


Modulation of Motor Vigor by Expectation of Reward Probability Trial-by-Trial Is Preserved in Healthy Ageing and Parkinson's Disease Patients

Margherita Tecilla,¹ Michael Großbach,² Giovanni Gentile,³ Peter Holland,¹ Sebastian Sporn,⁴ Angelo Antonini,³ and  Maria Herrojo Ruiz¹

¹Department of Psychology, Goldsmiths, University of London, London SE146NW, United Kingdom, ²Institute of Music Physiology and Musicians' Medicine, Hannover University of Music Drama and Media, Hannover 30175, Germany, ³Parkinson and Movement Disorders Unit, Study Center for Neurodegeneration (CESNE), Department of Neuroscience, University of Padua, Padua 35131, Italy, and ⁴Department of Clinical and Movement Neuroscience, Queen Square Institute of Neurology, University College London, London WC1N3BG, United Kingdom

Motor improvements, such as faster movement times or increased velocity, have been associated with reward magnitude in deterministic contexts. Yet whether individual inferences on reward probability influence motor vigor dynamically remains undetermined. We investigated how dynamically inferring volatile action-reward contingencies modulated motor performance trial-by-trial. We conducted three studies that coupled a reversal learning paradigm with a motor sequence task and used a validated hierarchical Bayesian model to fit trial-by-trial data. In Study 1, we tested healthy younger [HYA; 37 (24 females)] and older adults [HOA; 37 (17 females)], and medicated Parkinson's disease (PD) patients [20 (7 females)]. We showed that stronger predictions about the tendency of the action-reward contingency led to faster performance tempo, commensurate with movement time, on a trial-by-trial basis without robustly modulating reaction time (RT). Using Bayesian linear mixed models, we demonstrated a similar invigoration effect on performance tempo in HYA, HOA, and PD, despite HOA and PD being slower than HYA. In Study 2 [HYA, 39 (29 females)], we additionally showed that retrospective subjective inference about credit assignment did not contribute to differences in motor vigor effects. Last, Study 3 [HYA, 33 (27 females)] revealed that explicit beliefs about the reward tendency (confidence ratings) modulated performance tempo trial-by-trial. Our study is the first to reveal that the dynamic updating of beliefs about volatile action-reward contingencies positively biases motor performance through faster tempo. We also provide robust evidence for a preserved sensitivity of motor vigor to inferences about the action-reward mapping in aging and medicated PD.

Key words: action-reward contingency; motor vigor; Parkinson's disease; predictive coding; reward; uncertainty

Significance Statement

Navigating a world rich in uncertainty relies on updating beliefs about the probability that our actions lead to reward. Here, we investigated how inferring the action-reward contingencies in a volatile environment modulated motor vigor trial-by-trial in healthy younger and older adults, and in Parkinson's disease (PD) patients on medication. We found an association between trial-by-trial predictions about the tendency of the action-reward contingency and performance tempo, with stronger expectations speeding the movement. We additionally provided evidence for a similar sensitivity of performance tempo to the strength of these predictions in all groups. Thus, dynamic beliefs about the changing relationship between actions and their outcome enhanced motor vigor. This positive bias was not compromised by age or Parkinson's disease.

Received Aug. 18, 2022; revised Dec. 13, 2022; accepted Dec. 31, 2022.

Author contributions: G.G., P.H., S.S., A.A., and M.H.R. designed research; M.T. performed research; M.T., M.G., and M.H.R. analyzed data; M.T. wrote the first draft of the paper; M.G., P.H., and S.S. edited the paper; M.T. and M.H.R. wrote the paper.

We thank Osama Shah for programming the task in JavaScript. We also thank Caterina Tagliavini and McKenna Hedman for helping in data collection of Study 3.

The authors declare no competing financial interests.

Correspondence should be addressed to Maria Herrojo Ruiz at m.herrojo-ruiz@gold.ac.uk.

<https://doi.org/10.1523/JNEUROSCI.1583-22.2022>

Copyright © 2023 the authors

Introduction

The prospect of obtaining rewards invigorates motor performance, with incentives leading to faster and more accurate movements (Summerside et al., 2018; Sedaghat-Nejad et al., 2019; Codol et al., 2020). Several nonmutually exclusive mechanisms have been proposed to account for the beneficial effects of reward on movement. These include the reward-driven strengthening of motor representations at the cortical level (Galaro et al., 2019; Adkins and Lee, 2021), enhanced feedback-control processes (Padmala and Pessoa, 2011; Carroll et al., 2019; Manohar et

al., 2019), increased limb stiffness (Codol et al., 2020), and coarticulation (Sporn et al., 2022; Aves et al., 2021). Despite the growing number of studies demonstrating how rewards positively bias motor behavior, the evidence so far is limited to simple manipulations of reward magnitude (presence/absence; large/small). Yet, in our everyday life, we are exposed to environments rich in uncertainty, where adaptive behavior relies on estimating the changing relationship between actions and their outcomes. How beliefs about the probabilistic structure of reward contingencies modulate motor performance remains largely unexplored. In addition, whether this modulation is compromised with age and in neurologic conditions is unclear.

Hierarchical Bayesian inference models explain how individuals learn and make decisions under uncertainty (den Ouden et al., 2010; Feldman and Friston, 2010). On a neural level, processing uncertainty and updating beliefs about action-reward contingencies likely involves the anterior cingulate cortex (ACC; Behrens et al., 2007; Hayden et al., 2011), medial prefrontal cortex (mPFC; Rouault et al., 2019), and orbitofrontal cortex (OFC; Rolls et al., 2022). In multi-armed bandit and reversal learning tasks, these models describe learning as governed by inferences on the probabilistic stimulus-outcome mappings, as well as higher-level beliefs about the rate of change of these contingencies over time, labeled volatility (de Berker et al., 2016; Sheffield et al., 2022). In Bayesian predictive coding, beliefs about the probable causes of sensory data are updated via prediction errors (PEs) weighted by uncertainty or precision (Friston et al., 2014; Mathys et al., 2014). Thus, dynamic estimates of uncertainty allow for the expression of individual differences in belief updating. If motor vigor is modulated by beliefs about the action-reward contingencies, then individual differences in uncertainty estimates could explain differences in motor vigor. Alternatively, under equivalent signatures of decision-making behavior, individuals could exhibit differential sensitivity of motor performance to the expectation of reward probability.

We tested these hypotheses in three behavioral studies that used a reward-based motor decision-making task based on a reversal learning paradigm with changing stimulus-outcome contingencies over time.

In the first study, we investigated whether dynamic predictions about volatile action-reward contingencies influence motor sequence performance trial-by-trial. We additionally assessed whether the sensitivity of motor performance to the strength of these expectations undergoes changes in later stages of life and in patients with Parkinson's disease (PD) on their dopamine-replacement medication. This is motivated by the lack of evidence regarding how reward sensitivity and reversal learning interact to modulate motor vigor in PD and older adults. On the one hand, evidence supports preserved sensitivity to rewards and probabilistic learning in aging and medicated PD (Fera et al., 2005; Euteneuer et al., 2009; Aves et al., 2021). Yet other work suggests impoverished decision-making and reward-based learning in both groups. Specifically, aging and medicated PD can underperform in tasks using volatile probabilistic stimulus-outcome mappings (Cools et al., 2001; Eppinger et al., 2011; Nassar et al., 2016). However, the medication effects on decision-making in PD (on/off states) is still under debate (Ryterska et al., 2013; Kjær et al., 2018). Accordingly, whether aging and medicated PD can use their dynamic belief estimates to invigorate motor performance trial-by-trial remains unspecified.

In the second study, we evaluated the potential contribution of retrospective subjective inferences about credit assignment to

explain the motor vigor results. Last, we assessed how explicit beliefs about the reward tendency (confidence ratings) modulated motor performance trial-by-trial. This aimed at providing a more comprehensive understanding of the motor invigoration effect by beliefs about volatile reward probabilities.

Materials and Methods

Participants

All studies received ethical approval by the review board of Goldsmiths (healthy sample), University of London, and the Neurology Clinic, Padua University Hospital [Parkinson's disease (PD) sample]. Informed consent was acquired for each participant. Healthy younger (HYA) and older adults (HOA) were recruited through online advertisement and via the Research Participation Scheme (RPS) at Goldsmiths University, while PD were enrolled at the Neurology Clinic, Padua University Hospital.

Study 1

A total of 37 HYA (24 females, age 18–40, mean age 27.8, SEM 0.67; hereafter, we follow the intrinsic measures of precision for rounding descriptive and inferential statistics as reported by Cousineau, 2020), 20 PD patients (7 females, age 40–75, mean age 58.9, SEM 1.32), and an age-matched group of 37 HOA (17 females, age 40–75, mean age 61.5, SEM 1.25) participated in this research. The sample size for healthy samples was informed by previous work assessing differences between HYA and HOA in decision-making under uncertainty (de Boer et al., 2017; $N = 30, 30$) and our own work assessing group effects in parameters of hierarchical Bayesian models (Hein et al., 2021; Hein and Herrojo Ruiz, 2022; $N = 20, 20$). We increased the sample size to allow for variability being introduced because of the nature of the online study.

All participants were right-handed, had normal or corrected vision and were able to perform controlled finger movements. Amateur/professional pianists and participants diagnosed with a mental health disorder were excluded from the study. Additionally, exclusion criteria for PD patients were: implanted with deep brain stimulation (DBS), taking antidepressant medications, diagnosed with dementia and displaying tremor as an onset symptom. One PD patient declared to take Laroxyl, yet confirmed not to be diagnosed with depression. PD were evaluated through ITEL-Mini Mental State Examination (ITEL-MMSE; Metitieri et al., 2001), Unified Parkinson's Disease Rating Scale part III (UPDRS-III; Fahn and Elton, 1987), Hospital Anxiety and Depression Scale (HADS; Zigmond and Snaith, 1983), and State-Trait Anxiety Inventory (STAI Y2; Spielberger et al., 1983). Supplementary disease-related information was also gathered (Table 1). Patients completed the experiment in the ON medication state according to their usual dopamine-replacement treatment. The individual dopaminergic medication details were collected and converted to a levodopa-equivalent daily dose (LEDD) value (Table 1).

All participants took part in the study remotely (online), except for five PD patients, who completed the study in the laboratory facilities of the Neurology Clinic of Padua. An Italian translation of the original experimental instructions in English was created to test some of the HOA participants ($N = 24$) and all PD patients (for details on our control analyses to assess the effect of the language of the instructions, see Results). The previously validated Italian translations of the HADS, ITEL-MMSE, UDPRS-III, and STAI Y2 scales were used. HYA and HOA participants received a monetary compensation of £5 (€5 for those completing the task in Italian), which could be increased up to £10 (€10) as a function of their task performance. PD patients did not receive a monetary prize, in line with the clinical research policies at the Neurology Clinic of Padua.

Study 2

A separate sample of 39 HYA took part in Study 2, which was aimed at evaluating the potential contribution of subjective inferences about task-related reward (credit) assignment to explain our results (McDougle et al., 2016). HYA participants in this experiment were divided into two subsamples as a function of their reply (True/False) to a post-performance question (Q8; Table 2). Group Q8_T consisted of 26 participants (18 females, age 18–40, mean age 24.1, SEM 1.13) and Q8_F of 13 participants

Table 1. PD clinical information

Patient #	Age	UPDRS III ON	ITEL-MMSE	STAI Y2	HADS_A	HADS_D	Disease duration (years)	Main symptom	Most impaired side	Last drug intake (min)	LEDD	Active substance
1	57	38	22	51	6	3	10	R/B	SX	30	920	Benserazide, levodopa, rasagiline, ropinirole
2	46	17	22	40	10	16	7	R	SX	75	1197	Carbidopa, entacapone, levodopa
3	53	10	22	42	7	5	4	R/B	DX	120	100	Rasagiline
4	63	6	22	25	4	2	3	B	DX	720	50	Selegiline
5	57	6	22	33	7	7	2	R	DX	120	300	Benserazide, levodopa
6	53	22	19	53	9	8	23	R/LE	BOTH	130	420	Carbidopa, levodopa, rotigotine
7	62	24	22	33	4	3	11	T	DX	120	1105	Benserazide, levodopa, pramipexole
8	62	6	22	28	3	5	8	R/B/D	DX	75	450	Carbidopa, levodopa, opicapone, selegiline
9	62	17	22	25	4	3	8	T	SX	100	652	Benserazide, levodopa, pramipexole, selegiline
10	69	7	21	45	5	6	3	B	SX	120	300	Benserazide, levodopa
11	58	7	20	31	5	1	9	R	DX	30	970	Amantadine, carbidopa, entacapone, levodopa, pramipexole
12	54	25	19	32	2	5	7	R	SX	40	1780	Benserazide, levodopa, rasagiline, rotigotine
13	66	16	19	34	4	10	12	R/B	DX	150	1580	Amantadine, carbidopa, levodopa, opicapone, pramipexole, safinamide
14	53	21	22	44	5	5	8	R	BOTH	5	320	Ropinirole
15	55	4	22	37	4	1	2	R/T	DX	30	452	Benserazide, levodopa, pramipexole, rasagiline
16	69	13	20	35	1	0	7	B	SX	437	470	Benserazide, levodopa, ropinirole, selegiline
17	65	5	21	26	1	7	16	R/B	SX	360	100 + 3.9 ml/h	Levodopa, opicapone, pramipexole, trihexyphenidyl levodopa infusion gel
18	59	7	21	37	2	4	2	R/B	SX	5	150	Carbidopa, levodopa
19	58	8	22	30	1	4	5	R/T	DX	100	452	Benserazide, levodopa, pramipexole
20	56	17	22	40	6	8	6	R	DX	185	1110	Amantadine, benserazide, levodopa, pramipexole

Mini Mental State Examination (MMSE) predicted score = 1.01 × ITEL-MMSE score + 5.16; Unified Parkinson’s Disease Rating Scale part III ON medication (UPDRS III ON); State-Trait Anxiety Inventory, Form Y2 (STAI Y2); Hospital Anxiety and Depression Scale, anxiety subscale (HADS_A); Hospital Anxiety and Depression Scale, depression subscale (HADS_D); rigidity (R); bradykinesia (B); lack of energy (LE); tremor (T); dyskinesia (D); levodopa-equivalent daily dose (LEDD).

Table 2. Post-performance questionnaire

Please, indicate whether the following statements are True or False. Please note that performance errors mean pressing the wrong key(s) or key(s) in the wrong order, while bad choices mean playing a sequence that received no points on that attempt.

1. I made fewer than 10 performance errors [True/False]
2. I made between 10 and 30 performance errors [True/False]
3. I made more than 30 performance errors [True/False]
4. I recognized a performance error, because the tone sounded different from expected [True/False]
5. I recognized a performance error, because the finger movement felt different [True/False]
6. I memorized the sequences focusing on the finger movements, without paying attention to the tones [True/False]
7. I memorized the sequences focusing both on the finger movements and the tones [True/False]
8. I could *always* distinguish whether 0 points reflected a performance error or a bad decision [True/False]
9. I was *often* not sure whether 0 points reflected a performance error or a bad decision [True/False]

Post-performance questionnaire included in Study 2. Question 8 (Q8) is aimed at evaluating subjective inferences about the task-related credit assignment.

(11 females, age 18–40, mean age 25, SEM 1.7). The same inclusion/exclusion criteria and compensation as for HYA in Study 1 applied.

Study 3

For Study 3, we recruited 33 HYA (27 females, age 18–40, mean age 22.4, SEM 1.14) with the aim of understanding how trial-by-trial explicit confidence ratings about action-reward contingencies modulate motor performance. The same inclusion/exclusion criteria and compensation as for HYA in Study 1 applied.

Experimental design

In Studies 1 and 2, the experiment ran completely online on the Qualtrics platform (<https://www.qualtrics.com>) and was accessible through a study

link. The task was programmed in JavaScript and embedded into the Qualtrics form. We provide more details of the data acquisition below (see Acquisition of online data using JavaScript).

Participants performed a novel computerized reward-based motor decision-making task based on a reversal learning paradigm with changing stimulus-outcome contingencies over time (de Berker et al., 2016). Participants were instructed to play one of two sequences of finger movements on a virtual piano to express their decision, which is an extension of standard reversal learning and decision-making tasks that instruct participants to manifest their choice by pressing a right or left button (Hein et al., 2021).

The task consisted of a familiarization and a reward-based learning phase. In the familiarization phase participants learned how to play two short sequences (seq1 and seq2) of four finger presses each. Each sequence was uniquely represented by one of two different fractal images (Fig. 1A). They were asked to position their right hand on the keyboard as follows: index finger on “g” key, middle finger on “h” key, ring finger on “j” key, and little finger on “k” key. Each key press reproduced a distinct auditory tone, simulating a virtual piano. Participants were trained to press “g-j-h-k” for seq1 (red fractal) and “k-g-j-h” for seq2 (blue fractal). These sequences of key presses corresponded to the “E”, “G”, “F”, “A” notes and “A”, “E”, “G”, “F” notes on the virtual piano keyboard, respectively. Online videos showing the correct hand position on the keyboard and how to perform the two sequences were provided to increase inter-individual consistency. The familiarization phase terminated when an error-free performance was achieved for five times in succession for both sequences. The number of sequence renditions during familiarization was recorded and used for subsequent analyses.

The reward-based learning phase consisted of 180 trials. On each trial, participants were instructed to choose between two colored fractals (blue and red) and correctly play the associated sequence (seq1 and seq2) to receive a reward (five points; Fig. 1B). Trial-by-trial reward feedback about participants’ choices was provided on the screen (binary: “You earned 5 points!” or “You earned 0 points”). The reward probability associated with each sequence (or icon) changed every 30–42 trials (as in de Berker et al., 2016). The mapping governing the likelihood of sequences being rewarded was reciprocal [$p(\text{win}|\text{seq1}) = 1 - p(\text{win}|\text{seq2})$] and consisted of five stimulus-

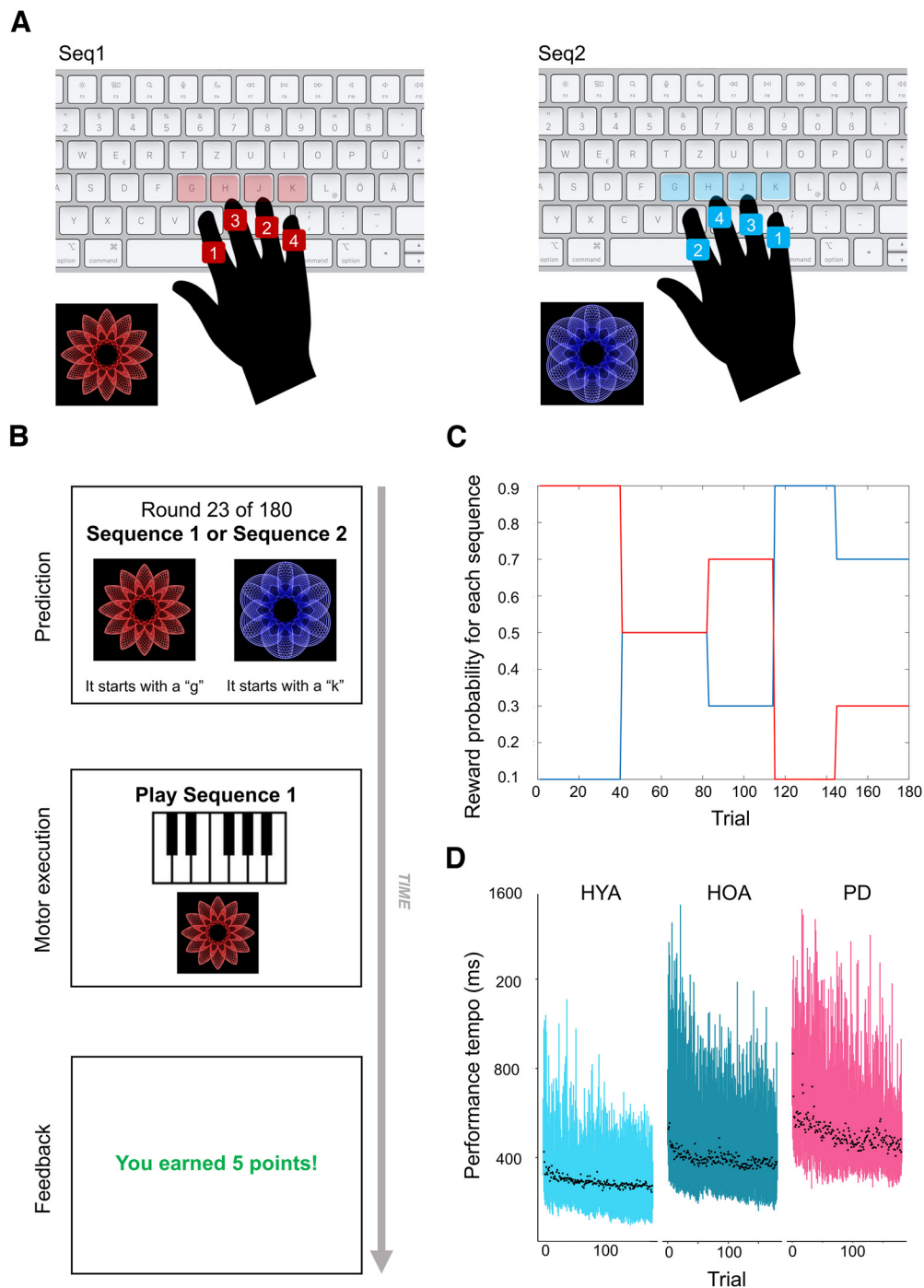


Figure 1. Task structure. **A**, In the task familiarization phase, participants learnt to play two sequences associated with two images (red fractal, seq1 “g-j-h-k”; blue fractal, seq2 “k-g-j-h”). **B**, On each trial of the reward-based learning phase, subjects decided which sequence to play to get the reward. The two icons were always either red or blue and presented to the left or right part of the screen, respectively. First, participants made a prediction about which sequence (associated to the corresponding icon) was more likely to give them a reward. When a decision was reached, they played the corresponding sequence using the keyboard. Finally, the outcome (win +5p or 0p) was revealed. In the example, the participant played seq1 and obtained five points, suggesting correct prediction and execution. In Study 3, participants were instructed to rate how certain they were of being rewarded on each trial after they performed their chosen sequence. Confidence ratings were provided by typing any number between 0 and 99 (not shown in the figure). **C**, Displays the typical subject-specific mapping of probabilistic stimulus–outcome contingency over the course of 180 trials. In the example, the order of reward mappings for the blue icon (and corresponding seq2) is 10–50–30–90–70% (reciprocal for red icon and corresponding seq1). In order to obtain the maximal reward, participants needed to track these changes and adapt their choices throughout the experiment. **D**, The trial by trial changes in performance tempo in milliseconds (mIKI; mean interkeystroke-intervals; for further details, see Behavioral and computational data analysis) for healthy younger adults (HYA; light blue), healthy older adults (HOA; dark blue), and patients with Parkinson’s disease (PD; in purple) across 180 trials in Study 1. Black dots represent the trial-by-trial within-group averages of performance tempo. Bars indicate 95% interval probabilities. Participants tended to play the sequences faster toward the end of the experiment, possibly reflecting a practice effect.

outcome contingency blocks (90/10, 70/30, 50/50, 30/70, 10/90; Fig. 1C). The order of the contingency blocks was randomly generated for each participant.

After the first key press, subjects had up to 5000 ms to perform the sequence, terminating in a Stop signal. Participants had no constraints on

initiating the sequence. Thus, reaction time (RT) included deliberation time, which reflects the time for deciding which sequence to play. Visual hints suggesting the first key to press for both sequences were displayed: “It starts with a “g”” for seq1 (red fractal); “It starts with a “k”” for seq2 (blue fractal).

Participants were instructed to press key “q” if they needed a reminder of the order of finger presses for each sequence. No participant required this reminder. As soon as participants completed the sequence within the allotted 5000 ms, the feedback was displayed.

Correctly playing the rewarded sequence added five points to the participants’ total score (win trial). Thus, receiving five points indicated that participants chose the rewarded sequence on the trial and did not make performance execution errors when playing it. Zero points, however, could reflect participants choosing an unrewarded sequence on that trial or, alternatively, choosing a rewarded sequence but performing it incorrectly (performance execution error; McDougle et al., 2016). No reward was provided when sequence performance exceeded the 5000-ms limit (no response trial) and participants were informed they played too slowly.

Thus, to maximize the total cumulative points over the experiment, participants had to infer the probability of reward associated with each sequence and adapt their choices when contingencies changed. They also had to perform the sequences correctly. Participants were informed at the beginning of the experiment that the stimulus–outcome mapping would change from time to time. However, they received no detailed information regarding the frequency or magnitude of those changes. We validated that each participant group completed the task correctly using two measures: (1) the percentage of trials that they performed either seq1 or seq2 (percPlayed); and (2) percPlayed by contingency phase. In the first case, percPlayed was used to demonstrate that participants did not have a preference toward one of the sequences, which could emerge if they perceived one sequence to be easier with regard to their motor skills. On average, we expected percPlayed to be 50% for each sequence type. Next, (2) was used to assess whether their chosen sequences tracked the contingency changes over time. To compute percPlayed by contingency phase, we estimated the rate of choosing seq1 in each contingency phase, separately in each participant. We then pooled these data across participants in each group, sorted by phases of increasing contingency values [0.1, 0.3, 0.5, 0.7, 0.9], as defined for seq1. See further details below (Behavioral and computational data analysis and Results).

In Study 2, we additionally asked participants at the end of the reward-based learning phase to reply to some questions about their performance. Similarly to Study 1, RT included deliberation time, as there were no constraints on initiating the sequence. We were particularly interested in assessing whether participants could correctly infer what zero points meant, that is, whether they could distinguish between a performance execution error or a decision to play a sequence that was unrewarded on the trial. Both scenarios would result in zero points. We reasoned that participants who could not always infer the meaning of zero might show a reduced invigoration effect. Table 2 lists the questions of the post-performance questionnaire, which required binary responses (True/False) and was designed based on previous work (McDougle et al., 2016; Herrojo Ruiz et al., 2017). The binary answer to Question 8 “I could always distinguish whether 0 points reflected a performance error or a bad decision” was used as criterion to split the control sample into Q8_T (i.e., participants were *always* sure about the hidden causes for the lack of reward) and Q8_F (i.e., participants were *not always* sure about the hidden causes for receiving zero points). Among other questions, participants were asked whether the subjective number estimate of performance errors was <10, between 10 and 30 or >30. This information was used to investigate whether Q8_T and Q8_F differed in the rate of subjective execution errors. The rationale here was that Q8_F participants relative to Q8_T could attribute more zeros to performance errors rather than inferring that their choice was not rewarded on that trial. Alternatively, they could misattribute zeros to bad decision outcomes. In both cases, their biased credit assignment would be reflected in a more pronounced difference between estimated and empirical error rates in Q8_F. However, their belief updating would differ; in the first case, Q8_F participants relative to Q8_T would not update their beliefs following a zero outcome, as this would be rendered as not informative feedback regarding the underlying probabilistic structure. Thus, differences in credit assignment could explain variation in decision-making and, potentially, associated motor vigor effects. Finally, we also assessed the strategy that participants used to memorize the sequences (79.5% of participants declared to have memorized the sequences focusing both on the finger movements and the tones; Q7).

In Study 3, we conducted an offline version of the task described above. The paradigm was coded in psychtoolbox (<http://psychtoolbox.org>) and run in MATLAB (version 2021b). In order to better capture measures of trial-wise RTs, excluding deliberation time, the 5000-ms time window for performing the sequence started at the fractals presentation (and not when the first key was pressed, as in Studies 1 and 2). Hence, reward delivery was contingent on RT and movement time (MT).

Importantly, after each sequence performance we asked participants how certain they were to be rewarded on that round (following Frömer et al., 2021). This aimed at unveiling a potential association between trial-by-trial explicit beliefs about the reward tendency (confidence ratings) and motor performance. Participants were instructed to type a number in the 0–99 range on the computer keyboard with their left hand. Value 0 denoted having no clue about receiving the points, while 99 reflected being absolutely certain of being rewarded. Participants were encouraged to explore the full 0–99 range. They were additionally asked to press the key “z” if they thought to have committed a performance execution error. This allowed us to estimate the percentage of correctly identified errors, which expands on Study 2 findings by informing about trial-by-trial (real-time) subjective inference on credit assignment. Participants had 3500 ms to complete the confidence rating, and the reward feedback was displayed after this interval.

Acquisition of online data using JavaScript

In Studies 1 and 2, because of the nature of the online experiment, cross-browser issues could emerge. A potential issue was that participants could use a variety of computer hardware, running on different web browsers, operating systems and keyboard types (e.g., tablets vs laptops). To mitigate the effect of hardware variability on the acquisition of motor performance data, we instructed participants to complete the task on a desktop or laptop computer. An inspection of browser user agent data suggests that the experiment was performed on a mixture of desktops or laptops running the Chrome and Safari browsers on Windows and Macintosh operating systems.

Timing data were collected using the web browser’s high-resolution timer. This browser resolution timer has an upper resolution limit of 2 ms on some web browsers. Therefore, all analysis scripts truncated timing data to 2-ms precision. When estimating the mean and SEM in time variables, we therefore considered a systematic error of 1 ms (2-ms precision means that our time measures were on average 1 ms too short).

For each participant, keypresses, timing data, points, contingency mapping, outcome, and other data were extracted on each trial, then stored and uploaded via JSON to the data folder in Pavlovia (see <https://gitlab.pavlovia.org/oshah001/reward-learning-experiment>).

The hierarchical Gaussian filter

To model intrasubject trial-by-trial performance in our task, we used a validated hierarchical Bayesian inference model, the Hierarchical Gaussian Filter (HGF; Mathys et al., 2011, 2014; Frässle et al., 2021). The HGF toolbox is an open source software and is freely available as part of TAPAS (<http://www.translationalneuromodeling.org/tapas>; Frässle et al., 2021). Here, we used the HGF version 6.1 implemented in MATLAB 2020b (MATLAB and Statistics Toolbox Release, The MathWorks). The HGF is a generative model that describes how individual agents learn about a hierarchy of hidden states in the environment, such as the latent causes of sensory inputs, probabilistic contingencies, and their changes over time (labeled volatility). Beliefs on each hierarchical level are updated through prediction errors (PEs) and scaled (weighted) by a precision ratio (precision as inverse variance or uncertainty). The precision ratio effectively operates as a learning rate, determining how much influence the uncertainty about the belief distributions has on the updating process (Mathys et al., 2011, 2014).

In our studies, the HGF was used to characterize subject-specific trial-by-trial trajectories of beliefs about stimulus–outcome contingencies (level 2) and their changes over time (environmental volatility, level 3). These belief distributions are Gaussian, summarized by the posterior mean (μ_2, μ_3) and the posterior variance (σ_2, σ_3). The latter represents uncertainty about the hidden states on those levels, that is, our imperfect knowledge about the true hidden states. On level 2, σ_2 is termed estimation or informational uncertainty. More generally, the inverse $1/\sigma$ is

termed precision, labeled π . The HGF provides trajectories of updated beliefs on the current trial, k , after observing the outcome (posterior mean $\mu_i^{(k)}$ for level $i = 2, 3$). Before observing the outcome, participants' predictions are denoted by the hat operator $\hat{\mu}_i^{(k)}$ and correspond to the values in the previous trial ($\mu_i^{(k-1)}$). As in previous work using reversal learning paradigms (Iglesias et al., 2013; Mathys et al., 2014; Hein et al., 2021), we modeled learning using the three-level HGF (HGF₃) for binary outcomes (Fig. 2A). In this hierarchical perceptual model, the hidden state on the lowest level, x_1 , represents the binary categorical variable of the experimental stimuli [for each trial k , $x_1^{(k)} = 0$ if the red icon/seq1 is rewarded (or blue/seq2 loses); $x_1^{(k)} = 1$ when red fractal/seq1 is not rewarded (or blue/seq2 wins)]. Higher in the hierarchy, x_2 reflects the true value of the tendency of the stimulus-outcome contingency, and x_3 the true volatility of the environment (i.e., of x_2). Belief updating in the HGF depends on various parameters, which can be estimated in each individual or fixed depending on the hypotheses. This allows for the assessment of individual learning characteristics. Here, we chose to individually estimate parameter ω_2 , representing the tonic (time-invariant) volatility on the second level, and ω_3 , denoting the tonic volatility on the third level. Generally, ω_2 and ω_3 parameters describe an individual's learning motif. Larger ω_2 values are associated with faster learning about stimulus outcomes, and thus greater update steps in μ_2 (see simulations in Hein et al., 2021). Similarly, greater levels of tonic volatility on level 3, ω_3 , increase the update steps on μ_3 . See details on our priors in Table 3. Using simulations to assess the accuracy of parameter estimation in the HGF₃, we and others have previously demonstrated that ω_2 can be estimated accurately, while ω_3 is not estimated well (Reed et al., 2020; Hein et al., 2021).

We then coupled the perceptual HGF model to a response model for binary outcomes, which defined how beliefs about the tendency of the stimulus-outcome contingencies were mapped onto decisions (e.g., which sequence should be chosen and played according to the beliefs on the current trial; Mathys et al., 2014). Our response model was the unit-square sigmoid observation model for binary responses (Iglesias et al., 2013; Mathys et al., 2014). This model estimates on each trial k the probability that the agent's response y is either 0 or 1 (Fig. 2B; $p[y^{(k)} = 1]$ and $p[y^{(k)} = 0]$), as a function of the predicted probability that the icon/sequence is rewarding. This mapping from beliefs to decisions depends on the response parameter ζ (interpreted as inverse decision noise). Higher ζ values indicate a greater probability for the agents to select the option that is more likely to be rewarding according to their beliefs. Simulations demonstrate that ζ is recovered well (Hein et al., 2021).

In the following, as stimuli (red and blue icons) are one-to-one associated with motor sequences (seq1 and seq2, respectively), we will use the term action-reward contingency when referring to stimulus-reward or stimulus-outcome mappings.

Models and priors

In line with previous work (Iglesias et al., 2013; Hein et al., 2021) we fitted the empirical data with different models. We started by modeling our data with the HGF₃ perceptual model + sigmoid response model, as described above. In this model, the third hierarchical level represents environmental volatility, that is the rate of change in the action-reward contingencies. In our paradigm the true volatility was constant across participants, as the reward contingencies changed approximately every 30–42 trials. In Study 1, using relatively uninformative priors for ω_2 , ω_3 as in previous work (prior mean -4 , -7 , respectively; prior variance 16 in both cases; de Berker et al., 2016; Iglesias et al., 2013; Hein et al., 2021) led to numerical instabilities in the HGF₃ in 20% of our participants across all groups, in particular in those exhibiting high win rates and thus learning well. The numerical instabilities also manifested when using tight priors (small variance of 4 or 1 in the prior distribution of ω_2 , ω_3), and when using prior values estimated in our data using an ideal observer model. An ideal observer is typically defined as the set of parameter values that minimize the overall surprise that an agent encounters when processing the series of inputs (see an application of an ideal observer model in Weber et al., 2020). It is likely that the divergence of the HGF₃ in 20% of our datasets is because of the trial number being smaller than in previous studies using the HGF₃ (180 instead of 320 or 400). We therefore proceeded to use the two-level HGF (HGF₂) in all our three studies, in which beliefs on

volatility on the third level are fixed. Priors for the perceptual HGF₂ model were chosen by simulating an ideal observer receiving the series of inputs that the participants observed. We then used the estimated posterior values on those model parameters as priors for the HGF₂ perceptual model coupled with our response model (Table 3). Complementing the HGF, we used two standard reinforcement learning models, the Rescorla–Wagner model (RW; fixed learning rate determined by PEs; Rescorla and Wagner, 1972) and Sutton K1 model (SK1; flexible learning rate driven by recent PEs; Sutton, 1992). Priors for reinforcement learning models were set according to previous literature (Diaconescu et al., 2014; Hein et al., 2021).

The different models (HGF₂, RW, SK1) were fitted to the trial-by-trial inputs and responses in each participant using the HGF toolbox, which generates maximum-a-posteriori (MAP) parameter estimates in each individual. To identify the model that explained the behavioral data across all participants best, we used random effects Bayesian model selection (BMS; through the freely available MACS toolbox, <https://github.com/JoramSoch/MACS>; Soch and Allefeld, 2018). Importantly, in Study 1 we used the same priors in all participant groups (HYA, HOA, PD) as in previous studies (Powers et al., 2017; Hein et al., 2021). Note, however, that recent computational modeling work suggests that using different prior values in each participant group may be more suitable to capture dissociable group effects (e.g., for mental health, see Valton et al., 2020). This approach, albeit interesting, would not favor a standard statistical comparison between groups: any between-group differences could be explained by the underlying models having been constructed differently.

Behavioral and computational data analysis

First, we validated the task by assessing (1) the percentage of trials that each sequence type was played (percPlayed) and (2) whether percPlayed followed the contingency changes (for details, see above, Experimental design). We additionally examined the percentage of trials in which each sequence type was played without performance execution errors (percCorrectlyPlayed).

General task performance in each participant was assessed by analyzing the percentage of errors (percError: rate of sequences with performance execution errors because of one or several wrong key presses), win rate (percWin: rate of trials in which the rewarded sequence is played without execution errors), the average of the trial-wise performance tempo [mIKI in milliseconds: trial-wise mean of the three interkeystroke-intervals (IKIs) across four key presses within the same trial; for trial-wise mIKI in Study 1, see Fig. 1D] and the mean of the trial-wise RT (in milliseconds: time interval between the fractal presentation and first key press). Importantly, mIKI is commensurate with movement time (MT), the time between the first and last key press ($MT = mIKI * 3$). Finally, we also assessed the number of sequence renditions that participants completed during the familiarization phase (rendFam: average of renditions across both sequence types). Time out trials and trials with performance execution errors were excluded from analyses on performance tempo and RT to avoid potential confounds, such as slowing following errors (Herrojo Ruiz et al., 2009).

Next, to investigate decision-making processes we analyzed group effects on three computational variables that characterized learning in each individual. The model that best explained the behavioral data across all participants according to BMS was the HGF₂ (see Results). We therefore assessed the perceptual model parameter ω_2 (subject-specific tonic volatility, which influences the speed of belief updating on level 2), ζ (the inverse decision noise of the response model), and the average across trials of σ_2 (posterior variance of the belief distribution). The quantity σ_2 is particularly interesting, as it represents informational uncertainty about the tendency of the action-reward contingency. Moreover, beliefs on level 2 are updated as a function of PEs about the stimulus-outcome mapping (the mismatch between the observed outcomes $u = 1$ or 0 and the agent's beliefs about the probability of such an outcome) and weighted by σ_2 (the precision ratio on level 2). Accordingly, if agents are more uncertain about the contingencies governing their environment, they will rely more on PEs to update their beliefs on that level.

To test our main research hypothesis that the strength of expectations about the action-reward contingency modulates the trial-by-trial motor performance, as a function of the group, we focused on the trajectory $\hat{\mu}_2$ (dropping trial index k for simplicity; prediction about the tendency of the action-reward contingency).

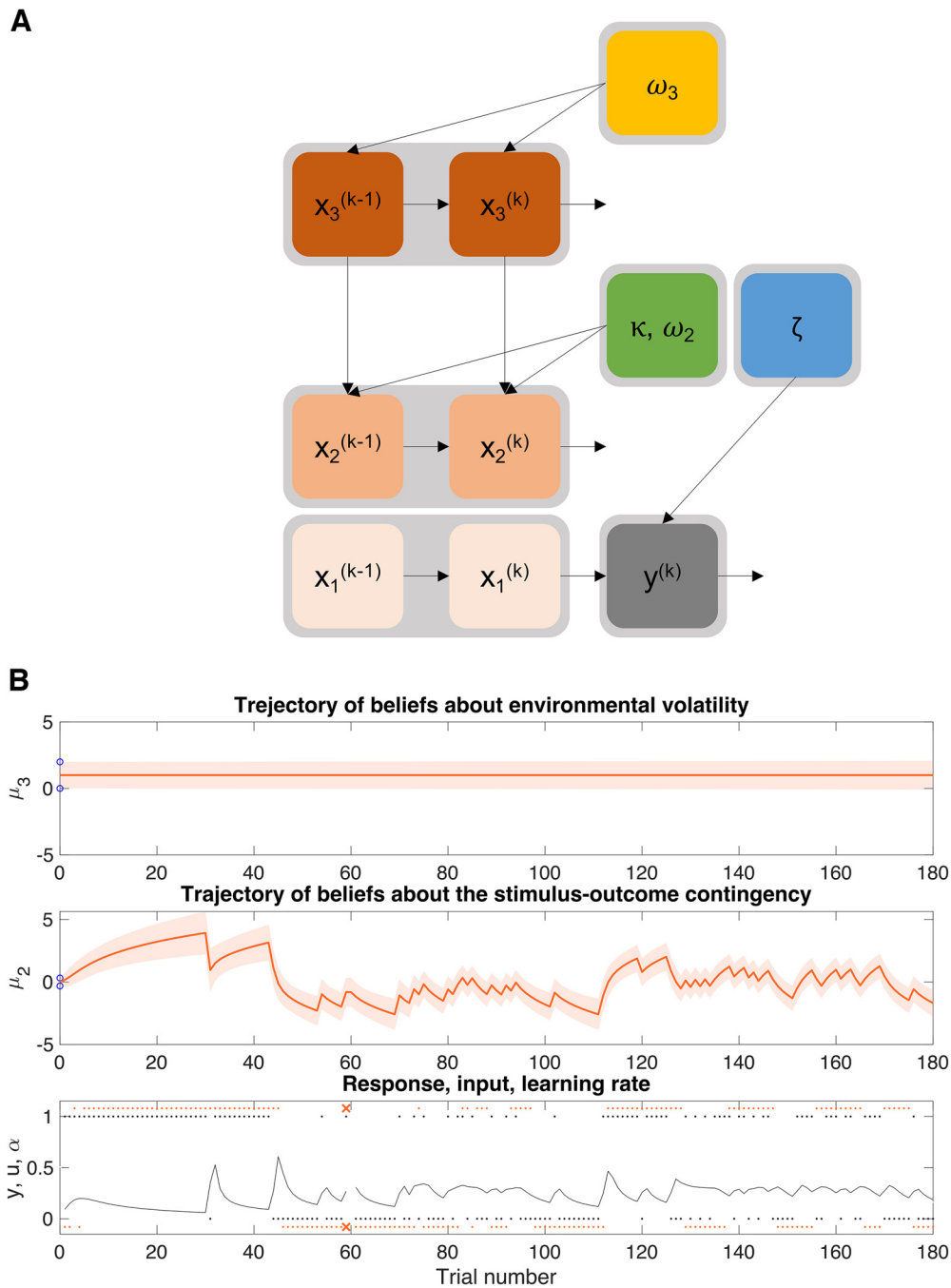


Figure 2. The hierarchical Gaussian filter (HGF) for binary outcomes. **A**, Illustration of the three-level HGF model (HGF₃) with relevant parameters modulating each level (adapted from Hein et al., 2021). Level x_i represents the binary categorical variable of the experimental stimuli on each trial k ; x_2 reflects the true value of the tendency of the stimulus-outcome contingency, and x_3 the true volatility of the environment. In our experiment, ω_2 , ω_3 and ζ were free parameters and were estimated by fitting individual responses and observed inputs with the HGF. κ represents the strength of coupling between level 2 and 3 (fixed to 1 in our study; data not shown in the text; for the model equations, see Mathys et al., 2014). **B**, In our three studies, the winning model was the two-level HGF (HGF₂), in which volatility was fixed across participants. Belief trajectories for the HGF₂ across the total 180 trials in a representative participant in Study 1. At the lowest level, black dots (u) represent the outcomes, denoting whether seq1 was rewarded or not [1 = seq1 wins (seq2 loses); 0 = seq2 wins (seq1 loses)]; orange dots (y) represent the participant's choices (1 = seq1 is played; 0 = seq2 is played); orange crosses depict performance execution errors; the black line is a subject-specific learning rate about stimulus outcomes (α for the full HGF equations, see Mathys et al., 2014). At the second level, μ_2 (σ_2) is the trial-by-trial trajectory of beliefs (mean and variance) about the tendency of the stimulus-outcome contingencies (x_2). A mean estimate μ_2 shifted toward positive values on the y -axis indicates that the participant had a greater expectation that seq1 was rewarded relative to seq2. In addition, larger (absolute) μ_2 values on that axis denote a stronger expectation that given the correct sequence choice a reward will be received. The trajectory of beliefs about phasic (log)volatility [μ_3 (σ_3)] is displayed at the top level. The true volatility in our task, x_3 , was constant, as the stimulus-outcome contingencies changed every 30–42 trials. In the winning model HGF₂, the degree of volatility was fixed across participants. Blue circles on the y -axis denote the upper and lower priors of the posterior distribution of beliefs, $\mu_i^{(0)} \pm \sigma_i^{(0)}$, $i = 2, 3$.

In Study 3, we also measured the explicit trial-wise confidence ratings (conf. number between 0 and 99) about the reward outcome to assess whether motor performance was sensitive to explicit beliefs about the reward tendency.

Statistical analyses

Bayesian analyses on Study 1

General task performance and computational variables. First, we calculated the mean and SEM as summary statistics for each of our general

Table 3. Means and variances of the priors on perceptual parameters and starting values of the beliefs of the winning HGF₂ model

Prior	Mean	Variance
κ (all)	1	0
ω_2 (Study 1)	−2.17	16
ω_2 (Study 2)	−2.16	16
ω_2 (Study 3)	−2.22	16
$\omega_3^{(0)}$ (all)	−7	0
$\mu_2^{(0)}$ (all)	0	0
$\sigma_2^{(0)}$ (all)	0.1	0
$\mu_3^{(0)}$ (all)	1	0
$\sigma_3^{(0)}$ (all)	1	0
ζ (all)	48	1

Free parameter ω_2 was estimated in its unbounded (linear) space. The prior values on ω_2 [mean (variance)] were: −2.17 (16), −2.16 (16), and −2.22 (16) for Studies 1, 2, and 3, respectively. These prior values were obtained using an ideal observer model that received the input that each participant had experienced. The response model parameter, ζ , was log-transformed, to allow for its estimation in an unbounded space. The remaining parameters were fixed and not estimated in each participant: $\sigma_2^{(0)}$, $\sigma_3^{(0)}$, κ , $\mu_2^{(0)}$, $\mu_3^{(0)}$. The coupling strength between levels 2 and 3 is κ , which was fixed to 1 (Hein et al., 2021). Among the fixed parameters, the following ones operate in their log-transformed space: $\sigma_2^{(0)}$, $\sigma_3^{(0)}$, κ , $\mu_2^{(0)}$. The prior variances are given in the space in which the parameters are typically estimated.

task performance (mIKI, RT, percError, percWin, rendFam) and computation variables (ω_2 , ζ , σ_2). Next, we evaluated between-group differences by computing Bayes Factors (BF) using the bayesFactor toolbox (<https://github.com/klabhub/bayesFactor>) in MATLAB. This toolbox implements tests that are based on multivariate generalizations of Cauchy priors on standardized effects (Rouder et al., 2012). For each dependent variable (DV), we calculated the BF on the model DV $\sim 1 + \text{group}$, where DV is explained by a fixed effect of group (HYA, HOA, PD). The model was fitted using the fitlme function of the MATLAB Statistics toolbox. Computing BF allowed us to quantify the evidence in support of the alternative hypothesis (full model, in our case assessing the main effect of the group) relative to the null model (intercept-only model, i.e., DV ~ 1). BF values were interpreted as in Andraszewicz et al. (2015). As BF is the ratio between the probability of the data being observed under the alternative hypothesis and the probability of the same data under the null hypothesis, a BF of 20 would indicate strong evidence for the alternative hypothesis. On the other hand, BF of 0.05 would provide strong evidence for the null hypothesis (for further details, see Andraszewicz et al., 2015, their Table 1). Accompanying the BF results, we provided the outcomes of standard one-way ANOVA for completion. In the case of main effects being observed in the group-level BF analysis, we conducted follow-up BF analyses on independent two-sample *t* tests.

When analyzing RT, we excluded outliers (RT values larger than three standard deviations above the mean) at the subject level. For BF analyses, we used the individual average across 180 trials for the mIKI, RT, and σ_2 variables. As mIKI and RT were not normally distributed, values were log-transformed (natural logarithm, log_mIKI and log_RT). The same preprocessing steps were applied to RT and mIKI values in Studies 2 and 3. The number of renditions during the familiarization phase was averaged between both types of sequence.

Sanity checks were performed to assess that participants chose to play each sequence as a function of the inferred action-reward contingencies and not based on individual sequence preferences. These were conducted by computing mean and SEM along with BF analyses for paired *t* tests on the percentage of trials each sequence type was (correctly) played (percPlayed; percCorrectlyPlayed; outcomes of standard paired *t* test reported for completion). We also report the group mean and SEM of percPlayed by contingency phases, which allowed us to observe whether participants' choices followed the changes in contingencies over time.

Assessing the association between predictions about the action-reward contingency and motor performance using Bayesian linear mixed models. Our main goal was to investigate whether trial-by-trial sequence performance tempo (mIKI) is modulated by the expectation about the tendency of the action-reward contingency ($\hat{\mu}_2$) in our participant

groups. In addition, we aimed to determine whether the group factor modulated the sensitivity of performance tempo to $\hat{\mu}_2$, resulting in different slopes of the association.

We addressed these questions by implementing a series of Bayesian linear mixed models (BLMMs) (R Core Team, 2022; version 4.0.3). We used the Bayesian regression models using Stan (brms; Bürkner, 2017, 2018, 2021) package, freely available on <https://cran.r-project.org/web/packages/brms/index.html>. Brms relies on the probabilistic programming language Stan, which implements Bayesian inference using Markov Chain Monte Carlo (MCMC) sampling methods to estimate approximate posterior probability distributions for model parameters.

In the HGF for binary categorical inputs, the sign of $\hat{\mu}_2$ (and similarly μ_2) is not informative, as it represents the tendency of an action-reward mapping for an arbitrary action (e.g., for seq1). Yet, we could similarly define the model in reference to the other action (e.g., seq2). In line with previous work (Stefanics et al., 2018; Hein and Herrojo Ruiz, 2022), we therefore took the absolute value of $\hat{\mu}_2$ ($|\hat{\mu}_2|$) for our analysis to represent the strength of predictions about the tendency of the action-reward mapping. Trials with greater $|\hat{\mu}_2|$ values are trials in which the participants will have a stronger expectation of receiving a reward, given they select the correct action. Thus, $|\hat{\mu}_2|$ represents the strength of the predictions. In one participant (HYA), we excluded $|\hat{\mu}_2|$ values of the last 27 trials, as the HGF trajectories diverged, despite the participant exhibiting normative learning patterns. Next, we centered the $|\hat{\mu}_2|$ values ($|\hat{\mu}_2|_c$) to allow the intercept estimate for mIKI to reflect the average $|\hat{\mu}_2|$ value. As for Bayesian ANOVAs (see General task performance and computational variables), mIKI was log-transformed to approach normality (log_mIKI). In one HOA participant, two log_mIKI values were discarded from the analyses as they were not registered correctly in the JSON file (i.e., represented an impossible value of mIKI ~ 50 ms).

In BLMM with brms, it is standard to select one group as reference for the parameter estimates. Brms then estimates the posterior distribution of parameter differences between each group and the reference group, as well as the posterior distributions of parameters in the reference group itself. We set HOA as the reference group, and therefore posterior distributions of between-group differences on response variables were assessed for HOA versus HYA and HOA versus PD.

We implemented six models of increasing complexity, with every model including a larger number of explanatory variables (Table 4). For simplicity, in the following we used variable label *y* to represent our dependent variable log_mIKI, and *x* to represent the explanatory variable $|\hat{\mu}_2|_c$. To answer our research questions, we primarily focused on: (1) the fixed effect of *x* [sensitivity (slope) of the performance tempo to the strength of predictions about the action-reward contingency in the reference group, HOA]; and (2) the interaction effect *x* * group [differences between groups in the sensitivity (slope) of the performance tempo to the strength of expectations about the action-reward mapping].

For each model we ran four independent chains with 5000 iterations each, of which the first 1000 were discarded as warmup. This resulted in a total of 16,000 posterior samples. In all models, we used a default prior distribution for the intercept, and a normal distribution for each fixed and random effect (fixed effects for group and *x*, normal [0,2]; interaction term group * *x*, normal [0,1]; random effects for intercept by subject and intercept by trial, normal [0,2]; random effect *x* by subject, normal [0,1]). The prior on the LKJ-Correlation, the correlation matrices in brms (Lewandowski et al., 2009), was set to 2 as recommended by Bürkner (2017). Chain convergence was assessed using the Gelman-Rubin statistics (R-hat < 1.1; Gelman and Rubin, 1992).

Models were compared using leave-one-out cross-validation of the posterior log-likelihood (LOO-CV) with Pareto-smoothed importance sampling (Vehtari et al., 2017). The identification of the best fitting model was based on the highest expected log point-wise predictive density (ELPD). We also checked that the absolute mean difference in ELPD between two models (elpd_diff in brms) exceeded twice the standard error of the differences (2*SE_diff). LOO-CV identified the most complex model (Table 4, model number 6) as the best fitting model (for further details, see Results). This model explained the performance tempo as the interaction between groups and the strength of the expectation about the action-reward contingency (in addition to main effects).

Table 4. Models of increasing complexity used for Bayesian linear mixed models analyses

Study #	Model #	Model
1–2	1	$y \sim 1 + (1 \text{subject})$
	2	$y \sim 1 + \text{group} + (1 \text{subject})$
	3	$y \sim 1 + \text{group} + x + (1 \text{subject})$
	4	$y \sim 1 + \text{group} * x + (1 \text{subject})$
	5	$y \sim 1 + \text{group} * x + (1 + x \text{subject})$
	6	$y \sim 1 + \text{group} * x + (1 + x \text{subject}) + (1 \text{trial})$
3	1	$y \sim 1 + (1 \text{subject})$
	2	$y \sim 1 + x + (1 \text{subject})$
	3	$y \sim 1 + x + (1 + x \text{subject})$
	4	$y \sim 1 + x + (1 + x \text{subject}) + (1 \text{trial})$

Models of increasing complexity used in Studies 1 and 2 (top) and Study 3 (bottom). In Studies 1 and 2, y corresponds to the motor performance (log_mIKI or log_RT); x is the unsigned centered value of the prediction about the tendency of the action-reward contingency ($|\hat{\mu}_2|_c$). This parameter represents the strength of the predictions. In model 1, y is explained by a fixed effect of the intercept and a random effect of intercept by subject (the latter accounts for repeated measurements); model 2 adds a fixed effect of group; model 3 includes the fixed effect of x , which allows to assess the sensitivity (slope) of performance tempo or RT to $|\hat{\mu}_2|_c$ in the reference group; model 4 incorporates the interaction term between group and x , which allows to investigate the between-group differences in the sensitivity (slope) of performance tempo or RT to $|\hat{\mu}_2|_c$; model 5 includes the random effect of $|\hat{\mu}_2|_c$ by subject; last, model 6 includes a random effect of intercept by trial. In Study 3, y corresponds to the motor performance (log_mIKI or log_RT); x is the confidence rating. In model 1, y is explained by a fixed effect of the intercept and a random effect of intercept by subject (the latter accounts for repeated measurements); model 2 adds a fixed effect of x , which allows to assess the sensitivity (slope) of performance tempo or RT to confidence ratings; model 3 includes the random effect of confidence ratings by subject; last, model 4 includes a random effect of intercept by trial.

Further, it modeled the effect of subjects on the intercept and $|\hat{\mu}_2|_c$ as a random effect, and the effect of trials on the intercept as a random effect. We reported for each parameter the posterior point estimate and the associated 95% credible interval (CI; for further details, see Results).

Because reward expectations could also modulate RT as shown previously (Codol et al., 2020), we conducted additional analyses to assess the effect of $|\hat{\mu}_2|$ on RT trial-by-trial. Further, we evaluated whether the group factor influences the sensitivity of RT to $|\hat{\mu}_2|$. In these analyses, we followed the same procedure as for the sequence performance tempo analysis. In particular, the associations between RT (log-transformed) and $|\hat{\mu}_2|_c$ were assessed by implementing and comparing six models of increasing complexity in brms (Table 4; for further details, see Results). RT values three standard deviations above the mean were excluded from statistical analyses. This approach was also followed in Studies 2 and 3. As for performance tempo, in Results, we use the variable label y for the dependent variable (log_RT) and x for $|\hat{\mu}_2|_c$.

Bayesian analyses on Study 2

As described above, in Study 2 participants were allocated to two different analysis groups (Q8_T and Q8_F) depending on their answer to a post-performance question (“I could *always* distinguish whether 0 points reflected a performance error or a bad decision”, binary answer: True/False). This allowed us to test the potential influence of subjective inferences about task-related reward assignment on the motor invigoration effect observed in Study 1. Specifically, we reasoned that participants who could not always infer the meaning of zero might show a reduced sensitivity of motor performance by beliefs about the reward tendency.

As for Study 1, we computed the mean and SEM as summary statistics for each dependent variable. Next, we used the bayesFactor toolbox to calculate the evidence in support of (or against) group differences in general task performance (mIKI, RT, percError, percWin) and computational variables (ω_2 , ζ , σ_2). We intentionally did not analyze the rate of sequence renditions during the familiarization phase as here we were only interested in assessing the role of subjective inferences about credit assignment on motor sequence performance decision-making behavior. We performed BF analysis on independent two-sample t tests to assess between group-differences on the variables of interest (results on standard independent t tests also reported for completion). RT and mIKI were log transformed and followed the same preprocessing steps as described for Study 1.

Next, to test potential between-group differences in the mIKI- $|\hat{\mu}_2|$ association, we implemented six BLMM of increasing complexity (same models as in Study 1; Table 4). As for Study 1, the most complex model (Table 4, model number 6) was identified as the best fit by LOO-CV (for further details, see Results). The same procedure was used to investigate the associations between RT with $|\hat{\mu}_2|$.

Finally, we evaluated whether Q8_T and Q8_F differed in the rate of retrospective subjective number estimate of performance errors. In particular, we were interested in assessing between-group differences in the tendency of under/overestimating the number of performance errors. For each participant, the rate of subjective performance execution errors (subjective_percError) was calculated through the post-performance questionnaire (see Questions 1, 2, 3; Table 2). We arbitrarily assigned a value of 0.028 (= 5/180) if subjects thought to have committed <10 performance errors; 0.111 (= 20/180) for between 20 and 40 estimated performance errors; 0.222 (= 40/180) for >40 subjective performance errors. To assess whether this rough estimate of the percentage of performance errors reflected a general over or underestimation of the true performance error rate in the total sample ($N = 39$), we first conducted a BF analysis on the correlation between the subjective and empirical error rates (Pearson’s r coefficient and p -value reported for completion). Next, we identified potential group-related systematic biases in the subjective estimate. This was done with a BF analysis using independent two-sample t tests on the normalized rate of subjective errors [(subjective_percError-percError)/percError; results on standard independent t tests reported for completion].

Bayesian analyses on Study 3

In Study 3, we aimed at assessing the association between trial-by-trial explicit beliefs about the reward tendency (confidence ratings) and motor performance. We were particularly interested in understanding whether being more certain (following Frömer et al., 2021) about obtaining the reward, given the right choice, would speed up motor responses.

First, following the same steps as for Studies 1 and 2, we calculated the mean and SEM as summary statistics for the general task performance variables (mIKI, RT, percWin, conf). Trial-by-trial confidence ratings were converted to a 0–0.99 scale.

We aimed to use the confidence rating as a predictor in our BLMM analyses to assess the sensitivity of motor performance (mIKI and RT) to explicit beliefs about the reward tendency. This was tested by implementing four BLMM of increasing complexity (Table 4).

As for Studies 1 and 2, we used the label y to represent our dependent variable (mIKI or RT), and x for the explanatory variable (conf). To test our hypothesis, we specifically focused on the fixed effect of x [sensitivity (slope) of the motor performance to the confidence ratings about the predicted outcome]. We used the same priors as in Study 1 for the corresponding factors. The most complex model number four and the model number 3 (Table 4) were identified as the best fit by LOO-CV for performance tempo and RT, respectively (for further details, see Results).

In addition, as a sanity check, we evaluated the association of confidence ratings with the strength of predictions about the action-reward contingency trial-by-trial. The investigation of motor vigor effects in Studies 1 and 2 assumed that the unsigned $|\hat{\mu}_2|$ values estimated in the HGF reflect the strength of participants’ expectation on the reward tendency. However, whether this HGF quantity reflects true explicit beliefs, assessed as confidence ratings, is not clear. We evaluated the association between confidence ratings and the unsigned $|\hat{\mu}_2|$ values using the formula $\text{conf} \sim 1 + |\hat{\mu}_2|_c + (1 + |\hat{\mu}_2|_c|\text{subj}) + (1|\text{trial})$ in brms. We chose a default prior distribution for the intercept, and a normal distribution for the fixed and random effects [fixed effect for $|\hat{\mu}_2|_c$, normal [0,2]]; random effects for intercept by subject and intercept by trial, normal [0,2]; random effect $|\hat{\mu}_2|_c$ by subject, normal [0,1]). The prior on the LKJ-Correlation was set to 2 as recommended by Bürkner (2017).

Finally, we provided summary statistics for the number of empirical performance errors and the number of subjective performance errors (how many times the “z” key was pressed throughout the experiment). This aimed at expanding on the findings of Study 2, informing about participants’ ability to correctly identify performance errors and thus infer the task-related credit assignment.

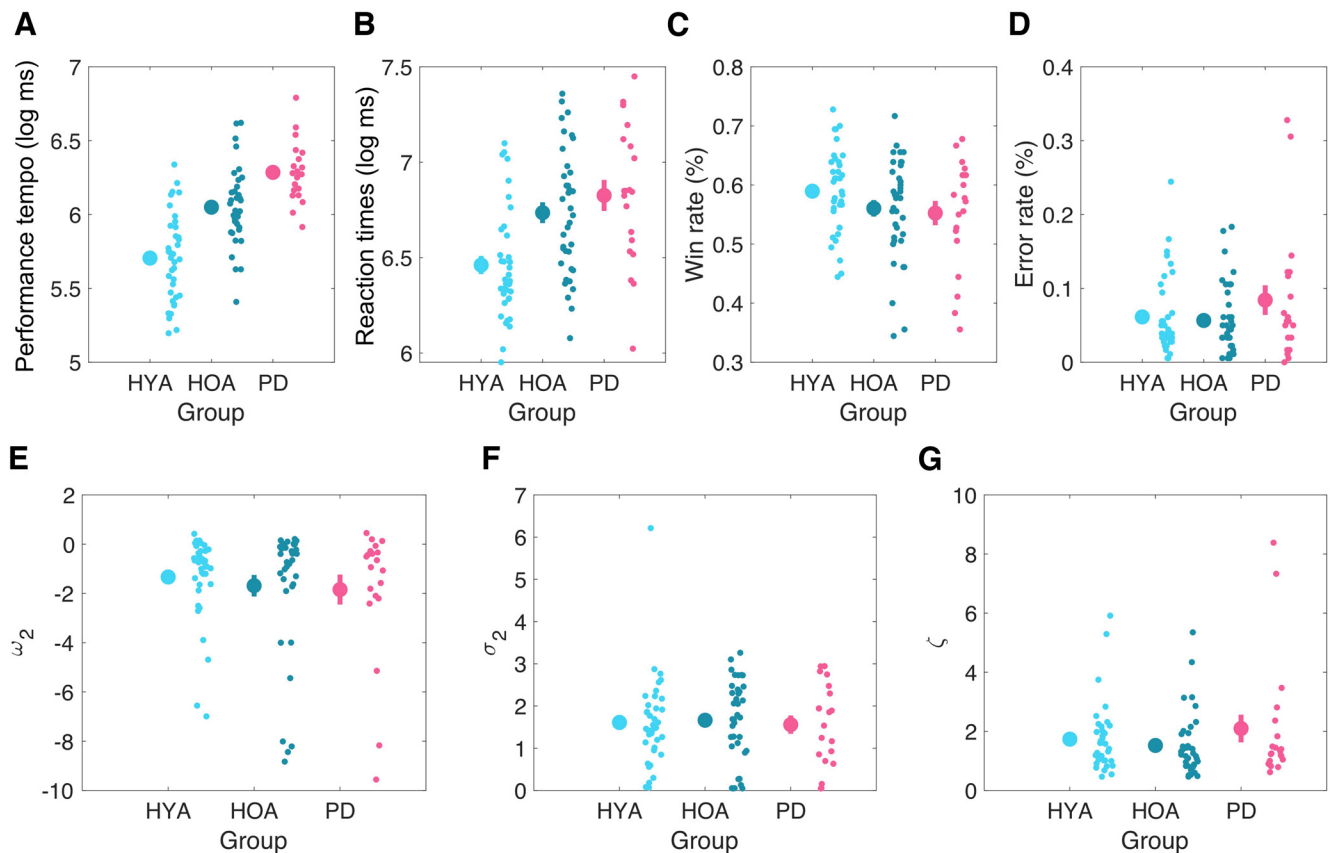


Figure 3. Markers of general task performance and decision-making across groups. Data presented for healthy younger adults (HYA; in light blue), healthy older adults (HOA; in dark blue), and patients with Parkinson's disease (PD; in purple) in Study 1. **A**, Performance tempo (mIKI, mean interkeystroke-interval, in milliseconds). **B**, Reaction time (RT; in milliseconds). **C**, Rate of win trials (percWin). **D**, Rate of performance execution errors (percError). **E**, Tonic volatility (ω_2). **F**, Informational uncertainty on level 2 (σ_2). **G**, Response model parameter (ζ). Values mIKI, RT and σ_2 are averaged across 180 trials within each participant. mIKI and RT values are log-transformed. In every plot, to the right of each mean (large dot) and SEM (denoted by the vertical bar), the individual data points in each group are shown to visualize group population variability.

Data and code availability

The data that support the main findings of these studies are available from the Open Science Framework Data Repository under the accession code 7kfbj (<https://osf.io/7kfbj/>).

Results

Study 1

Task validation. Participants played on average seq1 and seq2 50% of the trials (seq1: mean 0.490, SEM 0.008; seq2: mean 0.508, SEM 0.008). This suggests that they did not express a preference toward a sequence type (percPlayed, BF = 0.2295, moderate evidence in support of the null hypothesis for no differences in the percentage of performances by sequence type, $t_{(93)} = -1.204$, $p = 0.232$). Participants committed fewer performance execution errors in seq1 (mean 0.958, SEM 0.005) than seq2 (mean 0.922, SEM 0.008; percCorrectlyPlayed, BF = 1126.7, suggesting extreme evidence for alternative hypothesis that the rate of correct performance differed in seq1 and seq2, $t_{(93)} = 4.576$, $p < 0.001$). Next, we observed that percPlayed in each group successfully tracked the contingency changes over time. For true contingencies sorted according to increasing values, [0.1, 0.3, 0.5, 0.7, 0.9], HYA participants played the corresponding sequence at these rates: [0.18 (0.02), 0.33 (0.02), 0.48 (0.02), 0.67 (0.02), 0.81 (0.02)]. Similar values were obtained for HOA participants: [0.18 (0.02), 0.34 (0.02), 0.48 (0.02), 0.62 (0.02), 0.79 (0.02)]; and for PD patients: [0.16 (0.02), 0.32 (0.03), 0.47 (0.03), 0.63 (0.03), 0.79 (0.03)]. Accordingly, task performance demonstrated that each

group of participants learned to flexibly adapt to the changing contingencies over time.

General task performance. Overall, as expected, our analyses revealed between-group differences in performance tempo (mIKI in milliseconds, HYA: mean 300, SEM 15.8; HOA: mean 424, SEM 19.6; PD: mean 537, SEM 26.9; Fig. 3A), and reaction time (RT in milliseconds, HYA: mean 634, SEM 34.9; HOA: mean 838, SEM 49.4; PD: mean 918, SEM 77.5; Fig. 3B), with movements progressively slowing down in aging and PD patients. BF analyses on performance tempo yielded extreme evidence for a group effect (log_mIKI: BF = 1.1253e + 09, demonstrating extreme evidence for the alternative hypothesis; $F_{(2,91)} = 35.332$, $p < 0.001$). *Post hoc* pairwise *t* tests using BF showed extreme evidence for between-group differences in HYA versus HOA (BF = 1.2044e + 04) and in HYA versus PD (BF = 3.3592e + 07). We also found very strong evidence for the alternative hypothesis in HOA versus PD (BF = 32.591). Thus, performance tempo (and therefore movement time) was differently modulated between groups, with HYA being faster than HOA and PD, and HOA faster than PD. Regarding RT, there was extreme evidence supporting between-group differences (log_RT: BF = 404.521; $F_{(2,91)} = 11.383$, $p < 0.001$). BF analysis on *post hoc* independent two-sample *t* tests revealed extreme evidence for between-group differences in HYA versus HOA (BF = 109.444) and HYA versus PD (BF = 239.335). Yet, we only found anecdotal evidence in support of the null hypothesis in HOA versus PD (BF = 0.403).

Hence, despite HYA displaying shorter RTs than HOA and PD, our analyses suggest similar RTs in HOA and PD.

In addition, we found anecdotal evidence supporting that groups differed in the number of sequence renditions during the familiarization phase (rendFam, HYA: mean 5.6, SEM 0.1; HOA: mean 6.0, SEM 0.2; PD: mean 7.1, SEM 0.8; BF = 1.733; $F_{(2,91)} = 4.448$, $p = 0.014$). *Post hoc* BF analyses to assess differences between pairs of groups revealed anecdotal and moderate evidence for between-group differences in HYA and HOA (BF = 1.900) and HYA and PD (BF = 3.030), respectively. Still, HOA and PD practiced the two sequences to a similar extent (BF = 0.853, revealing anecdotal evidence for the null hypothesis). Of note, practicing more during familiarization was not associated with better win rates or average performance tempo during task completion. A correlation analysis across all participants between the number of repetitions during familiarization and these variables demonstrated moderate evidence for null correlation effects (percWin: BF = 0.290; Pearson $r = -0.134$, $p = 0.200$; log_mIKI: BF = 0.397; Pearson $r = 0.158$, $p = 0.131$; note that we excluded one PD patient who practiced 21 times during familiarization as outlier in this correlation analysis).

The group effects observed above were not accompanied by a dissociation between groups in the win rate or the rate of performance execution errors (Fig. 3C,D). BF analysis on win rates provided moderate evidence for the lack of a group effect (percWin, HYA: mean 0.590, SEM 0.012; HOA: mean 0.561, SEM 0.014; PD: mean 0.553, SEM 0.021; BF = 0.210, supporting moderate evidence for the null hypothesis; $F_{(2,91)} = 1.848$, $p = 0.163$). A similar outcome was observed in the analysis of performance execution error rates (percError, HYA: mean 0.061, SEM 0.009; HOA: mean 0.057, SEM 0.008; PD: mean 0.084, SEM 0.020; BF = 0.146, moderate evidence for the null hypothesis; $F_{(2,91)} = 1.456$, $p = 0.239$). In sum, we found moderate evidence that HYA, HOA, and PD did not differ in either the rate of win or error trials.

Computational parameters. Decision-making was assessed by looking at between-group differences in the computational variables ω_2 , ζ , and σ_2 . After excluding the HGF₃ from model comparison because of numerical instabilities, BMS was conducted on the HGF₂ and two reinforcement learning models (RW, SK1) using the individual log-model evidence (LME) values provided by the HGF toolbox. The winning model was the HGF₂, with an exceedance probability of 0.95 and an expected frequency of 0.90. Of note, although the HGF₃ model was not included in BMS, a qualitative comparison of LME values for the HGF₃ and HGF₂ models in the 80% participants in which HGF₃ did not lead to numerical instabilities revealed extremely similar values (LME differences < 1). This observation suggested that both models described behavior in our task with constant true volatility to a similar degree.

Overall, we found no group effect on the signatures of reward-based learning and decision-making in our volatile task (Fig. 3E–G). BF analysis on ω_2 demonstrated strong evidence for the absence of a main effect of group (HYA: mean -1.332 , SEM 0.282; HOA: mean -1.686 , SEM 0.438; PD: mean -1.843 , SEM 0.609; BF = 0.059; $F_{(2,91)} = 0.380$, $p = 0.685$). Similarly, we found strong evidence in favor of a lack of group effect on the informational uncertainty about beliefs on the tendency of the action-reward contingency, σ_2 (HYA: mean 1.610, SEM 0.177; HOA: mean 1.663, SEM 0.158; PD: mean 1.559, SEM 0.218; BF = 0.045; $F_{(2,91)} = 0.074$, $p = 0.928$). Last, groups exhibited a similar mapping from beliefs to responses, driven by the response model parameter ζ (HYA: mean 1.735, SEM

0.191; HOA: mean 1.523, SEM 0.176; PD: mean 2.095, SEM 0.469; BF = 0.114, demonstrating moderate evidence for the null hypothesis; $F_{(2,91)} = 1.1495$, $p = 0.321$).

A direct comparison between the Italian HOA subsample and (Italian) PD sample revealed anecdotal or moderate evidence in support of the null hypothesis when assessing general performance and decision-making variables (exception for log_mIKI). These findings thus converge with the outcomes of the full HOA sample analysis. On the other hand, the very strong evidence in support of group effects on the performance tempo in the full sample was only anecdotal when directly comparing Italian HOA and PD samples on this variable (log_mIKI: BF = 2.556; $t_{(42)} = -2.348$, $p = 0.024$). These results suggested that Italian healthy aging was associated with slower performance tempo relative to United Kingdom healthy aging participants (log_mIKI: BF = 6.637; $t_{(35)} = 2.871$, $p = 0.007$; moderate evidence supporting differences in performance tempo). Hence, between-group effects on general task performance and decision-making cannot be accounted for by language differences.

Sensitivity of motor performance to the strength of expectations about the action-reward contingency. For performance tempo, LOO-CV identified the most complex model (model number 6) as the best fit. The absolute mean difference in ELPD between the winning model and the second best fitting model (elpd_diff) was -665.8557 and the standard error of the differences (SE_diff) equals 39.0404 (elpd_diff > 2*SE_diff). When ELPD differences between two models are larger than four, and also if the number of observations is > 100, and the model is moderately well specified, then the standard error is a good estimate of the uncertainty in the difference between models (Vehtari et al., 2017; Sivula et al., 2022). Posterior predictive checks revealed that the best model had strong predictive power for the range of the DV (Fig. 4A). In the following we use variable label y to represent our dependent variable log_mIKI (in log-ms), and x to represent the explanatory variable $|\hat{\mu}_2|_{c.c.}$. Table 5 presents a summary of the posterior distributions for the winning model.

First, we found that groups differed in performance tempo, as expected. This is in line with our previous between-group analyses showing a progressive slowness in execution tempo in HOA and PD. The posterior estimate for the intercept in the reference group, HOA, was 6.00, CI = [5.91, 6.09] (in milliseconds, 404, CI = [368, 443]). The distribution of the differences between intercepts in HOA and HYA had a posterior estimated value of -0.34 , CI = $[-0.47, -0.21]$ (in milliseconds, -116 , CI = $[-163, -70]$), while the distribution of the differences between intercepts in HOA and PD yielded a posterior point estimate of 0.25, CI = [0.09, 0.41] (in milliseconds, 114, CI = [41, 192]). As neither of the two distributions overlapped with zero, we concluded that HYA performed the sequences faster than HOA, while PD was slower than HOA (Fig. 4B).

Next, we evaluated how the strength of predictions about the action-reward contingency modulated performance tempo on a trial-by-trial basis. The analyses supported our hypothesis, showing that stronger expectations about the reward contingency invigorated motor performance through faster execution tempo. Here, we focused on the distribution of the fixed effect of x (slope of the association between y and x) in the reference group, HOA. This distribution informs about the sensitivity of the performance tempo to the strength of predictions about the action-reward contingency in HOA. The posterior estimate of x was equal to -0.04 , CI = $[-0.07, -0.01]$. As the distribution did not include zero, this highlights a negative relationship

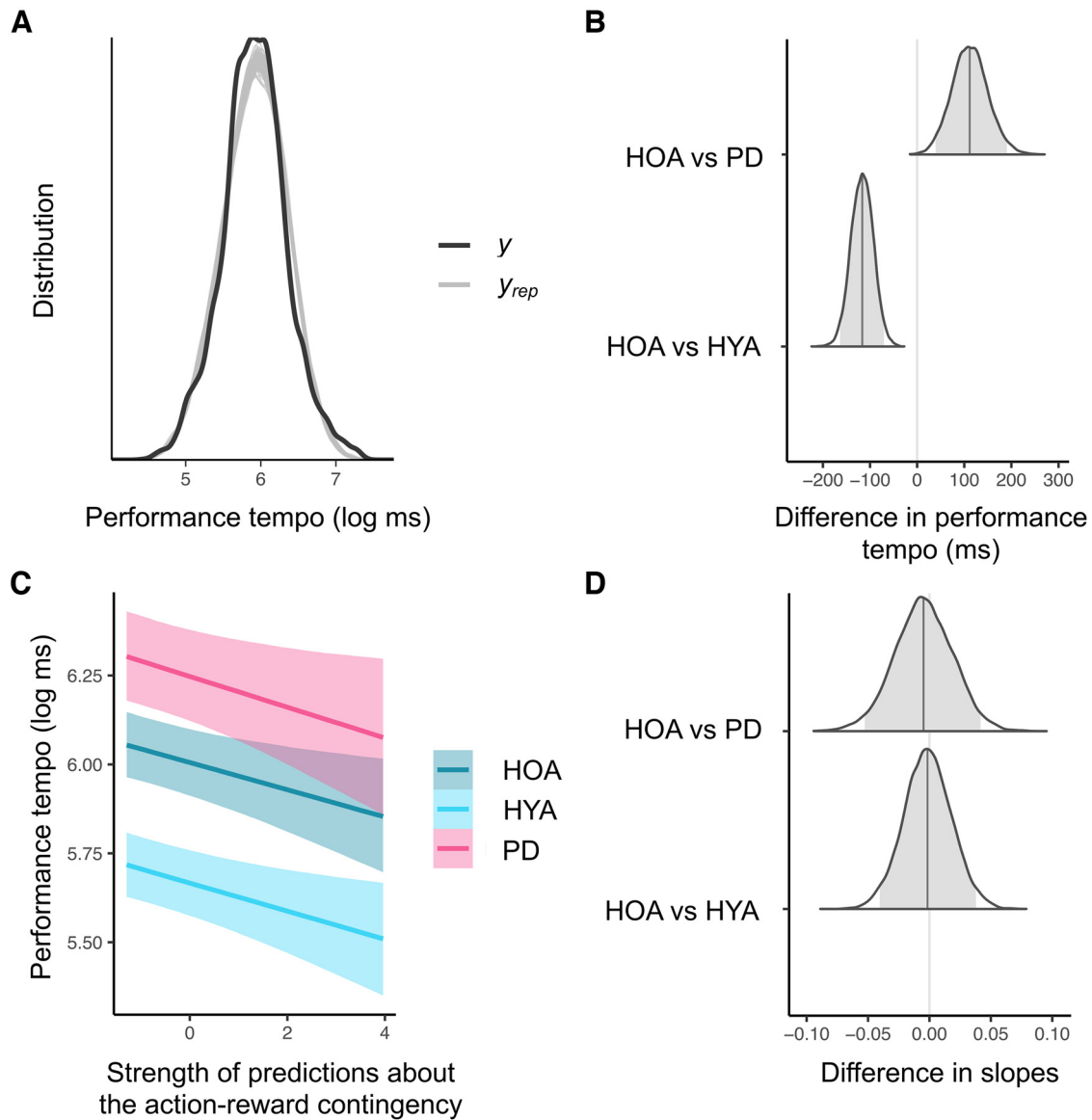


Figure 4. Invigoration of performance tempo by beliefs is preserved in healthy aging and in Parkinson's disease. Bayesian linear mixed model [BLMM; model number 6, $y \sim 1 + \text{group} * x + (1 + x|\text{subject}) + (1|\text{trial})$] with healthy older adults (HOA) as the reference group in Study 1. **A**, Illustration of the posterior predictive checks where the distribution of the observed outcome variable (y , in our case performance tempo) is compared with simulated datasets (y_{rep}) from the posterior predictive distribution (100 draws). **B**, Distributions of the difference in milliseconds between performance tempo (intercept) in HOA and healthy younger adults (HYA), and in HOA and patients with Parkinson's disease (PD). For each distribution, the gray vertical bar indicates the posterior point estimate, while the gray area under the curve represents the 95% credible interval (CI). In the current plot, CIs do not overlap with zero (the null hypothesis). This indicates that there is a 95% probability of between-group differences in performance tempo. **C**, Results of the BLMM analysis. We analyzed how the strength of predictions about the action-reward contingency modulates performance tempo separately for HYA (in light blue), HOA (in dark blue), and PD (in purple). Here, mIKI (performance tempo: mean interkeystroke-interval) values are represented in the log-scale. The negative slopes suggest that stronger predictions about the action-reward contingency are associated with faster performance tempo. **D**, Distributions of the difference between slopes in HOA versus HYA, and HOA versus PD. Here, as CIs include zero, we can conclude with 95% probability that groups do not differ in how the strength of predictions about the reward contingency influences motor performance tempo. Thus, the sensitivity of performance tempo to the strength of predictions about the reward mapping is not differently modulated between groups.

between performance tempo and the strength of expectations about the action-reward contingency in the reference group (Fig. 4C).

We were also interested in evaluating between-group differences in the sensitivity of performance tempo to the strength of expectations about the action-reward contingency. This was conducted by assessing the distribution of the interaction effect $\text{group} * x$ on the slope. Both the posterior distributions of slope differences between HOA and HYA and between HOA and PD overlapped with zero, suggesting that the sensitivity was similar between groups (HOA vs HYA: posterior estimate = -0.00 , CI =

$[-0.04, 0.04]$; HOA vs PD: posterior estimate = -0.00 , CI = $[-0.05, 0.04]$; Fig. 4D).

Overall, our BLMM analysis demonstrated that motor performance tempo was influenced trial-by-trial by the strength of predictions about the tendency of the action-reward contingency, with stronger expectations leading to faster execution tempo. However, the sensitivity of performance tempo to the strength of these predictions was not differently modulated between groups, suggesting that all groups could successfully use the inferred predictions to invigorate their motor performance to a similar degree.

Table 5. Summary of the posterior distributions for the fixed effects of the best fitting Bayesian linear mixed models

Study #	Dependent variable	Fixed effect	Estimate	I-95% CI	u-95% CI	R-hat		
1	Performance tempo	y: HOA	6.00	5.91	6.09	1.00		
		y: HOA vs HYA	−0.34	−0.47	−0.21	1.00		
		y: HOA vs PD	0.25	0.09	0.41	1.00		
		x: HOA	−0.04	−0.07	−0.01	1.00		
		group * x: HOA vs HYA	−0.00	−0.04	0.04	1.00		
		group * x: HOA vs PD	−0.00	−0.05	0.04	1.00		
	Reaction times	y: HOA	6.65	6.54	6.75	1.01		
		y: HOA vs HYA	−0.28	−0.42	−0.13	1.00		
		y: HOA vs PD	0.09	−0.08	0.27	1.00		
		x: HOA	−0.02	−0.04	0.01	1.00		
		group * x: HOA vs HYA	−0.01	−0.05	0.03	1.00		
		group * x: HOA vs PD	−0.03	−0.07	0.02	1.00		
	2	Performance tempo	y: Q8 _T	5.62	5.51	5.72	1.00	
			y: Q8 _T vs Q8 _F	0.07	−0.11	0.25	1.01	
x: Q8 _T			−0.04	−0.06	−0.01	1.00		
group * x: Q8 _T vs Q8 _F			−0.00	−0.04	0.04	1.00		
Reaction times		y: Q8 _T	6.24	6.13	6.34	1.00		
		y: Q8 _T vs Q8 _F	−0.01	−0.19	0.18	1.00		
		x: Q8 _T	−0.02	−0.04	0.002	1.00		
		group * x: Q8 _T vs Q8 _F	0.01	−0.03	0.04	1.00		
		3	Performance tempo	y	5.82	5.73	5.91	1.00
				x	−0.04	−0.08	−0.001	1.00
Reaction times	y		6.47	6.37	6.58	1.00		
	x		−0.10	−0.20	0.01	1.00		

Estimates, credible intervals (CIs) and R-hat values for the fixed effects of the best fitting models in Studies 1 and 2 [model number 6: $y \sim 1 + \text{group} * x + (1 + x|\text{subject}) + (1|\text{trial})$] and in Study 3 [model number 4: $y \sim 1 + x + (1 + x|\text{subject}) + (1|\text{trial})$] for performance tempo; model number 3: $y \sim 1 + x + (1 + x|\text{subject})$ for reaction times]. In Study 1, y: HOA refers to the posterior estimate for the intercept in the reference group (healthy older adults, HOA). y: HOA versus HYA and y: HOA versus PD reflect the posterior distributions of the differences between intercepts [HOA vs healthy younger adults (HYA); HOA vs Parkinson's patients (PD), respectively]. x: HOA is the posterior distribution of the association (slope) between motor performance (either performance tempo or reaction times) and the strength of predictions about the action-reward contingency in the reference group. group * x: HOA versus HYA and group * x: HOA versus PD are the posterior distributions of slope differences between HOA and HYA and between HOA and PD, respectively. In Study 2, y: Q8_T refers to the posterior estimate for the intercept in the reference group (participants that replied True to Question 8, Q8_T). y: Q8_T versus Q8_F reflects the posterior distribution of the difference between intercepts [Q8_T vs participants that replied False to Question 8 (Q8_F)]. x: Q8_T is the posterior distribution of the association (slope) between motor performance (either performance tempo or reaction times) and the strength of predictions about the action-reward contingency in the reference group. The upper bound of the CI for the slope effect in the BLMM analyses for RT is given with three decimal digits to demonstrate that 0 was included in the 95% CI. group * x: Q8_T versus Q8_F is the posterior distribution of slope difference between Q8_T and Q8_F. In Study 3, y refers to the posterior estimate for the intercept. x is the posterior distribution of the association (slope) between motor performance (either performance tempo or reaction times) and the confidence ratings. The upper bound of the 95% CI estimate of the slope effect in the BLMM analyses for performance tempo was −0.001, when considering three decimal digits. In all studies, I-95% CI and u-95% CI refer to the lower and upper bound of the credible intervals of the posterior distributions of the fixed effects. For each parameter, we also reported the corresponding Gelman–Rubin statistics (R-hat values). Values <1.1 indicate chain convergence (Gelman and Rubin, 1992).

In a separate analysis, we determined whether the motor inigation effect extended to the RT, reflecting the time to initiate the sequence performance (first key press). As for performance tempo, LOO-CV identified model 6 as the best fit (elpd_diff = −378.2718, SE_diff = 30.69 148; elpd_diff > 2*SE_diff) and posterior predictive checks demonstrated good predictive power for the range of the DV albeit less so than for performance tempo (Fig. 5A). On the other hand, Gelman–Rubin statistics (R-hat values) demonstrated an excellent chain convergence. Table 5 presents a summary of the posterior distributions for the winning model.

Our brms analysis on the best fitting model revealed shorter RT in HYA compared with HOA, with no differences emerging between HOA and PD. The posterior point estimate for the intercept in the reference group, HOA, was 6.65, CI = [6.54, 6.75] (in milliseconds, 771, CI = [693, 856]). The distribution of the differences between intercepts in HOA and HYA was centered at −0.28, CI = [−0.42, −0.13] (in milliseconds, −188, CI = [−289, −88]), which did not overlap with zero. On the other hand, the distribution of the differences between intercepts in HOA and PD yielded a posterior point estimate of 0.09, CI = [−0.08, 0.27] (in milliseconds, 77, CI = [−65, 231]) and included zero (Fig. 5B). These results demonstrated that HYA initiated the sequence faster than HOA, consistent with our mIKI group results, whereas PD and HOA had a similar RT intercept.

Regarding the association between the strength of predictions about the action-reward contingency and RT, we observed no trial-by-trial modulation and no group effects. The distribution of the fixed effect of x (slope of the association between y and x in the reference group, HOA) had a posterior point estimate of −0.02, CI [−0.04, 0.01]. As the distribution's center overlapped with zero, this demonstrates that the strength of predictions about the action-reward contingency did not modulate RT in this group (Fig. 5C). Potential between-group differences in the slope were assessed by investigating the distribution of the interaction effect group * x. Both the posterior distributions of slope differences between HOA and HYA and between HOA and PD included zero (HOA vs HYA: posterior estimate = −0.01, CI = [−0.05, 0.03]; HOA vs PD: posterior estimate = −0.03, CI = [−0.07, 0.02]; Fig. 5D). This outcome supported that the sensitivity of RT to the strength of expectations about the reward mapping did not differ between groups. Thus, the strength of predictions about the action-reward contingency invigorated performance tempo on a trial-by-trial basis without affecting the RT.

Study 2

Subjective inference about task-related reward assignment

We conducted Bayesian analyses on the HYA sample of Study 2 to evaluate whether subjective inferences about the hidden causes for the absence of reward could modulate the motor inigation effect observed in Study 1.

Overall, our analyses provided anecdotal and moderate evidence for the lack of differences between Q8_T and Q8_F in the main markers of general task performance (log_mIKI: BF = 0.417; $t_{(37)} = -0.795, p = 0.432$; log_RT: BF = 0.329; $t_{(37)} = 0.156, p = 0.877$; percWin: BF = 0.408; $t_{(37)} = 0.758, p = 0.453$; percError: BF = 0.596; $t_{(37)} = -1.252, p = 0.219$; for summary statistics, see Fig. 6A–D).

Random effects Bayesian model selection yielded substantially greater evidence in favor of model HGF₂ (exceedance probability 0.94, and expected frequency 0.68). Using this model to characterize decision-making processes in Q8_T and Q8_F samples, we observed that a BF analysis on ω_2, ζ and σ_2 provided anecdotal evidence for the absence of a group effect (ω_2 : BF = 0.560; $t_{(37)} = -1.183, p = 0.244$; ζ : BF = 0.445; $t_{(37)} = 0.895, p = 0.377$; σ_2 : BF = 0.463; $t_{(37)} = -0.951, p = 0.348$; for summary statistics, see Fig. 6E–G).

Hence, whether participants were *always* certain (Q8_T) or not (Q8_F) of the implications of receiving zero points, their general motor sequence performance and decision-making behavior seemed similar, albeit this interpretation is based on anecdotal evidence.

We further investigated whether not being *always* sure about the causes for the lack of reward could impact the sensitivity of

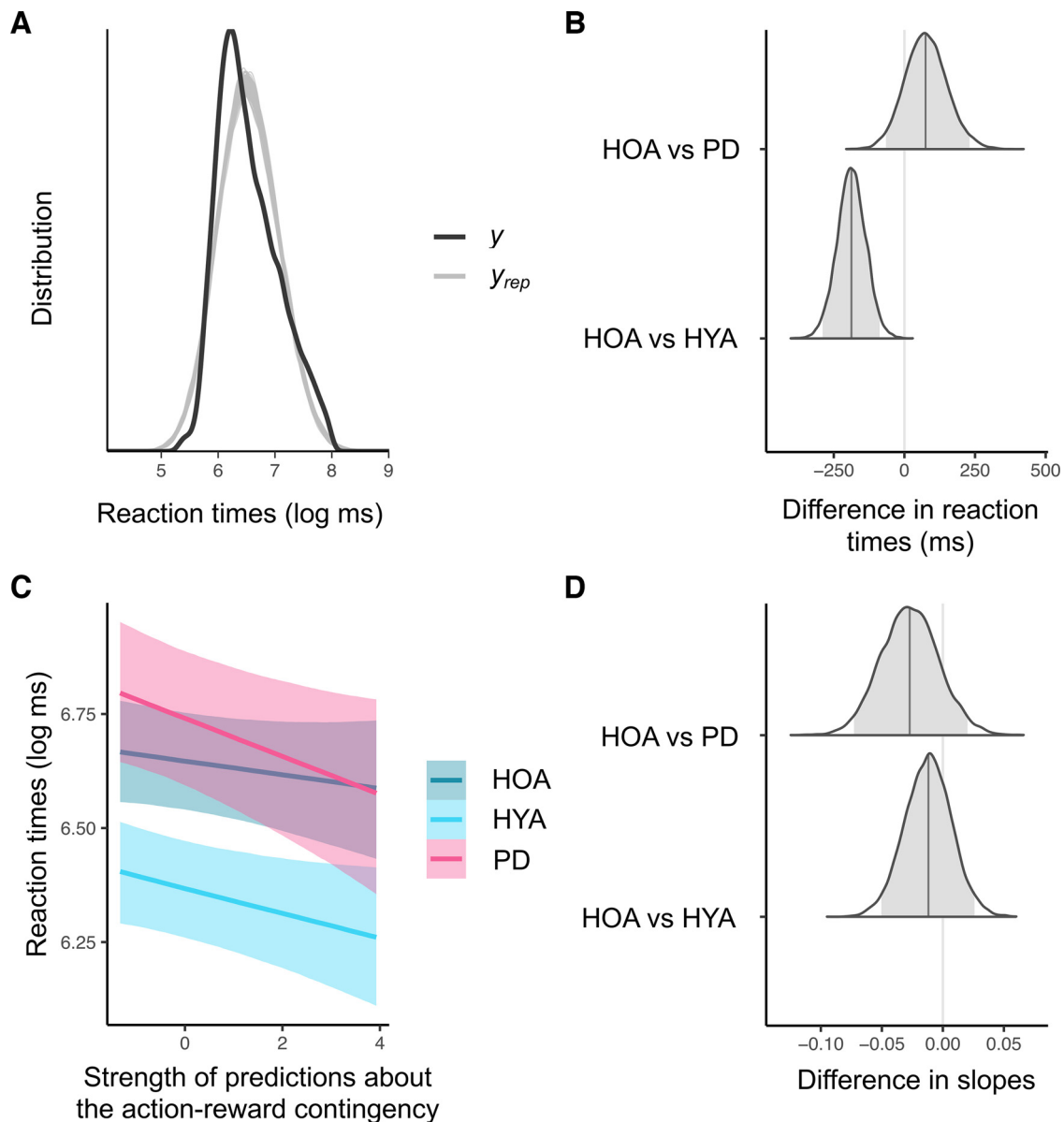


Figure 5. Motor vigor effects on reaction times across healthy young, older, and Parkinson's participants. Bayesian linear mixed model [BLMM; model number 6, $y \sim 1 + \text{group} * x + (1 + x|\text{subject}) + (1|\text{trial})$] with healthy older adults (HOA) as the reference group in Study 1. **A**, Illustration of the posterior predictive checks where the distribution of the observed outcome variable y , in our case reaction times (RT) is compared with simulated datasets (y_{rep}) from the posterior predictive distribution (100 draws). **B**, Distributions of the difference in milliseconds between RT (intercept) in HOA and healthy younger adults (HYA), and in HOA and patients with Parkinson's disease (PD). For each distribution, the gray vertical bar indicates the posterior point estimate, while the gray area under the curve represents the 95% credible interval (CI). In the current plot, CI of the bottom distribution does not overlap with zero (the null hypothesis). This indicates that there is 95% probability of between-group differences in RT. On the other hand, the distribution at the top includes zero. This suggests that there is 95% probability of HOA and PD not differing in RT. **C**, Results of the BLMM analysis. We analyzed how the strength of predictions about the action-reward contingency modulates RT separately for HYA (in light blue), HOA (in dark blue), and PD (in purple). Here, RT values are represented in the log-scale. We found no modulation of RT by the strength of expectations about the reward mapping. **D**, Distributions of the difference between slopes in HOA versus HYA, and HOA versus PD. Here, as CIs include zero, we can conclude with 95% probability that groups do not differ in how the strength of predictions about the reward contingency influences RT. Thus, the sensitivity of RT to the strength of predictions about the reward mapping is not differently modulated between groups.

motor performance (mIKI and RT) to the strength of predictions about the action-reward contingency. As for the main experiment, LOO-CV identified the most complex model (model number 6) as the best fit (mIKI, $\text{elpd_diff} = -144.9434$, $\text{SE_diff} = 20.33661$; $\text{elpd_diff} > 2 * \text{SE_diff}$; RT, $\text{elpd_diff} = -106.3677$, $\text{SE_diff} = 17.4019$; $\text{elpd_diff} > 2 * \text{SE_diff}$). Table 5 presents a summary of the posterior distributions for the winning models.

For performance tempo, the posterior predictive checks demonstrated a very strong predictive power for the range of DV values in the best model (Fig. 7A). Consistent with our previous BF

analyses on mIKI, the distribution of the differences between intercepts in Q8_T and Q8_F overlapped with zero, suggesting that subjective inferences about credit assignment did not impact performance tempo (Fig. 7B). BLMM analyses also revealed a negative association (slope) between the strength of predictions about the action-reward contingency and performance tempo. This replicates our findings in Study 1, showing that stronger predictions about the reward contingencies are followed by faster execution tempo (Fig. 7C). Yet, no between-group slope differences were observed. Thus, subjective inferences about the causes for the absence of reward did not modulate the sensitivity of

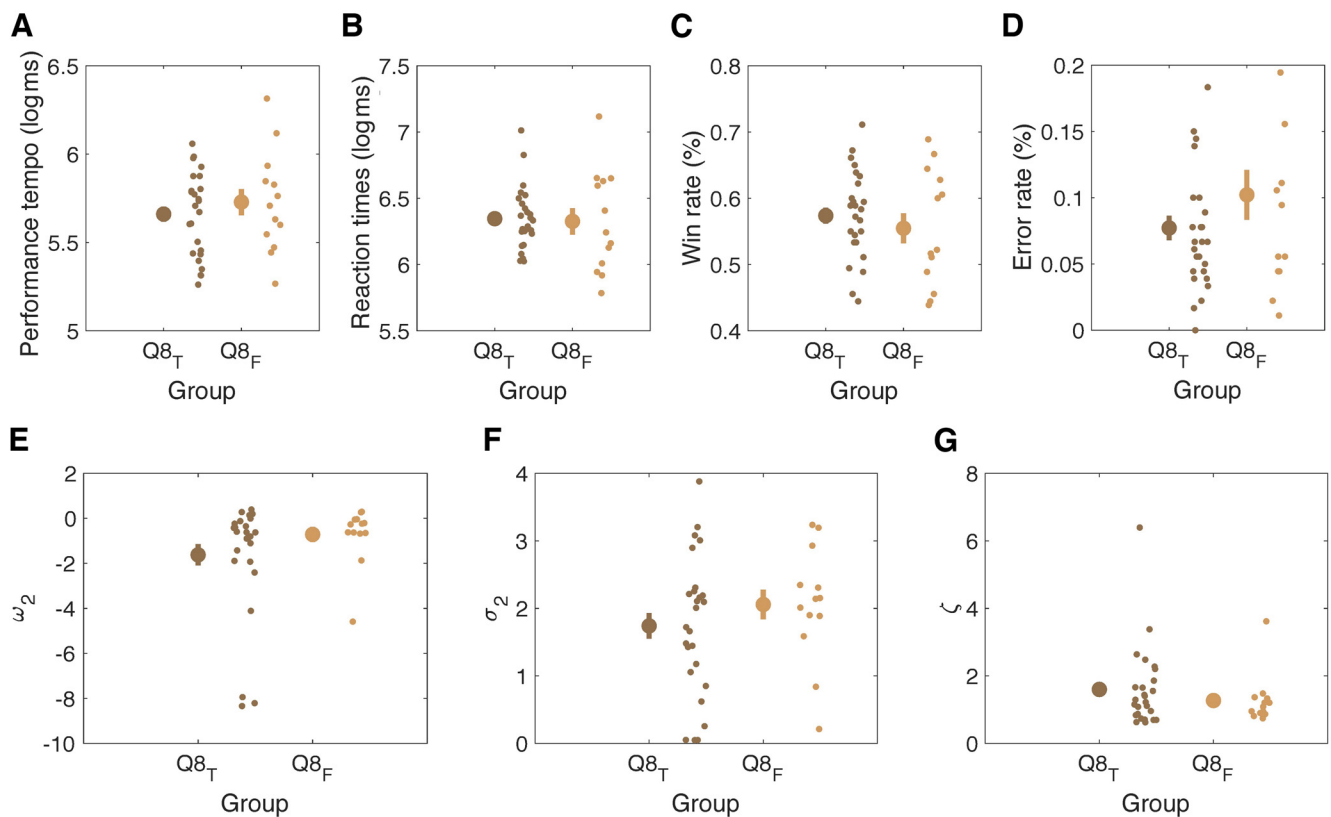


Figure 6. Effect of retrospective credit assignment on general task performance and decision-making. Markers of general task performance and decision-making in participants that replied True to Question 8 (Q8_T; in dark brown) and participants that replied False to Question 8 (Q8_F; in light brown) in the post-performance questionnaire (see Table 2) in Study 2. **A**, Performance tempo (mIKI, mean interkeystroke-interval; in milliseconds, Q8_T: mean 287, SEM 13.2; Q8_F: mean 307, SEM 27.2). **B**, Reaction times (RT; in milliseconds, Q8_T: mean 564, SEM 30.5; Q8_F: mean 555, SEM 68.7). **C**, Rate of win trials (percWin; Q8_T: mean 0.574, SEM 0.013; Q8_F: mean 0.555, SEM 0.024). **D**, Rate of performance execution errors (percError; Q8_T: mean 0.077, SEM 0.010; Q8_F: mean 0.102, SEM 0.020). **E**, Tonic volatility (ω_2 ; Q8_T: mean -1.624 , SEM 0.510; Q8_F: mean -0.715 , SEM 0.357). **F**, Informational uncertainty on level 2 (σ_2 ; Q8_T: mean 1.740, SEM 0.203; Q8_F: mean 2.057, SEM 0.237). **G**, Response model parameter (ζ ; Q8_T: mean 1.599, SEM 0.237; Q8_F: mean 1.271, SEM 0.206). Values mIKI, RT, and σ_2 are averaged across 180 trials within each participant. mIKI and RT values are log-transformed. In every plot, to the right of each mean (large dot) and SEM (denoted by the vertical bar) are displayed the individual data points in each group to visualize group population variability.

performance tempo to the strength of expectations about the action-reward contingency (Fig. 7D).

Regarding RT, the predictive power for the range of RT values was weaker compared with performance tempo (Fig. 8A), yet Gelman–Rubin statistics demonstrated an excellent chain convergence (R-hat values equal to 1.00). BLMM analyses showed no differences between Q8_T and Q8_F (intercepts) on RT, which is in line with our BF results (Fig. 8B). We found no robust evidence for an association (slope) between the strength of predictions about the action-reward contingency and RT (Fig. 8C). The 95% CI of the slope distribution ranged from -0.04 to 0.00 . A closer look at the upper bound of the distribution including three decimal digits revealed a value of 0.002 , demonstrating that 0 was marginally part of the 95% CI. This outcome suggests that RT is not robustly modulated by the strength of predictions about the action-reward contingency, unlike performance tempo.

No between-group slope differences were observed. Thus, as for performance tempo, subjective inferences about credit assignment did not modulate the association between RT and the strength of expectations about the action-reward contingency (Fig. 8D).

Finally, we investigated the effect of differences in inferences about reward assignment on the post-performance subjective error rate. First, the subjective error rate estimation was validated by computing BF analysis on the correlation between subjective

and empirical error rates. Results provided strong evidence for a positive association in the full sample ($N = 39$; BF = 10.204; $r = 0.448$, $p = 0.004$). Next, we found no support for between-group differences in the subjective error rate (BF = 0.432, demonstrating anecdotal evidence for the null hypothesis; $t_{(36)} = -0.850$, $p = 0.401$). Thus, being not *always* sure about the causes for the lack of reward did not influence the rate of subjective number estimate of performance errors.

To conclude, our analyses provided evidence for the lack of differences between Q8_T and Q8_F in the evaluated parameters, suggesting that subjective inferences about task-related credit assignment do not modulate decision-making, general motor performance or the association between expectation on reward probability and motor vigor. Thus, even if the groups in Study 1 would have had differences in credit assignment, it is unlikely that this would have led to a modulation of group effects. In addition, here, we found further support for our main research hypothesis, whereby stronger predictions about the action-reward contingency enhanced motor vigor through faster movement.

Study 3

Sensitivity of motor performance to confidence ratings about reward. In this study we focused our BLMM analysis on the association between motor performance (mIKI and RT) and confidence ratings to investigate how explicit beliefs about the reward

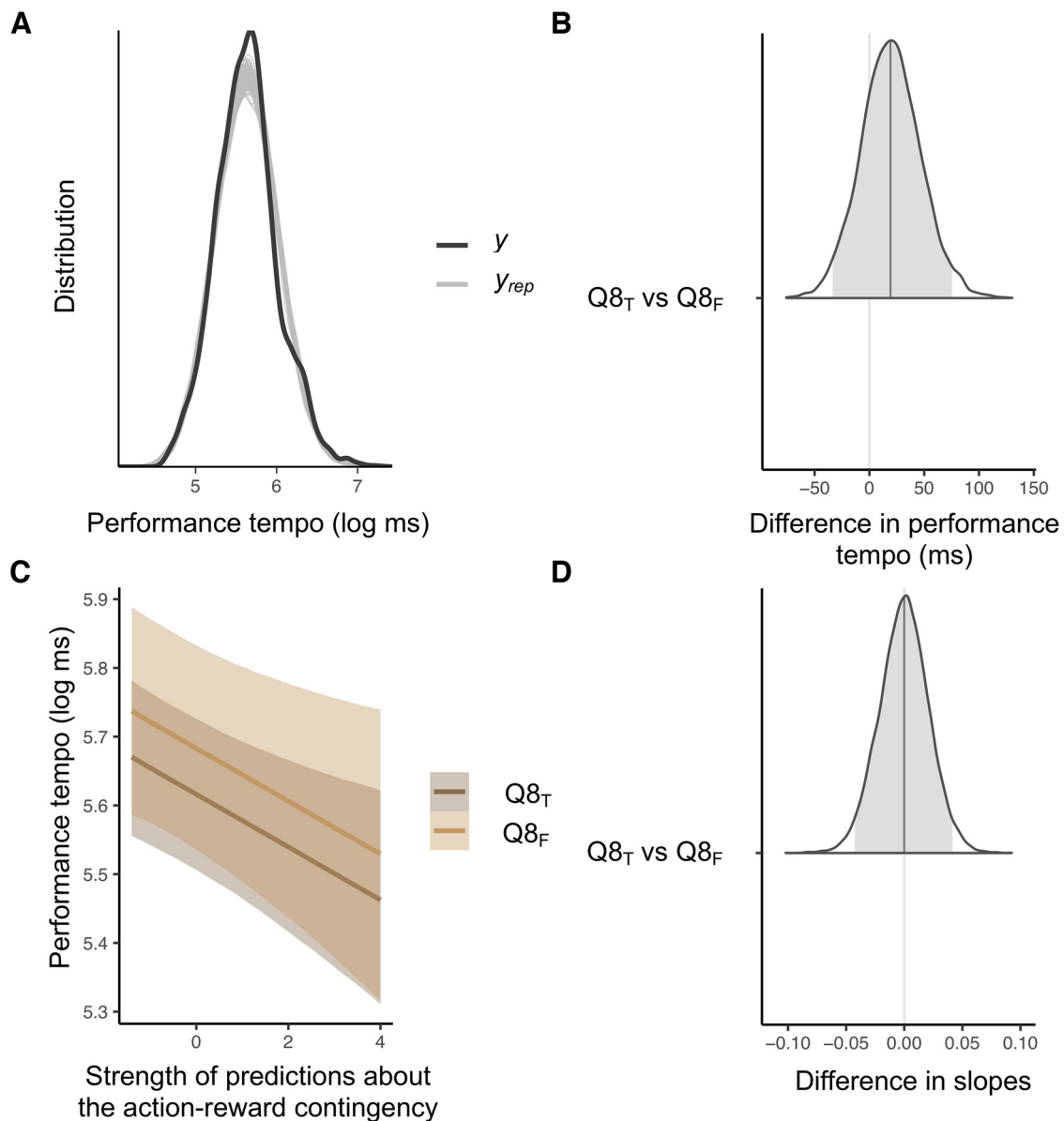


Figure 7. No effect of retrospective credit assignment on motor vigor: performance tempo. Bayesian linear mixed models [BLMMs; model number 6, $y \sim 1 + \text{group} * x + (1 + x | \text{subject}) + (1 | \text{trial})$] with participants that replied True to Question 8 ($Q8_T$; see Table 2) as reference group in Study 2. **A**, Illustration of the posterior predictive checks where the distribution of the observed outcome variable (y , in our case performance tempo) is compared with simulated datasets (y_{rep}) from the posterior predictive distribution (100 draws). **B**, Distribution of the difference in milliseconds between performance tempo (intercept) in $Q8_T$ and in participants that replied False to Question 8 ($Q8_F$; see Table 2). The gray vertical bar indicates the posterior point estimate, while the gray area under the curve represents the 95% credible interval (CI). In the current plot, CI does overlap with zero (the null hypothesis). This indicates that there is 95% probability of no between-group differences in performance tempo. **C**, Results of the BLMM analysis. We analyzed how the strength of predictions about the action-reward contingency modulates performance tempo separately for $Q8_T$ (in dark brown) and $Q8_F$ (in light brown). Here, mIKI (performance tempo: mean interkeystroke-interval) values are represented in the log-scale. The negative slopes suggest that stronger predictions about the action-reward contingency are associated with faster performance tempo, which replicates our findings in the main experiment (see Fig. 4C). **D**, Distribution of the difference between slopes in $Q8_T$ and $Q8_F$. Here, as CIs include zero, we can conclude with 95% probability that groups do not differ in how the strength of predictions about the reward contingency influences motor performance tempo. Thus, the sensitivity of performance tempo to the strength of predictions about the reward mapping is not differently modulated between groups.

outcome modulated motor vigor. Table 5 presents a summary of the posterior distributions for the winning models.

For performance tempo, LOO-CV identified the most complex model (model number 4) as the best fit (mIKI, $\text{elpd_diff} = -112.4178$, $\text{SE_diff} = 15.74263$; $\text{elpd_diff} > 2 * \text{SE_diff}$). The posterior predictive checks demonstrated that the observed outcome variable y overlapped well with the simulated datasets y_{rep} from the posterior predictive distribution (Fig. 9A). The y distribution exhibited two peaks, however, denoting two modes of mean performance tempo in our sample. The BLMM analyses showed a

negative association (slope) between the confidence ratings and the performance tempo, with stronger explicit beliefs about the reward tendency speeding up performance (Fig. 9B). The slope estimate was -0.04 (95% CI from -0.08 to -0.001 , including three decimal digits in the upper bound; Fig. 9C).

In the case of RT, LOO-CV identified the model number 3 as the best fit ($\text{elpd_diff} = -45.046830$, $\text{SE_diff} = 18.255767$; $\text{elpd_diff} > 2 * \text{SE_diff}$). This model did not include trials as random effect. The posterior predictive checks showed in this case that the y and y_{rep} distributions overlapped perfectly (Fig. 9D).

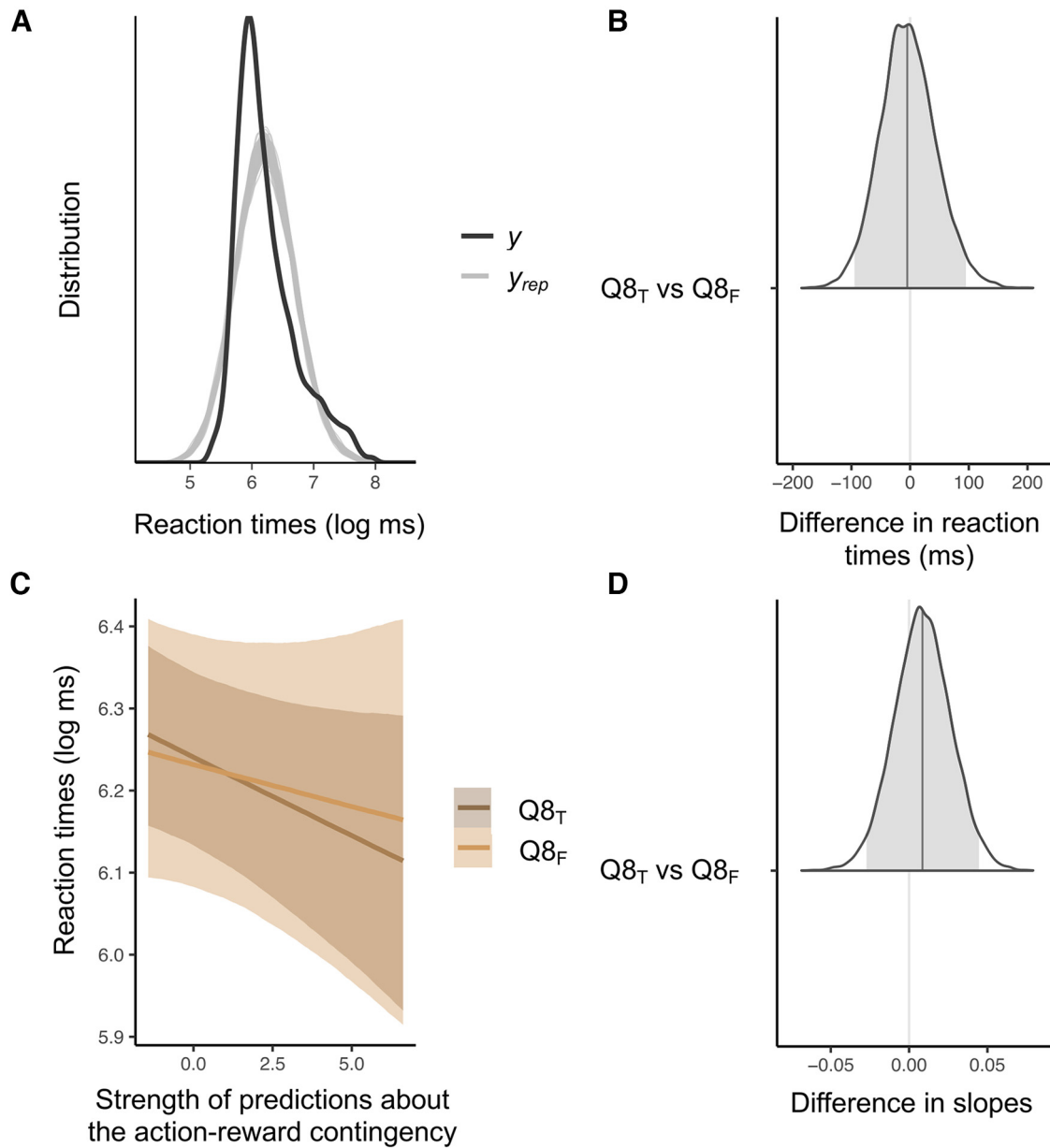


Figure 8. No effect of retrospective credit assignment on motor vigor: reaction times. Bayesian linear mixed models [BLMMs; model number 6, $y \sim 1 + \text{group} * x + (1 + x|\text{subject}) + (1|\text{trial})$] with participants that replied True to Question 8 (Q8_T; see Table 2) as reference group in Study 2. **A**, Illustration of the posterior predictive checks where the distribution of the observed outcome variable (y , in our case RT) is compared with simulated datasets (y_{rep}) from the posterior predictive distribution (100 draws). **B**, Distribution of the difference in milliseconds between RT (intercept) in Q8_T and in participants that replied False to Question 8 (Q8_F; see Table 2). The gray vertical bar indicates the posterior point estimate, while the gray area under the curve represents the 95% credible interval (CI). In the current plot, CI does overlap with zero (the null hypothesis). This indicates that there is 95% probability of no between-group differences in performance tempo. **C**, Results of the BLMM analysis. We analyzed how the strength of predictions about the action-reward contingency modulates RT separately for Q8_T (in dark brown) and Q8_F (in light brown). Here, RT values are represented in the log-scale. We found no robust evidence for a modulation of RT by the strength of expectations about the reward mapping. The upper bound of the distribution including three decimal digits revealed a value of 0.002, demonstrating that 0 was marginally part of the 95% CI. **D**, Distribution of the difference between slopes in Q8_T and Q8_F. Here, as CIs include zero, we can conclude with 95% probability that groups do not differ in how the strength of predictions about the reward contingency influences RT. Thus, the sensitivity of RT to the strength of predictions about the reward mapping is not differently modulated between groups.

As opposed to performance tempo, we found no robust modulation of RT by confidence ratings (Fig. 9E). The 95% CI of the slope distribution ranged from -0.20 to 0.01 . Thus, a zero effect was a credible value of the slope distribution (Fig. 9F).

Overall, these results support the conclusion that being more certain about obtaining the reward speeds up performance tempo, and thus movement time, without having a clear effect on RT. This expands our previous findings on the computational parameter $|\hat{\mu}_2|_c$, supporting a motor invigoration effect by explicit beliefs about the reward tendency under volatility.

In a separate sanity check, we assessed whether our measure of confidence was correlated with $|\hat{\mu}_2|$ in the HGF₂. This would suggest that implicit beliefs about the tendency of the action-reward contingency, captured with computational modeling, can be a proxy for explicit ratings about the confidence of reward delivery. Indeed, a BLMM analysis demonstrated a strong association between $|\hat{\mu}_2|$ and confidence ratings. The posterior point estimate for the intercept was 0.53 , CI = $[0.47, 0.59]$. The distribution of the fixed effect of the association between $|\hat{\mu}_2|$ and the confidence ratings had a posterior point estimate of 0.09 ,

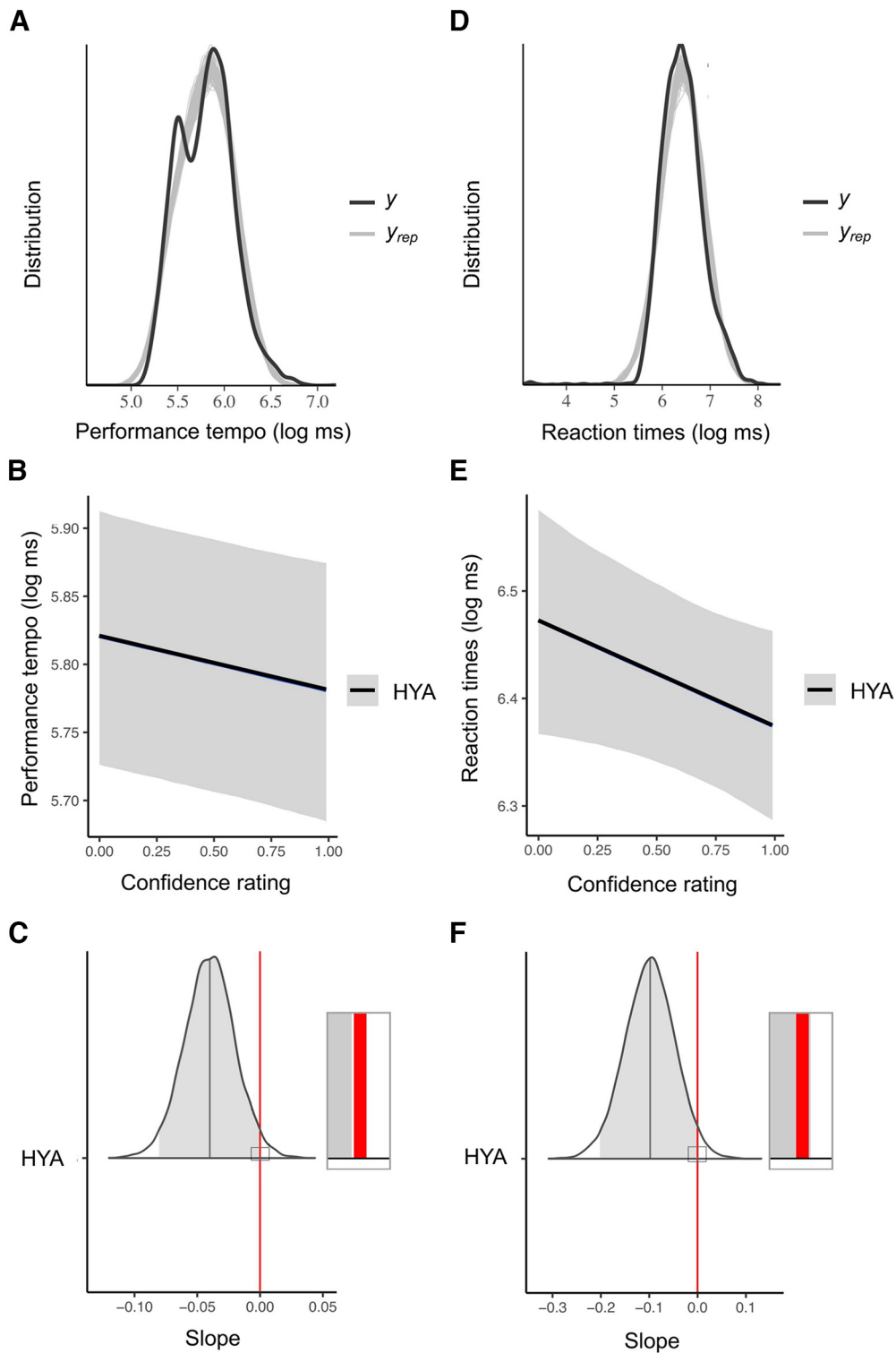


Figure 9. Explicit confidence ratings invigorate performance tempo. Bayesian linear mixed model [BLMM; model number 4, $y \sim 1 + x + (1 + x|subject) + (1|trial)$] in Study 3 for performance tempo (left) and model number 3, $y \sim 1 + x + (1 + x|subject)$ in Study 3 for reaction times (RT; right). **A**, Illustration of the posterior predictive checks where the distribution of the observed outcome variable (y , in our case performance tempo) is compared with simulated datasets (y_{rep}) from the posterior predictive distribution (100 draws). **B**, Results of the BLMM analysis. We analyzed how explicit beliefs about the reward tendency (confidence ratings) modulate performance tempo. Here, mIKI (performance tempo: mean interkeystroke-interval) values are represented in the log-scale. The negative slope had a point estimate of -0.04 [95% credible interval (CI) from -0.08 to -0.001 , including three decimal digits in the upper bound]. The 95% CI did not include zero. This suggests that being more certain about receiving a reward outcome is associated with faster performance tempo, which replicates our findings with the computational parameter $|\hat{\mu}_2|$ (see Figs. 4C, 7C). **C**, Distribution of the slope. The gray vertical bar indicates the posterior point estimate, while the gray area under the curve represents the 95% CI. The vertical red line denotes zero. **D**, Illustration of the posterior predictive checks where the distribution of the observed outcome variable (y , in our case RT) is compared with simulated datasets (y_{rep}) from the posterior predictive distribution (100 draws). **E**, Results of the BLMM analysis. Here, RT values are represented in the log-scale. We found no robust evidence for a modulation of RT by the strength of expectations about the reward mapping (95% CI from -0.20 to 0.01). **F**, Distribution of the slope. The gray vertical bar indicates the posterior point estimate, while the gray area under the curve represents the 95% CI. The vertical red line denotes zero.

CI [0.04, 0.14]. R-hat values were below 1.1, indicating chain convergence (Gelman and Rubin, 1992).

Last, descriptive statistics of performance variables in this task revealed values consistent with HYA samples in Studies 1 and 2 (mIKL, in milliseconds, mean 335, SEM 14.4; RT, in milliseconds, mean 662, SEM 26.7; percWin, mean 0.542, SEM 0.011; conf, mean 0.527, SEM 0.028). Also, out of the 180 trials, participants made 9.1 (SEM 1.6) performance errors on average, while they subjectively reported making 4.8 (SEM 0.7) errors. Thus, they subjectively reported only 53% of the performance errors they committed.

Discussion

We investigated how predictions about the tendency of the action-reward contingency invigorated motor performance trial-by-trial in healthy younger adults (HYA), in medicated Parkinson's disease patients (PD), and in an age-matched sample of healthy older adults (HOA). The task was a combination of a standard reversal learning and decision-making paradigm with a motor sequence task. We fitted the trial-by-trial behavioral data using the Hierarchical Gaussian Filter (HGF; Mathys et al., 2011, 2014; Frässle et al., 2021) and performed Bayesian analyses [Bayes factor and Bayesian linear mixed models (BLMMs)].

Study 1 showed a trial-by-trial modulation of performance tempo, commensurate with movement time, by the strength of expectations about the action-reward contingencies. The invigoration effect was limited to performance tempo and was not observed for reaction time (RT). Moreover, BLMM revealed a similar sensitivity of performance tempo to these predictions in our three groups. This provides compelling evidence for a preservation of motor invigoration by expectations of reward probability in HOA and PD, expanding the understanding on how reward sensitivity and reversal learning interact to modulate motor vigor in aging and medicated PD.

Previous investigations of the beneficial effects of reward on motor behavior (e.g., faster and more accurate motor performance; Sedaghat-Nejad et al., 2019) have been limited to manipulations of reward magnitude (presence/absence; large/small) in deterministic contexts (Codol et al., 2020; Aves et al., 2021; Sporn et al., 2022). Our findings expand on computational work that demonstrated the updating of beliefs in a perceptual task to speed RT (Marshall et al., 2016). The authors found that, as participants learned to track the transition probabilities between stimuli, different decision-making variables affected RT. Our results show that the trial-by-trial influence of motor vigor by belief updating can be extended beyond the perceptual domain to learning about action-reward contingencies.

Despite the preserved motor invigoration effect in HOA and PD, we found extreme evidence for between-group differences in the mean performance tempo. HYA were faster than HOA and PD, and HOA quicker than PD. The slower sequence execution in HOA is consistent with a general slowness of hand movements in later stages of life (Ketcham et al., 2002; Aves et al., 2021). Regarding PD, the slower performance is likely explained by a sequence effect (SE). SE is a common bradykinetic symptom in PD, which manifests through slower and attenuated sequential movements (Kang et al., 2010). Dopamine (DA) intake does not ameliorate symptoms associated with SE, suggesting a non-DA involvement in the pathophysiology of this effect (Bologna et al., 2016). Similar results were found for RT, with HYA displaying shorter RT than HOA and PD. Yet, RT did not dissociate between HOA and PD.

We additionally found evidence for similar win and error rates in our three groups. Empirical findings on reward learning in aging and medicated PD have been mixed. Some studies have shown reduced probabilistic and reversal learning in older adults and PD ON medication, suggesting difficulties in establishing new stimulus-outcome associations and updating reward beliefs (Cools et al., 2001; Eppinger et al., 2011; Nassar et al., 2016). Consistent with this, de Boer et al. (2017) demonstrated poorer probabilistic reversal learning in aging compared with young participants, with the attenuation of the anticipatory values signals in the prefrontal brain accounting for the impoverished performance. However, other work argued for preserved reward sensitivity and learning in older adults and medicated PD (Fera et al., 2005; Euteneuer et al., 2009; Aves et al., 2021). Specifically, PD ON medication have been found to successfully learn from rewards, and exhibit deficits in reversal learning exclusively for negative feedback (Frank et al., 2004; Levy-Gigi et al., 2019). Also, Hird et al. (2022) reported that age does not modulate the invigorating effect of reward on motor responses. This is consistent with our findings, highlighting a preserved motor invigoration effect by reward in aging and medicated PD.

Our groups did not differ in the main markers of decision-making. We provided some evidence for the absence of a group effect on tonic volatility (ω_2 ; index of individual learning about the action-reward mapping under volatility; Hein et al., 2021), estimated uncertainty about the action-reward tendency (σ_2) and on the mapping from beliefs to responses (ζ). Accordingly, belief updating in our task with changing action-reward contingencies was comparable across HYA, HOA, and PD groups.

One aspect that was not identified in Study 1 was whether participants correctly inferred the hidden causes for the lack of reward (McDougle et al., 2016). Study 2 demonstrated that retrospective subjective inference about credit assignment did not contribute to differences in general motor performance, decision-making, motor vigor or the subjective estimate of performance errors. Because the feedback that participants received was veridical (unlike in McDougle et al., 2016), the effects of misattribution of the causes of zero reward in our study are likely very small, as the anecdotal evidence suggests. A limitation of this study, however, was that it relied on retrospective self-report. Accordingly, we conducted a third study to determine whether trial-by-trial explicit beliefs about the reward tendency (confidence ratings) are associated with faster motor performance.

Study 3 demonstrated that performance tempo is associated with confidence ratings trial-by-trial: being more certain about obtaining the reward speeded up the movement. Moreover, the confidence ratings were robustly correlated with the strength of the predictions. This outcome supports that implicit beliefs about the tendency of the action-reward contingency, captured with computational modeling, can be a proxy for explicit ratings about the confidence of reward delivery.

The invigoration effect of beliefs (both implicit and explicit) did not extend to RT. Accordingly, across our three studies, RT was not robustly modulated in the same dynamic trial-wise manner as performance tempo was. In Studies 1 and 2, RT included deliberation time (no constraints on initiating the sequence), which could have introduced noise to the RT distribution and weakened the motor vigor effects. By contrast, RT in Study 3 excluded deliberation time.

According to current hypotheses, motor vigor is based on trading-off future efforts and gains, reflecting a subject's willingness to invest energy to harvest future rewards (Shadmehr, 2010; Yoon et al., 2020). Specifically, it increases when the

option is inferred to be valuable and decreases for perceived effort. This has been demonstrated both for movement times and RT (Summerside et al., 2018; Codol et al., 2020). It follows that changes in vigor should be modulated by inferences on the tendency of reward probability. We demonstrated that exclusively performance tempo, commensurate with movement time, is affected by beliefs about the action-reward contingency on a trial-by-trial basis. The lack of robust invigoration effects on RT is consistent with sequential planning effects introducing noise to the RT distribution. Recent work has demonstrated that the preparatory state of discrete sequential finger movements reflects sequence planning skills (Mantziara et al., 2021). Accordingly, RT in our task would include trial-by-trial variability in sequence preparation, which may mask the underlying motor vigor effects. A prediction for future work would be a trial-by-trial invigoration of RT, beyond movement time, in motor tasks that do not require preparation of discrete movements.

A limitation of the present work is that, because of the nature of our online experiment, we only tested PD ON medication. Future work should investigate the effect of DA on the trial-by-trial association between the expectations of reward probability and motor vigor. Interestingly, a recent study by Hird et al. (2022) found only a weak association between dopamine D1 receptor availability and the invigorating effect of reward. This outcome, together with our finding of preserved dynamic motor vigor effects in medicated PD, raises an interesting question: if motor vigor and learning are driven by the dopaminergic system as previously postulated (Balleine et al., 2007; Eppinger et al., 2011), how robust is this association in more complex scenarios rich in uncertainty and with changing reward probabilities over time? Our results suggest that DA-replacement therapy could restore putative decision-making deficits during learning in volatile environments in PD.

In addition, the interplay between dynamic decision-making and motor performance might be driven by several neurotransmitter systems linked to precision weighting of prediction errors during belief updating: acetylcholine (Moran et al., 2013), noradrenaline (Dayan and Yu, 2006), in addition to dopamine (Iglesias et al., 2013; Haarsma et al., 2021). On a neural level, learning uncertain stimulus-reward contingencies relies on the ACC, OFC, and portions of the mPFC (Hayden et al., 2011; Rouault et al., 2019; Rolls et al., 2022). The mPFC is also involved in mapping beliefs to actions during exploration-exploitation (Domenech et al., 2020). Follow-up neuroimaging studies could assess the role of these regions in the motor vigor effects reported here, including the preserved effects in aging and PD.

To conclude, this study is the first to demonstrate that inferring the probabilistic reward mappings positively biases motor performance through faster performance tempo. Additionally, we provided novel evidence for a preserved sensitivity of the motor invigoration effects in HOA and PD. Thus, healthy young, old and medicated PD can similarly obtain benefits in their motor performance when updating beliefs about the volatile action-reward contingencies.

References

- Adkins TJ, Lee TG (2021) Reward modulates cortical representations of action. *Neuroimage* 228:117708.
- Andraszewicz S, Scheibehenne B, Rieskamp J, Grasman R, Verhagen J, Wagenmakers EJ (2015) An introduction to Bayesian hypothesis testing for management research. *J Manag* 41:521–543.
- Aves P, Moreau L, Alghamdi A, Sporn S, Galea JM (2021) Age-related differences in reward-based modulation of sequential reaching performance. *bioRxiv* 461920. <https://doi.org/10.1101/2021.09.27.461920>.
- Balleine BW, Delgado MR, Hikosaka O (2007) The role of the dorsal striatum in reward and decision-making. *J Neurosci* 27:8161–8165.
- Behrens TEJ, Woolrich MW, Walton ME, Rushworth MFS (2007) Learning the value of information in an uncertain world. *Nat Neurosci* 10:1214–1221.
- Bologna M, Leodori G, Stirpe P, Paparella G, Colella D, Belvisi D, Fasano A, Fabbrini G, Berardelli A (2016) Bradykinesia in early and advanced Parkinson's disease. *J Neurol Sci* 369:286–291.
- Bürkner PC (2017) brms: an R package for Bayesian multilevel models using Stan. *J Stat Softw* 80:1–28.
- Bürkner PC (2018) Advanced Bayesian multilevel modeling with the R package Brms. *R J* 10:395–411.
- Bürkner PC (2021) Bayesian item response modeling in R with Brms and Stan. *J Stat Softw* 100:1–54.
- Carroll TJ, McNamee D, Ingram JN, Wolpert DM (2019) Rapid visuomotor responses reflect value-based decisions. *J Neurosci* 39:3906–3920.
- Codol O, Holland PJ, Manohar SG, Galea JM (2020) Reward-based improvements in motor control are driven by multiple error-reducing mechanisms. *J Neurosci* 40:3604–3620.
- Cools R, Barker RA, Sahakian BJ, Robbins TW (2001) Enhanced or impaired cognitive function in Parkinson's disease as a function of dopaminergic medication and task demands. *Cereb Cortex* 11:1136–1143.
- Cousineau D (2020) How many decimals? Rounding descriptive and inferential statistics based on measurement precision. *J Math Psychol* 97:102362.
- Dayan P, Yu AJ (2006) Phasic norepinephrine: a neural interrupt signal for unexpected events. *Network* 17:335–350.
- 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.
- de Boer L, Axelsson J, Riklund K, Nyberg L, Dayan P, Bäckman L, Guitart-Masip M (2017) Attenuation of dopamine-modulated prefrontal value signals underlies probabilistic reward learning deficits in old age. *Elife* 6:e26424.
- den Ouden HEM, Daunizeau J, Roiser J, Friston KJ, Stephan KE (2010) Striatal prediction error modulates cortical coupling. *J Neurosci* 30:3210–3219.
- Diaconescu AO, Mathys C, Weber LAE, Daunizeau J, Kasper L, Lomakina EI, Fehr E, Stephan KE (2014) Inferring on the intentions of others by hierarchical Bayesian learning. *PLoS Comput Biol* 10:e1003810.
- Domenech P, Rheims S, Koechlin E (2020) Neural mechanisms resolving exploitation-exploration dilemmas in the medial prefrontal cortex. *Science* 369:eabb0184.
- Eppinger B, Hämmerer D, Li SC (2011) Neuromodulation of reward-based learning and decision making in human aging. *Ann N Y Acad Sci* 1235:1–17.
- Euteneuer F, Schaefer F, Stuermer R, Boucsein W, Timmermann L, Barbe MT, Ebersbach G, Otto J, Kessler J, Kalbe E (2009) Dissociation of decision-making under ambiguity and decision-making under risk in patients with Parkinson's disease: a neuropsychological and psychophysiological study. *Neuropsychologia* 47:2882–2890.
- Fahn S, Elton RL (1987) The unified Parkinson's disease rating scale. In: *Recent developments in Parkinson's disease, Vol 2* (Fahn S, Marsden CD, Calne DB, Goldstein M, eds), pp 293–304. Florham Park: Macmillan Health Care Information.
- Feldman H, Friston KJ (2010) Attention, uncertainty, and free-energy. *Front Hum Neurosci* 4:215.
- Fera F, Weickert TW, Goldberg TE, Tessitore A, Hariri A, Das S, Lee S, Zolnick B, Meeter M, Myers CE, Gluck MA, Weinberger DR, Mattay VS (2005) Neural mechanisms underlying probabilistic category learning in normal aging. *J Neurosci* 25:11340–11348.
- Frank MJ, Seeberger LC, O'reilly RC (2004) By carrot or by stick: cognitive reinforcement learning in parkinsonism. *Science* 306:1940–1943.
- Frässle S, et al. (2021) TAPAS: an open-source software package for translational neuromodeling and computational psychiatry. *Front Psychiatry* 12:68081.
- Friston KJ, Stephan KE, Montague R, Dolan RJ (2014) Computational psychiatry: the brain as a phantastic organ. *Lancet Psychiatry* 1:148–158.
- Frömer R, Nassar MR, Bruckner R, Stürmer B, Sommer W, Yeung N (2021) Response-based outcome predictions and confidence regulate feedback processing and learning. *Elife* 10:e62825.

- Galaro JK, Celnik P, Chib VS (2019) Motor cortex excitability reflects the subjective value of reward and mediates its effects on incentive-motivated performance. *J Neurosci* 39:1236–1248.
- Gelman A, Rubin DB (1992) Inference from iterative simulation using multiple sequences. *Stat Sci* 7:457–472.
- Haarsma J, Fletcher PC, Griffin JD, Taverne HJ, Ziauddeen H, Spencer TJ, Miller C, Katthagen T, Goodyer I, Diederer KM, Murray GK (2021) Precision weighting of cortical unsigned prediction error signals benefits learning, is mediated by dopamine, and is impaired in psychosis. *Mol Psychiatry* 26:5320–5333.
- Hayden BY, Heilbronner SR, Pearson JM, Platt ML (2011) Surprise signals in anterior cingulate cortex: neuronal encoding of unsigned reward prediction errors driving adjustment in behavior. *J Neurosci* 31:4178–4187.
- Hein TP, Herrojo Ruiz M (2022) State anxiety alters the neural oscillatory correlates of predictions and prediction errors during reward-based learning. *Neuroimage* 249:118895.
- Hein TP, de Fockert J, Ruiz MH (2021) State anxiety biases estimates of uncertainty and impairs reward learning in volatile environments. *Neuroimage* 224:117424.
- Herrojo Ruiz M, Jabusch HC, Altenmüller E (2009) Detecting wrong notes in advance: neuronal correlates of error monitoring in pianists. *Cereb Cortex* 19:2625–2639.
- Herrojo Ruiz M, Maess B, Altenmüller E, Curio G, Nikulin VV (2017) Cingulate and cerebellar beta oscillations are engaged in the acquisition of auditory-motor sequences. *Hum Brain Mapp* 38:5161–5179.
- Hird EJ, Beierholm U, De Boer L, Axelsson J, Backman L, Guitart-Masip M (2022) Dopamine and reward-related vigor in younger and older adults. *Neurobiol Aging* 118:34–43.
- Iglesias S, Mathys C, Brodersen KH, Kasper L, Piccirelli M, den Ouden HEM, Stephan KE (2013) Hierarchical prediction errors in midbrain and basal forebrain during sensory learning. *Neuron* 80:519–530.
- Kang SY, Wasaka T, Shamim EA, Auh S, Ueki Y, Lopez GJ, Kida T, Jin S-H, Dang N, Hallett M (2010) Characteristics of the sequence effect in Parkinson's disease: sequence effect in Parkinson's disease. *Mov Disord* 25:2148–2155.
- Ketcham CJ, Seidler RD, Van Gemmert AWA, Stelmach GE (2002) Age-related kinematic differences as influenced by task difficulty, target size, and movement amplitude. *J Gerontol B Psychol Sci Soc Sci* 57:P54–P64.
- Kjær SW, Damholdt MF, Callesen MB (2018) A systematic review of decision-making impairments in Parkinson's disease: dopaminergic medication and methodological variability. *Basal Ganglia* 14:31–40.
- Levy-Gigi E, Haim-Nachum S, Hall JM, Crouse JJ, Winwood-Smith R, Lewis SJ, Moustafa AA (2019) The interactive effect of valence and context on reversal learning in individuals with Parkinson's disease. *Neurosci Lett* 692:216–224.
- Lewandowski D, Kurowicka D, Joe H (2009) Generating random correlation matrices based on vines and extended onion method. *J Multivar Anal* 100:1989–2001.
- Mantziara M, Ivanov T, Houghton G, Kornysheva K (2021) Competitive state of movements during planning predicts sequence performance. *J Neurophysiol* 125:1251–1268.
- Manohar SG, Muhammed K, Fallon SJ, Husain M (2019) Motivation dynamically increases noise resistance by internal feedback during movement. *Neuropsychologia* 123:19–29.
- Marshall L, Mathys C, Ruge D, de Berker AO, Dayan P, Stephan KE, Bestmann S (2016) Pharmacological fingerprints of contextual uncertainty. *PLoS Biol* 14:e1002575.
- Mathys C, Daunizeau J, Friston KJ, Stephan KE (2011) A Bayesian foundation for individual learning under uncertainty. *Front Hum Neurosci* 5:39.
- 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.
- McDougle SD, Boggess MJ, Crossley MJ, Parvin D, Ivry RB, Taylor JA (2016) Credit assignment in movement-dependent reinforcement learning. *Proc Natl Acad Sci U S A* 113:6797–6802.
- Metitieri T, Geroldi C, Pezzini A, Frisoni GB, Bianchetti A, Trabucchi M (2001) The Itel-MMSE: an Italian telephone version of the mini-mental state examination. *Int J Geriatr Psychiatry* 16:166–167.
- Moran RJ, Campo P, Symmonds M, Stephan KE, Dolan RJ, Friston KJ (2013) Free energy, precision and learning: the role of cholinergic neuro-modulation. *J Neurosci* 33:8227–8236.
- Nassar MR, Bruckner R, Gold JI, Li S-C, Heekeren HR, Eppinger B (2016) Age differences in learning emerge from an insufficient representation of uncertainty in older adults. *Nat Commun* 7:11609.
- Padmala S, Pessoa L (2011) Reward reduces conflict by enhancing attentional control and biasing visual cortical processing. *J Cogn Neurosci* 23:3419–3432.
- Powers AR, Mathys C, Corlett PR (2017) Pavlovian conditioning-induced hallucinations result from overweighting of perceptual priors. *Science* 357:596–600.
- R Core Team (2022) R: a language and environment for statistical computing. In: R foundation for statistical computing. Vienna: Austria. Available at <https://www.R-project.org/>.
- Reed EJ, Uddenberg S, Suthaharan P, Mathys CD, Taylor JR, Groman SM, Corlett PR (2020) Paranoia as a deficit in non-social belief updating. *Elife* 9:e56345.
- Rescorla RA, Wagner AR (1972) A theory of Pavlovian conditioning: variations in the effectiveness of reinforcement and nonreinforcement. *Curr Res Theory* 2:64–99.
- Rolls ET, Deco G, Huang CC, Feng J (2022) The human orbitofrontal cortex, vmPFC, and anterior cingulate cortex effective connectome: emotion, memory, and action. *Cereb Cortex* 33:330–356.
- Rouault M, Drugowitsch J, Koechlin E (2019) Prefrontal mechanisms combining rewards and beliefs in human decision-making. *Nat Commun* 10:301.
- Rouder JN, Morey RD, Speckman PL, Province JM (2012) Default Bayes factors for ANOVA designs. *J Math Psychol* 56:356–374.
- Ryterska A, Jahanshahi M, Osman M (2013) What are people with Parkinson's disease really impaired on when it comes to making decisions? A meta-analysis of the evidence. *Neurosci Biobehav Rev* 37:2836–2846.
- Sedaghat-Nejad E, Herzfeld DJ, Shadmehr R (2019) Reward prediction error modulates saccade vigor. *J Neurosci* 39:5010–5017.
- Shadmehr R (2010) Control of movements and temporal discounting of reward. *Curr Opin Neurobiol* 20:726–730.
- Sheffield JM, Suthaharan P, Leptourgos P, Corlett PR (2022) Belief updating and paranoia in individuals with schizophrenia. *Biol Psychiatry Cogn Neurosci Neuroimaging* 7:1149–1157.
- Sivula T, Magnusson M, Matamoros AA, Vehtari A (2022) Uncertainty in Bayesian leave-one-out cross-validation based model comparison. *arXiv* 2008.10296. <https://doi.org/10.48550/arXiv.2008.10296>.
- Soch J, Allefeld C (2018) MACS – a new SPM toolbox for model assessment, comparison and selection. *J Neurosci Methods* 306:19–31.
- Spielberger CD, Gorsuch RL, Lushene R, Vagg PR, Jacobs GA (1983) Manual for the state-trait anxiety inventory. Palo Alto: Consulting Psychologists Press.
- Sporn S, Chen X, Galea JM (2022) The dissociable effects of reward on sequential motor behavior. *J Neurophysiol* 128:86–104.
- Stefanics G, Heinzle J, Horváth AA, Stephan KE (2018) Visual mismatch and predictive coding: a computational single-trial ERP study. *J Neurosci* 38:4020–4030.
- Summerside EM, Shadmehr R, Ahmed AA (2018) Vigor of reaching movements: reward discounts the cost of effort. *J Neurophysiol* 119:2347–2357.
- Sutton RS (1992) Gain adaptation beats least squares? Paper presented at the 7th Yale Workshop on Adaptive and Learning Systems. New Haven, May.
- Valton V, Wise T, Robinson OJ (2020) Recommendations for Bayesian hierarchical model specifications for case-control studies in mental health. *arXiv* 2011.01725. <https://doi.org/10.48550/arXiv.2011.01725>.
- Vehtari A, Gelman A, Gabry J (2017) Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Stat Comput* 27:1413–1432.
- Weber LA, Diaconescu AO, Mathys C, Schmidt A, Komater M, Vollenweider F, Stephan KE (2020) Ketamine affects prediction errors about statistical regularities: a computational single-trial analysis of the mismatch negativity. *J Neurosci* 40:5658–5668.
- Yoon T, Jaleel A, Ahmed AA, Shadmehr R (2020) Saccade vigor and the subjective economic value of visual stimuli. *J Neurophysiol* 123:2161–2172.
- Zigmond AS, Snaith RP (1983) The hospital anxiety and depression scale. *Acta Psychiatr Scand* 67:361–370.