pittenger, C. Arbitration between action strategies in obsessive-compulsive disorder. *Neuroscientist* **22**, 188–198 (2016).
  22. Doll, B. B., Bath, K. G., Daw, N. D. & Frank, M. J. Variability in dopamine genes dissociates model-based and model-free reinforcement learning. *J. Neurosci.* **36**, 1211–1222 (2016).
  23. Russek, E. M., Momennejad, I., Botvinick, M. M., Gershman, S. J. & Daw, N. D. Predictive representations can link model-based reinforcement learning to model-free mechanisms. *PLoS Comput. Biol.* **13**, e1005768 (2017).
  24. Linnebank, F. E., Kindt, M. & de Wit, S. Investigating the balance between goal-directed and habitual control in experimental and real-life settings. *Learn. Behav.* **46**, 306–319 (2018).
  25. Gläscher, J., Daw, N., Dayan, P. & O'Doherty, J. P. States versus rewards: dissociable neural prediction error signals underlying model-based and model-free reinforcement learning. *Neuron* **66**, 585–595 (2010).
  26. Kool, W., Gershman, S. J. & Cushman, F. A. Cost-benefit arbitration between multiple reinforcement-learning systems. *Psycholog. Sci.* **28**, 1321–1333 (2017).
  27. Pezzulo, G., Rigoli, F. & Friston, K. J. Hierarchical active inference: a theory of motivated control. *Trends Cogn. Sci.* **22**, 294–306 (2018).
  28. Schad, D. J. et al. Processing speed enhances model-based over model-free reinforcement learning in the presence of high working memory functioning. *Front. Psychol.* **5**, 1450 (2014).
  29. Ruan, Z. et al. Impairment of arbitration between model-based and model-free reinforcement learning in obsessive-compulsive disorder. *Front. Psychiatry* **14**, 1162800 (2023).
  30. Lee, S. W., Shimojo, S. & O'Doherty, J. P. Neural computations underlying arbitration between model-based and model-free learning. *Neuron* **81**, 687–699 (2014).
  31. Kim, D., Park, G. Y., O' Doherty, J. P. & Lee, S. W. Task complexity interacts with state-space uncertainty in the arbitration between model-based and model-free learning. *Nat. Commun.* **10**, 5738 (2019).
  32. Lockwood, P. L., Klein-Flügge, M. C., Abdurahman, A. & Crockett, M. J. Model-free decision making is prioritized when learning to avoid harming others. *Proc. Natl Acad. Sci.* **117**, 27719–27730 (2020).
  33. Otto, A. R., Raio, C. M., Chiang, A., Phelps, E. A. & Daw, N. D. Working-memory capacity protects model-based learning from stress. *Proc. Natl Acad. Sci.* **110**, 20941–20946 (2013).
  34. Park, H., Lee, D. & Chey, J. Stress enhances model-free reinforcement learning only after negative outcome. *PLoS One* **12**, e0180588 (2017).
  35. Radenbach, C. et al. The interaction of acute and chronic stress impairs model-based behavioral control. *Psychoneuroendocrinology* **53**, 268–280 (2015).
  36. Wirz, L., Bogdanov, M. & Schwabe, L. Habits under stress: mechanistic insights across different types of learning. *Curr. Opin. Behav. Sci.* **20**, 9–16 (2018).
  37. Cremer, A., Kalbe, F., Gläscher, J. & Schwabe, L. Stress reduces both model-based and model-free neural computations during flexible learning. *NeuroImage* **229**, 117747 (2021).
  38. Wyckmans, F. et al. The modulation of acute stress on model-free and model-based reinforcement learning in gambling disorder. *J. Behav. Addictions* **11**, 831–844 (2022).
  39. Gershman, S. J. Empirical priors for reinforcement learning models. *J. Math. Psychol.* **71**, 1–6 (2016).
  40. Yokoi, A., Arbuckle, S. A. & Diedrichsen, J. The role of human primary motor cortex in the production of skilled finger sequences. *J. Neurosci.* **38**, 1430–1442 (2018).
  41. Kinrade, N. P., Jackson, R. C. & Ashford, K. J. Reinvestment, task complexity and decision making under pressure in basketball. *Psychol. Sport Exerc.* **20**, 11–19 (2015).
  42. Chib, V. S., De Martino, B., Shimojo, S., & O'Doherty, J. P. Neural mechanisms underlying paradoxical performance for monetary incentives are driven by loss aversion. *Neuron* **74**, 582–594 (2012).
  43. Mobbs, D. et al. Choking on the money: reward-based performance decrements are associated with midbrain activity. *Psychol. Sci.* **20**, 955–962 (2009).
  44. Anthony, J.P. & Delgado, M. R. Stress and decision making: effects on valuation, learning, and risk-taking. *Curr. Opin. Behav. Sci.* **14**, 33–39 (2009).
  45. Baumeister, R. F. Choking under pressure: self-consciousness and paradoxical effects of incentives on skillful performance. *J. Personal. Soc. Psychol.* **46**, 610 (1984).
  46. Hill, D. M., Hanton, S., Matthews, N. & Fleming, S. Choking in sport: A review. *Int. Rev. Sport Exerc. Psychol.* **3**, 24–39 (2010).
  47. DeCaro, M. S., Thomas, R. D., Albert, N. B. & Beilock, S. L. Choking under pressure: multiple routes to skill failure. *J. Exp. Psychol.: Gen.* **140**, 390 (2011).
  48. Feher da Silva, C., Lombardi, G., Edelson, M. & Hare, T. A. Rethinking model-based and model-free influences on mental effort and striatal prediction errors. *Nature Human Behaviour* **7**, 956–969 (2023).
  49. Feher da Silva, C. & Hare, T. A. Humans primarily use model-based inference in the two-stage task. *Nat. Human Behaviour* **4**, 1053–1066 (2020).
  50. Byrne, K. A., Silasi-Mansat, C. D. & Worthy, D. A. Who chokes under pressure? The Big Five personality traits and decision-making under pressure. *Personal. Individ. Differ.* **74**, 22–28 (2015).
  51. Ma, D. S., Correll, J. & Wittenbrink, B. The Chicago face database: A free stimulus set of faces and norming data. *Behav. Res. methods* **47**, 1122–1135 (2015).

## Acknowledgements

We are grateful to Satoshi Tada for technical assistance, and Peter Karagiannis for editing an early version of the manuscript. This work was supported by KAKENHI (22H05155), JST CREST (JPMJCR22P4), and JST Moonshot R&D (JPMJMS2011) to M.H. The funders had no role in study design, data collection and analysis, decision to publish or preparation of the manuscript.

## Competing interests

The authors declare no competing interests.

### Additional information

**Supplementary information** The online version contains supplementary material available at <https://doi.org/10.1038/s44271-024-00123-3>.

**Correspondence** and requests for materials should be addressed to Pranav Sankhe or Masahiko Haruno.

**Peer review information** *Communications psychology* thanks Laura Bustamante and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editors: Patricia Lockwood and Marike Schiffer. A peer review file is available.

**Reprints and permissions information** is available at <http://www.nature.com/reprints>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Open Access** This article is licensed under a Creative Commons Attribution-NonCommercial-NoDerivatives 4.0 International License, which permits any non-commercial use, sharing, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if you modified the licensed material. You do not have permission under this licence to share adapted material derived from this article or parts of it. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by-nc-nd/4.0/>.

© The Author(s) 2024