

Domain-Specific Working Memory, But Not Dopamine-Related Genetic Variability, Shapes Reward-Based Motor Learning

 Peter Holland,*  Olivier Codol,* Elizabeth Oxley, Madison Taylor, Elizabeth Hamshere, Shadiq Joseph, Laura Huffer, and  Joseph M. Galea

School of Psychology and Centre for Human Brain Health, University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom

The addition of rewarding feedback to motor learning tasks has been shown to increase the retention of learning, spurring interest in its possible utility for rehabilitation. However, motor tasks using rewarding feedback have repeatedly been shown to lead to great interindividual variability in performance. Understanding the causes of such variability is vital for maximizing the potential benefits of reward-based motor learning. Thus, using a large human cohort of both sexes ($n = 241$), we examined whether spatial (SWM), verbal, and mental rotation (RWM) working memory capacity and dopamine-related genetic profiles were associated with performance in two reward-based motor tasks. The first task assessed the participant's ability to follow a slowly shifting reward region based on hit/miss (binary) feedback. The second task investigated the participant's capacity to preserve performance with binary feedback after adapting to the rotation with full visual feedback. Our results demonstrate that higher SWM is associated with greater success and an enhanced capacity to reproduce a successful motor action, measured as change in reach angle following reward. In contrast, higher RWM was predictive of an increased propensity to express an explicit strategy when required to make large reach angle adjustments. Therefore, SWM and RWM were reliable, but dissociable, predictors of success during reward-based motor learning. Change in reach direction following failure was also a strong predictor of success rate, although we observed no consistent relationship with working memory. Surprisingly, no dopamine-related genotypes predicted performance. Therefore, working memory capacity plays a pivotal role in determining individual ability in reward-based motor learning.

Key words: genetics; motor learning; reaching; reward; working memory

Significance Statement

Reward-based motor learning tasks have repeatedly been shown to lead to idiosyncratic behaviors that cause varying degrees of task success. Yet, the factors determining an individual's capacity to use reward-based feedback are unclear. Here, we assessed a wide range of possible candidate predictors, and demonstrate that domain-specific working memory plays an essential role in determining individual capacity to use reward-based feedback. Surprisingly, genetic variations in dopamine availability were not found to play a role. This is in stark contrast with seminal work in the reinforcement and decision-making literature, which show strong and replicated effects of the same dopaminergic genes in decision-making. Therefore, our results provide novel insights into reward-based motor learning, highlighting a key role for domain-specific working memory capacity.

Introduction

When performing motor tasks under altered environmental conditions, adaptation to the new constraints occurs through the

recruitment of a variety of systems (Taylor and Ivry, 2014). Arguably, the most studied of those systems is cerebellum-dependent adaptation, which consists of the implicit and automatic recalibration of mappings between actual and expected outcomes through sensory prediction errors (Tseng et al., 2007; Morehead et al., 2017). In addition to cerebellar adaptation, other work has demonstrated the involvement of a cognitive, deliberative process whereby motor plans are adjusted based on structural understanding of the task

Received Jan. 20, 2019; revised Sept. 17, 2019; accepted Sept. 24, 2019.

Author contributions: P.H., O.C., and J.M.G. designed research; P.H., O.C., E.O., M.T., E.H., S.J., and L.H. performed research; P.H. and O.C. analyzed data; P.H., O.C., E.O., M.T., E.H., S.J., L.H., and J.M.G. edited the paper; O.C. wrote the first draft of the paper; J.M.G. contributed unpublished reagents/analytic tools.

This work was supported by the European Research Council Starting Grant MotMotLearn 637488.

The authors declare no competing financial interests.

*P.H. and O.C. contributed equally to this work.

Correspondence should be addressed to Peter Holland at P.J.Holland@bham.ac.uk.

<https://doi.org/10.1523/JNEUROSCI.0583-19.2019>

Copyright © 2019 the authors

(Taylor and Ivry, 2011; Bond and Taylor, 2015). We label this process “explicit control” (Codol et al., 2018; Holland et al., 2018), although it has also been referred to as strategy (Taylor and Ivry, 2011) or explicit re-aiming (Morehead et al., 2015). Recently, it has been proposed that reinforcement learning, whereby the memory of successful or unsuccessful actions is strengthened or weakened, respectively, may also play a role (Huang et al., 2011; Izawa and Shadmehr, 2011; Shmuelof et al., 2012). Such reward-based reinforcement has been assumed to be an implicit and automatic process (Haith and Krakauer, 2013). However, recent evidence suggests that phenomena attributed to reinforcement-based learning during visuomotor rotation tasks can largely be explained through explicit processes (Codol et al., 2018; Holland et al., 2018).

One outstanding feature of reinforcement-based motor learning is the great variability expressed across individuals (Therrien et al., 2016, 2018; Codol et al., 2018; Holland et al., 2018). What factors underlie such variability is unclear. If reinforcement is explicitly grounded, it could be argued that individual working memory capacity (WMC), which is reliably related to the propensity to use explicit control in classical motor adaptation tasks (Anguera et al., 2010, 2012; Christou et al., 2016; Holland et al., 2018; Sidarta et al., 2018), would also predict performance in reinforcement-based motor learning. Anguera et al. (2010) demonstrated that mental rotation WMC (RWM), unlike other forms of working memory (WM), such as verbal WMC (VWM), correlates with explicit control. Recently, Christou et al. (2016) reported similar results with spatial WMC (SWM). If this extends to reward-based motor learning, this would strengthen the proposal that it bears a strong explicit component.

Another potential contributor to this variability is genetic profile. In previous work (Codol et al., 2018; Holland et al., 2018), we argue that reinforcement-based motor learning performance relies on a balance between exploration and exploitation of the task space, a feature reminiscent of structural learning and reinforcement-based decision-making (Sutton and Barto, 1998; Daw et al., 2005; Frank et al., 2009). A series of studies from Frank and colleagues suggest that individual tendencies to express explorative/exploitative behavior can be predicted based on dopamine-related genetic profile (Frank et al., 2007, 2009; Doll et al., 2016). Reinforcement has consistently been linked to dopaminergic function in a variety of paradigms; thus, such a relationship could also be expected in reward-based motor learning (Pekny et al., 2015). Specifically, Frank and colleagues focused on catecholamine-O-methyl-transferase (COMT), dopamine- and cAMP-regulated neuronal phosphoprotein (DARPP32), and dopamine receptor D2 (DRD2), and suggest a distinction between COMT-modulated exploration and DARPP32- and DRD2-modulated exploitation (Frank et al., 2009).

Consequently, we investigated the influence of WMC (RWM, SWM, and VWM) and genetic variations in dopamine metabolism (DRD2, DARPP32, and COMT) on individuals' ability to perform reward-based motor learning. We examined this using two established reward-based motor learning tasks. First, a task analogous to a gradually introduced rotation (Holland et al., 2018) required participants to learn to adjust the angle at which they reached to a slowly and secretly shifting reward region (Acquire); second, a task with an abruptly introduced rotation (Shmuelof et al., 2012; Codol et al., 2018) required participants to preserve performance with reward-based feedback after adapting to a visuomotor rotation (Preserve). The use of these two tasks enabled us to examine whether similar predictors of performance explained participant's capacity to acquire and preserve behavior with reward-based feedback.

Materials and Methods

Before the start of data collection, the sample size, variables of interest, and analysis method were preregistered. The preregistered information, data, and analysis code can be found online at <https://osf.io/j5v2s/> and <https://osf.io/rmwc2/> for the Preserve and Acquire tasks, respectively.

Participants

A total of 121 participants (30 male, mean age: 21.06 years, range: 18–32 years) and 120 participants (16 male, mean age: 19.24 years, range: 18–32 years) were recruited for the Acquire and Preserve tasks, respectively. All participants provided informed consent and were remunerated with either course credit or money (£7.50/h). All participants were free of psychological, cognitive, motor, or uncorrected visual impairment. The study was approved by and performed in accordance with the local research ethics committee of the University of Birmingham (Birmingham, UK).

Experimental design

Participants were seated before a horizontally fixed mirror reflecting a screen placed above, on which visual stimuli were presented. This arrangement resulted in the stimuli appearing at the level on which participants performed their reaching movements. The Acquire (gradual) and Preserve (abrupt) tasks were performed on two different stations, with a KINARM (BKIN Technology; sampling rate: 1000 Hz) and a Polhemus 3SPACE Fastrak tracking device (sampling rate: 120 Hz), used respectively. The Acquire task was run using Simulink (The MathWorks) and Dexterit-E (BKIN Technology), whereas the Preserve task was run using MATLAB (The MathWorks) and Psychophysics toolbox (Brainard, 1997). The Acquire task used the same paradigm and equipment as in Holland et al. (2018), with the exception of the maximum reaction time, which was increased from 0.6 to 1 s; and the maximum movement time, which was reduced from 1 to 0.6 s. The Preserve task used the same setup and display as in Codol et al. (2018); however, the number of “refresher” trials during the binary feedback blocks was increased from one to two in every 10 trials. The designs were kept as close as possible to their respective original publications to promote replication and comparability across studies. In both tasks, reaching movements were made with the dominant arm. Both the Acquire and Preserve tasks have previously been examined in isolation from each other: Acquire task (Therrien et al., 2016, 2018; Cashaback et al., 2017, 2019; Holland et al., 2018) and Preserve task (Shmuelof et al., 2012; Codol et al., 2018), and we maintain this distinction here. However, it should be noted that the two tasks are essentially visuomotor rotation tasks. One of the aims of this study was to determine whether similar mechanisms underlie the use of binary feedback in both the learning of a gradual rotation and maintenance of a previously learnt abrupt rotation. Therefore, despite the similarities, we analyze the results of each task in isolation in addition to comparing the results across tasks.

Reaching tasks

Acquire task. Participants performed 670 trials, each of which followed a stereotyped timeline. The starting position for each trial was in a consistent position ~30 cm in front of the midline and was indicated by a red circle (1 cm radius). After holding the position of the handle within the starting position, a target (red circle, 1 cm radius) appeared directly in front of the starting position at a distance of 10 cm. Participants were instructed to make a rapid “shooting” movement that passed through the target. If the cursor position at a radial distance of 10 cm was within a reward region ($\pm 5.67^\circ$, initially centered on the visible target; Fig. 1A, gray region), the target changed color from red to green and a green tick was displayed just above the target position, informing participants of the success of their movement. However, if the cursor did not pass through the reward region, the target disappeared from view and no tick was displayed, signaling failure (binary feedback). After each movement, the robot returned to the starting position and participants were instructed to passively allow this.

For the first 10 trials, the position of the robotic handle was displayed as a white cursor (0.5 cm radius) on screen. Following this practice block, the cursor was extinguished for the remainder of the experiment and

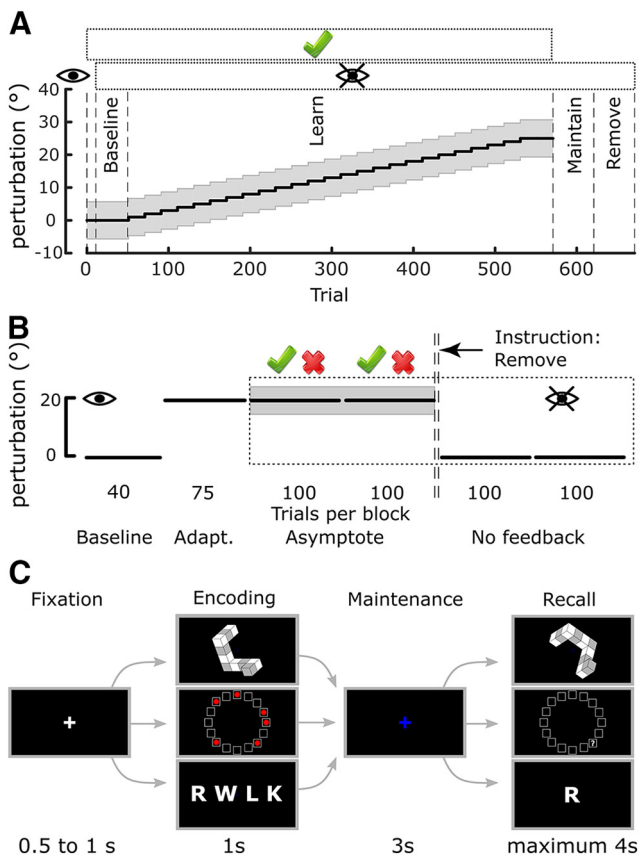


Figure 1. Experimental design. **A**, Time course of the Acquire task with the different experimental periods labeled. Gray region represents the reward region, which gradually rotated away from the visual target after the initial baseline period. Rectangle enclosing the green tick above the axes represents trials in which reward was available. Rectangle with the “eye” symbol represents when vision was not available. **B**, Time course of the Preserve task. After adapting to an initial rotation with vision available, vision was removed (eye symbol) and reward-based feedback was introduced (tick and cross above the axes). Before the no-feedback blocks, participants were instructed to remove any strategy they had been using. **C**, WMC tasks. The three tasks followed a stereotyped timeline with only the items to be remembered differing. Each trial consisted of four phases (Fixation, Encoding, Maintenance, and Recall) with the time allocated to each displayed below.

participants only received binary feedback. The baseline block consisted of the first 40 trials under binary feedback. During this period, the reward region remained centered on the visible target. Subsequently, unbeknownst to the participant, the reward region rotated in steps of 1° every 20 trials; the direction of rotation was counterbalanced across participants. After reaching a rotation of 25°, the reward region was held constant for an additional 20 trials. Performance during these last 20 trials was used to determine overall task success. Subsequently, binary feedback was removed, and participants were instructed to continue reaching as they were (maintain block) for the following 50 trials. Following this, participants were then informed that the reward region shifted during the experiment but not of the magnitude or the direction of the shift. They were then instructed to return to reaching in the same manner as they were at the start of the experiment (remove block, 50 trials). During the learning phase of the task, participants were given a 1 min rest after Trials 190 and 340.

Preserve task. Participants performed 515 trials in total. On each trial, participants were instructed to make a rapid “shooting” movement that passed through a target (white circle, radius: 0.125 cm) visible on the screen. The starting position for each trial was indicated by a white square (width: 1 cm) ~30 cm in front of the midline, and the target was located at angle of 45° from the perpendicular in a counterclockwise direction at a distance of 8 cm. The position of the tracking device attached to the fingertip was displayed as a cursor (green circle, radius: 0.125 cm). When

the radial distance of the cursor from the starting position exceeded 8 cm, the cursor feedback disappeared, and the end position was displayed instead.

First, participants performed a baseline period of 40 trials, during which the position of the cursor was visible, and the cursor accurately reflected the position of the fingertip. In the adaptation block (75 trials), participants were exposed to an abruptly introduced 20° clockwise visuo-motor rotation of the cursor feedback (Fig. 1B). Subsequently, all visual feedback of the cursor was removed, and participants received only binary feedback. If the end position of the movement fell within a reward region, the trial was considered successful and a tick was displayed; otherwise, a cross was displayed. The reward region was centered at a clockwise rotation of 20° with respect to the visual target with a width of 10°; that is, it was centered on the direction that successfully accounted for the previously experienced visuomotor rotation. Binary feedback was provided for 200 trials divided into 2 blocks of 100 trials (asymptote blocks). Furthermore, participants experienced 2 “refresher” trials for every 10 trials, where rotated visual feedback of the cursor position was again accessible (Shmuelof et al., 2012; Codol et al., 2018). This represents an increase compared with Codol et al. (2018) because participants in this study tended to have poorer performance under binary feedback, possibly due to a fatigue effect following the WM tasks (Anguera et al., 2012) (see Discussion). Finally, two blocks (100 trials each) with no performance feedback were used to assess retention of the perturbation (no-feedback blocks). Before the first of those two blocks, participants were informed of the visuomotor rotation, asked to stop accounting for it through aiming off target and to aim straight at the target.

WM tasks

Participants performed three WM tasks, all of which followed the same design with the exception of the nature of the items to be remembered (Fig. 1C). All WM tasks were run using MATLAB (The MathWorks) and Psychophysics toolbox (Brainard, 1997). At the start of each trial, a white fixation cross was displayed in the center of the screen for a period of 0.5–1 s randomly generated from a uniform distribution (fixation period in Fig. 1C). In the encoding period, the stimulus to be remembered was displayed for 1 s and then subsequently replaced with a blue fixation cross for the maintenance period, which persisted for 3 s. Finally, during the recall period, participants were given a maximum of 4 s to respond by pressing one of three keys on a keyboard with their dominant hand. The “1” key indicated that the stimuli presented in the recall period was a “match” to that presented in the encoding period; the “2” key indicated a “nonmatch”; and “3” indicated that the participant was unsure as to the correct answer. Each WM task contained 5 levels of difficulty with the 12 trials presented for each, 6 of which were trials in which “match” was the correct answer and 6 in which “nonmatch” was the correct answer. Consequently, each WM task consisted of 60 trials, and the order in which the tasks were performed was pseudorandomized across participants. Before the start of each task, participants performed 10 practice trials to familiarize themselves with the task and instructions. For both the Acquire and Preserve tasks, the WM tasks were performed in the same experimental session as the reaching. However, in the case of the Acquire task, the WM tasks were performed after the reaching task, whereas for the Preserve task the WM tasks were performed first.

In the RWM task (Fig. 1C, top row), the stimuli consisted of six 2D representations of 3D shapes drawn from an electronic library of the Shepard and Metzler type stimuli (Peters and Battista, 2008). The shape presented in the recall period was always the same 3D shape presented in the encoding period after undergoing a screen-plane rotation of 60°, 120°, 180°, 240°, or 300°. In “match” trials, the only transform applied was the rotation; however, in “nonmatch” trials, an additional vertical axis mirroring was also applied. The difficulty of mental rotation has been demonstrated to increase with larger angles of rotation (Shepard and Metzler, 1971); therefore, the different degrees of rotation corresponded to the 5 levels of difficulty. However, given the symmetry of two pairs of rotations (60 and 300, 120 and 240), these 5 levels were collapsed to 3 for analysis.

In the SWM task (Fig. 1C, middle row), stimuli in the encoding period consisted of a variable number of red circles placed within 16 squares

arranged in a circular array (McNab and Klingberg, 2008). In the recall period, the array of squares was presented without the red circles; and instead, a question mark appeared in one of the squares. Participants then answered a question (“Was there a red dot in the square marked by a question mark?”) by pressing a corresponding key. In “match” trials, the question mark appeared in one of the squares previously containing a red circle; and in “nonmatch” trials, it appeared in a square that was previously empty. Difficulty was scaled by varying the number of red circles (i.e., the number of locations to remember) from 3 to 7.

In the VWM task (Fig. 1C, bottom row), participants were presented with a list of a variable number of consonants during the encoding period. In the recall period, a single consonant was presented, and participants answered the question (“Was this letter included in the previous array?”). Thus, the letter could either be drawn from the previous list (“match” trials) or have been absent from the previous list (“nonmatch” trials). Difficulty in this task was determined by the length of the list to be remembered, ranging from 5 to 9.

Both the SWM and RWM tasks have been suggested to fall under the general umbrella term of spatial ability (Buszard and Masters, 2018). However, Miyake et al. (2001) suggest that, although both RWM and short-term storage of spatial information (i.e., SWM) are within the spatial domain, RWM appears to rely more heavily on executive function and SWM on basic short-term storage of spatial information. Furthermore, previous studies have found relationships between motor learning and this SWM task (Christou et al., 2016; Vandevorde and Orban de Xivry, 2019) and tasks similar to our RWM task (Anguera et al., 2010). Therefore, we included both tasks to investigate whether there was any severability in their relationships with reaching performance and leveraged our use of two separate reaching tasks and large cohorts to probe whether this was due to specific task parameters.

Genetic sample collection and profiling

COMT is thought to affect DA function mainly in the PFC (Egan et al., 2001; Goldberg et al., 2003), a region known for its involvement in WM and strategic planning (Anguera et al., 2010; Doll et al., 2015), whereas DARPP32 and DRD2 act mainly in the basal ganglia to promote exploitative behavior, possibly by promoting selection of the action to be performed (Frank et al., 2009). Consequently, we focused here on single nucleotide polymorphisms (SNPs) related to those genes: RS4680 (COMT) and RS907094 (DARPP32). Regarding DRD2, there are two potential SNPs available, RS6277 and RS1800497. Although previous studies focusing on exploration and exploitation have assessed RS6277 expression (Frank et al., 2007, 2009; Doll et al., 2016), it should be noted that this SNP varies greatly across ethnic groups, with some allelic variations being nearly completely absent in non-Caucasian-European groups (e.g., see RS6277 in Auton et al., 2015). This has likely been inconsequential in previous work, as Caucasian-European individual represented the majority of sampled groups; here, however, this represents a critical shortcoming, as we aim at investigating a larger and more representative population, including other ethnic groups. Consequently, we based our analysis on the RS1800497 allele of the DRD2 gene (Pearson-Fuhrhop et al., 2013).

At the end of the task, participants were asked to produce a saliva sample, which was collected, stabilized, and transported using Oragene. DNA saliva collection kits (OG-500, DNAGenotek). Participants were requested not to eat or drink anything except water for at least 2 h before sample collection. Once data collection was completed across all participants, the saliva samples were sent to LGC (<https://www.lgcgroup.com/>) for DNA extraction (per Oragene protocols: <https://www.dnagenotek.com/>) and genotyping. SNP genotyping was performed using the KASP SNP genotyping system. KASP is a competitive allele-specific PCR incorporating a FRET quencher cassette. Specifically, the SNP-specific KASP assay mix (containing two different, allele-specific, competing forward primers) and the universal KASP master mix (containing FRET cassette plus Taq polymerase in an optimized buffer solution) were added to DNA samples and a thermal cycling reaction performed, followed by an endpoint fluorescent read according to the manufacturer’s protocol. All assays were tested on in-house validation DNA before being run on project samples. No-template controls were used, and 5% of the samples had duplicates included on each plate to

enable the detection of contamination or nonspecific amplification. All assays had >90% call rates. Following completion of the PCR, all genotyping reaction plates were read on a BMG PHERAStar plate reader. The plates were recycled until a laboratory operator was satisfied that the PCR had reached its endpoint. In-house Kraken software then automatically called the genotypes for each sample, with these results being confirmed independently by two laboratory operators. Furthermore, the duplicate saliva samples collected from 5% of participants were checked for consistency with the primary sample. No discrepancies between primary samples and duplicates were discovered.

Data analysis

Acquire task. Reach trials containing movement times over 0.6 or <0.2 s were removed from analysis (6.9% of trials). The endpoint angle of each movement was defined at the time when the radial distance of the cursor exceeded 10 cm. This angle was defined in relation to the visible target with positive angles indicating clockwise rotations. Endpoint angles and target angles for participants who experienced the counterclockwise rotations were sign-transformed. The explicit component of retention was defined as the difference between the mean reach angle of the maintain block and the remove block, whereas the implicit component was the difference between the mean reach angle of the remove block and baseline (Werner et al., 2015). Participants that achieved a mean reach angle within the reward region during the final 20 trials before the maintain block were considered “successful” in learning the rotation; otherwise, they were considered “unsuccessful.” As in Holland et al. (2018), for unsuccessful participants, the largest angle of rotation at which the mean reach angle fell within the reward region was taken as the end of successful performance, and only trials before this point were included for further analysis. Success rate (SR) was defined as the percentage of trials during the learning blocks in which the endpoint angle was within the reward region. To examine the effect of reward on the change in endpoint angle on the subsequent trial, we examined the magnitude and variability of changes in endpoint angle between consecutive trials (Therrien et al., 2016, 2018; Holland et al., 2018; Sidarta et al., 2018). To calculate the median absolute change following rewarded (ΔR) and unrewarded (ΔP) trials, we extracted the changes in reach angle following each trial type and calculated the median of the absolute values of these changes for each participant. These measures therefore represent the median of the magnitude of changes in reach angle, regardless of direction. Furthermore, to examine the variability of trial-by-trial adjustments ($MAD[\Delta R]$ and $MAD[\Delta P]$ for rewarded and unrewarded trials, respectively), we calculated the median absolute deviation of the changes in reach angle. It is important to note that ΔR and ΔP are calculated from the absolute magnitude of the changes in reach angle, whereas $MAD[\Delta R]$ and $MAD[\Delta P]$ are calculated from the nonabsolute values (including the direction of change).

Preserve task. Reach trials containing movement times over 1 s were removed from analysis (2.38% of trials). The endpoint angle for each movement was defined at the time that the radial distance of the cursor from the start position exceeded 8 cm. Trials in which the error was >80° were excluded from further analysis (0.94% of trials). As in Codol et al. (2018), learning rate was calculated by fitting an exponential function to the angular error between cursor and target for trials in the adaptation block, with the β value taken as the learning rate (mean $R^2 = 0.34 \pm 0.15$). The β estimates attained from all fits were first sign-transformed and then log-transformed to counteract skewness before entering the regression analysis. Using this method, a value close to 0 indicated faster learning, whereas more negative values indicated slower learning. Similar to Codol et al. (2018), SR, corresponding to percentage of rewarded trials, was measured separately in the first 30 and last 170 trials of the asymptote blocks and labeled early and late SR, respectively. This reflects a dichotomy between a dominantly exploration-driven early phase and a later exploitation-driven phase. The analysis of changes in reach angle (ΔR and ΔP) was confined to the last 170 trials of the asymptote blocks. Implicit retention was defined as the difference between the average baseline reach direction and the mean reach direction of the last 20 trials of the last no-feedback block (Codol et al., 2018). Analysis of changes in

reach angle following rewarded trials was not preregistered but was included *post hoc*.

Exploratory analysis of reaching data. To understand which outcome variables in the reaching tasks were predictive of overall task success, we split the learning period into two sections for every participant. We assessed trial-by-trial changes in endpoint angle in the first section and compared them with SR in the second section. For the Acquire task, we assessed trial-by-trial adjustments during the learning block, excluding the final 20 trials, and compared them with SR in the last 20 trials of the learning block. In the Preserve task, we measured adjustments in the first 100 trials of the asymptote blocks and compared them with SR in the last 100 trials of the asymptote blocks.

WM tasks. WM performance was defined as the average percentage of correct responses across the three highest levels of difficulty for each task. In the case of the RWM task, the symmetrical arrangement of the angles of rotation in effect produced three levels of difficulty and therefore all trials were analyzed.

Genetics. Genes were linearly encoded, with heterozygote alleles being 0, homozygote alleles bearing the highest dopaminergic state being 1, and homozygote alleles bearing the lowest dopaminergic state being -1 (Table 1). All groups were assessed for violations of the Hardy–Weinberg equilibrium. The participant pool in the Preserve task was in Hardy–Weinberg equilibrium for all three genes considered. In the Acquire task population, COMT and DRD2 were in Hardy–Weinberg equilibrium, but DARPP32 was not ($p = 0.002$), with too few heterozygotes. Therefore, the DARPP32 alleles were recoded, with the heterozygotes (0) and the smallest homozygote group (C:C, -1) combined and recoded as 0.

Statistical analysis

Regressions were performed using the linear Lasso method (Tibshirani, 1996) (*lasso* function in MatLab's *Statistics and Machine Learning Toolbox*). Lasso regression uses a shrinkage method that allows for some predictors to be shrunk to a value of 0, effectively removing them from the regression model. Therefore, the method acts as a selection method for predictors in a way analogous to stepwise regression. We used a 10-fold cross validation approach to calculate the mean squared error (MSE) over a range of values of a penalty term λ . Specifically, as λ increases, the shrinkage of predictor values increases. For $\lambda = 0$, the model reduced to a standard linear regression, as all predictors were included without any shrinkage. Cross validation protects against the problem of overfitting by calculating the MSE on data “unseen” by the model during fitting. For any given outcome variable, if its MSE(λ) function exhibited a minimum value within its defined boundaries, the model associated with that minimum value was considered selected. If no minimum was observed, this signified that an empty model was a better fit than any other possible model. If such minimum was detected in the MSE(λ) function, the β estimates from that model (i.e., at that value of λ) were taken. We repeated this procedure 1000 times to obtain the distribution of the true β from the estimates (Hastie et al., 2015). In order for a potential variable to be considered a selected predictor, that predictor should be selected (i.e., $\beta \neq 0$) in at least 80% of the repetitions. The threshold of 80% was chosen as to maintain sufficient sensitivity while still returning relatively sparse models. We report the median β estimate in the text for all selected predictors.

To understand what genetic and WM factors are predictive of performance in the Acquire task, we performed a lasso regression of the seven predictors (three allelic variations, three WM, and ethnicity) onto each of several outcome measures representative of performance: SR, implicit and explicit retention, ΔR , MAD[ΔR], ΔP , and MAD[ΔP]. For the Preserve task, we performed separate lasso regressions using the same seven predictors for the following outcome variables: baseline reach direction as a control variable, learning rate in the adaptation block, early and late SR in the asymptote blocks (first 30 and last 170 trials) (Codol et al., 2018), retention in the no-feedback blocks, and ΔR and ΔP during the asymptote blocks. We adopted a parsimonious approach when interpreting the results of the regression analysis and gave particular credence to results reproduced by the analysis across both tasks.

Before the regression analysis, all predictors and predicted variables were standardized (z -scored). For all nonordinal variables, individual

Table 1. Coding for SNPs^a

SNP	Location	Allele code -1	Allele code 0	Allele code 1
rs4680	COMT	G:G (val:val) 31, 33	A:G (met:val) 68, 61	A:A (met:met) 17, 21
rs1800497	DRD2	T:T (lys:lys) 8, 7	T:C (lys:glu) 48, 51	C:C (glu:glu) 64, 62
rs907094	DARPP32	C:C 10, 21	C:T 54, 38	T:T 56, 62

^aThe name of the SNP is provided along with the code assigned to each allele. The numbers indicate the counts for the specific allele in the two tasks (Preserve, Acquire).

data were considered outliers if further than 3 SDs from the mean and were removed before standardization. Multicollinearity of predictors was also assessed before regression with Belsley Collinearity Diagnostics (*collintest* function in MATLAB's *Econometrics Toolbox*), and no predictors were found to exhibit condition indexes >30 , indicating acceptable levels of collinearity. When considering retention for both tasks, unsuccessful participants were removed from the regression analysis. We further characterized the relationships between predictor variables by combining the data for the two tasks for the WM tasks and the genetic codes ($N = 241$). We analyzed relationships between the WM tasks with correlations and between genetics and WM tasks with one-way ANOVAs.

Exploratory mediation analysis. We performed a mediation analysis to test whether the relationship between SWM and SR was mediated by ΔR . Our hypothesis was that higher SWM enables smaller changes after correct trials (ΔR), and this then explains the relationship between SWM and SR. To ensure that separate trials were used in the calculation of ΔR and SR, we split the trials into two equally sized folds. The SR was then calculated for onefold as a percentage of correct trials, and ΔR was calculated as the median absolute change of reach angle after correct trials in the other fold. For the Acquire task, only successful subjects were included in the mediation analysis. We used Baron and Kenny's three step mediation analysis (Baron and Kenny, 1986): first regress SR on SWM, then regress ΔR on SWM, and finally regress SR on both SWM and ΔR . Subsequently, we performed a Sobel test to determine whether there was a significant reduction in the relationship between SWM and SR when including ΔR . The Sobel test examines whether the amount of variance in SR explained by SWM is significantly reduced by including the mediator (Sobel, 1986). For a significant effect to be found, SWM must be a significant predictor of ΔR and ΔR must also be a significant predictor of SR after controlling for the effect of SWM on SR. We repeated this procedure 1000 times with the allocation of trials to each fold randomized on each repetition. We present results in terms of the 95% CIs for the R^2 values for each of the regressions and the median p value of the Sobel test, along with the associated 95% CIs. An alternative possibility to the hypothesized model is that the relationship between SWM and ΔR is mediated by SR. To compare the size of the mediation effect for these alternate models, we follow the Mackinnon and Dwyer (1993) procedure and normalize the size of the indirect effect by dividing it by the sum of the direct and indirect effects. This analysis allows to express the mediation effect in terms of percentage of the total effect. We present the median of the normalized value for the 1000 repetitions on both the hypothesized and alternate models.

Results

Acquire task

In the Acquire task, participants had to learn to compensate for a secretly shifting reward region to obtain successful feedback (Figs. 2, 3). As in Holland et al. (2018), approximately one-fourth (28.1%) of participants failed to learn to compensate for the full extent of the rotation (Fig. 3A). The inability of a significant proportion of participants to learn the full extent of the rotation is also consistent with previous reports in reward-based motor learning paradigms (Saijo and Gomi, 2010; Therrien et al., 2016, 2018; Codol et al., 2018; Cashaback et al., 2019). Successful participants retained most of the learnt rotation (mean $80.7 \pm 28.0\%$

SD) in the maintain block. This level of retention is in accordance with that reported previously in similar paradigms (Therrien et al., 2016; Holland et al., 2018). However, upon being asked to remove any strategy they had been using, their performance returned to near-baseline levels. Consequently, a large explicit component to retention was found for successful participants (Fig. 3B), whereas both successful and unsuccessful participants manifest a small but nonzero implicit component ($t_{(86)} = 9.90$, $p = 7.43 \times 10^{-16}$, $d = 1.061$ and $t_{(33)} = 4.53$, $p = 7.39 \times 10^{-5}$, $d = 0.776$, respectively; Figure 3C). The persistent implicit retention is a common finding of retention periods in which no visual feedback is provided and may reflect a combination of implicit reinforcement (Shmuelof et al., 2012), use-dependent plasticity (Diedrichsen et al., 2010), perceptual bias (Vindras et al., 1998), or perceptual recalibration (Modchalingam et al., 2019). Furthermore, in accordance with Holland et al. (2018), we found that participants made larger ($t_{(120)} = 15.80$, $p = 4.32 \times 10^{-31}$, $d = 1.900$) and more variable changes in reach angle following unrewarded trials ($t_{(120)} = 14.54$, $p = 3.144 \times 10^{-28}$, $d = 1.667$; Fig. 3D–G). However, in participants who would go on to fail, the posterror adjustments were smaller than in successful participants ($t_{(119)} = 3.33$, $p = 0.001$, $d = 0.672$; Fig. 3D). Changes following rewarded trials were similar between the groups ($t_{(119)} = 0.71$, $p = 0.48$, $d = 0.143$; Fig. 3F, G). The results obtained in this sample ($N = 121$) therefore replicate results from a previous study ($N = 30$) and provide further confirmation that performance in this task is fundamentally explicitly driven (Holland et al., 2018).

To understand what genetic and WM factors are predictive of performance in the reaching task, we performed a lasso regression of the seven predictors (three allelic variations, three WM, and ethnicity) onto each of several outcome measures representative of performance: SR, implicit and explicit retention, ΔR , $MAD[\Delta R]$, ΔP , and $MAD[\Delta P]$.

For SR, SWM, RWM, and DRD2 were selected as predictors (median $\beta = 0.31$, 0.06, and 0.03, respectively; Fig. 4A), with the strongest predictor being SWM. Figure 5 displays the effect of the strongest predictor selected for each outcome variable and shows that there was a positive relationship between SWM and SR (Fig. 5A). To ensure that the relationship between SWM and SR was not due to failure at a later point in the task, SR was measured during the initial period in which subjects who could not fully account for the displacement are still successful; for those who could, the full learning block was included.

Next, retention was assessed by splitting up the explicit and implicit components, such as in Holland et al. (2018). No predictor was related to the implicit component, but the explicit com-

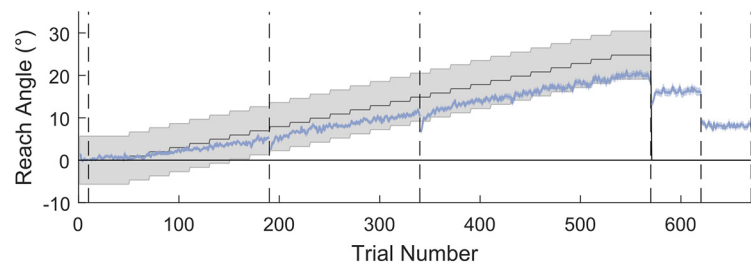


Figure 2. Reaching performance in the Acquire task. Gray region represents the gradually rotating rewarded region. Blue line indicates mean reach angle for each trial. Shaded blue region represents SEM. Vertical dashed lines indicate different experiment blocks or breaks. Average performance for the full cohort falls within the reward region and demonstrates a clear reduction in reach angle when asked to remove strategy. $N = 121$.

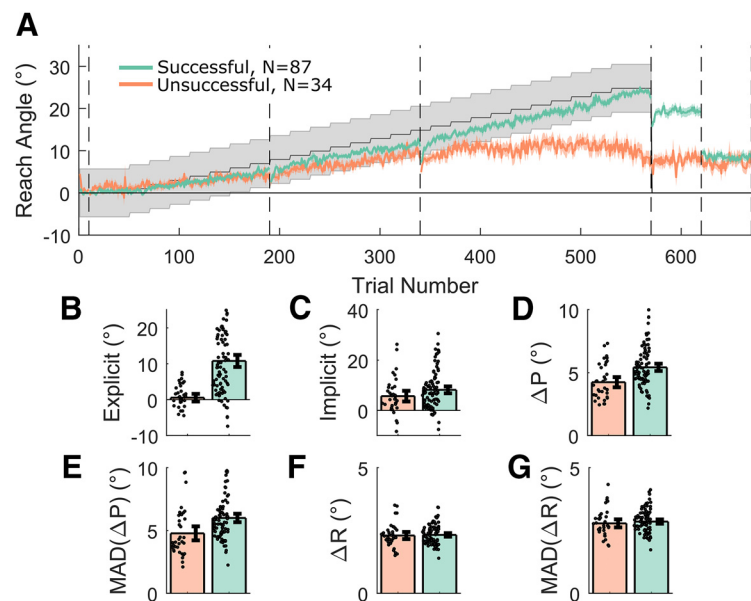


Figure 3. Acquire task split by success at final angle. **A**, Average reach angle for the successful (green) and unsuccessful (orange) groups. Shaded regions represent SEM. Gray shaded region represents the rewarded region. Despite similar initial performance, a clear divergence can be seen between the two groups and an explicit component to retention is only visible in the successful group, whereas implicit retention is similar between groups. **B–G**, Subplots displaying derived measures, which acted as outcome variables for the regression analysis, separated into successful and unsuccessful participants overlaid with individual data points. Error bars indicate 95% bootstrapped CIs. ΔR and ΔP refer to changes made in reach angle after rewarded and unrewarded trials, respectively. **D, F**, Bar plots represent the median absolute change. **E, G**, Median absolute deviation of the changes in angle after each trial type.

ponent was strongly and positively associated with RWM ($\beta = 0.27$; Figs. 4B, 5B) with a weaker association between DARPP32 and explicit retention ($\beta = 0.03$). These results suggest positive relationships for both RWM and SWM with task performance: greater RWM predicts a greater contribution of explicit processes to learning, whereas greater SWM predicts a greater percentage of correct trials.

In Holland et al. (2018), the amplitude of the changes in reach angle participants made following unrewarded trials was found to be predictive of task success; that is, greater ΔP was predictive of an increased chance of overall task success. Thus, it could be that RWM and SWM, which are shown to associate with performance in this study, are themselves predictors of changes in reach angle. Conformingly, the regression results demonstrated that a large ΔR was inversely related to SWM ($\beta = -0.11$; Figs. 4F, 5D), as was $MAD[\Delta R]$ ($\beta = -0.17$; Figs. 4G, 5E). The results indicate that greater SWM was predictive of smaller and less variable changes in reach angle after successful trials, suggesting high

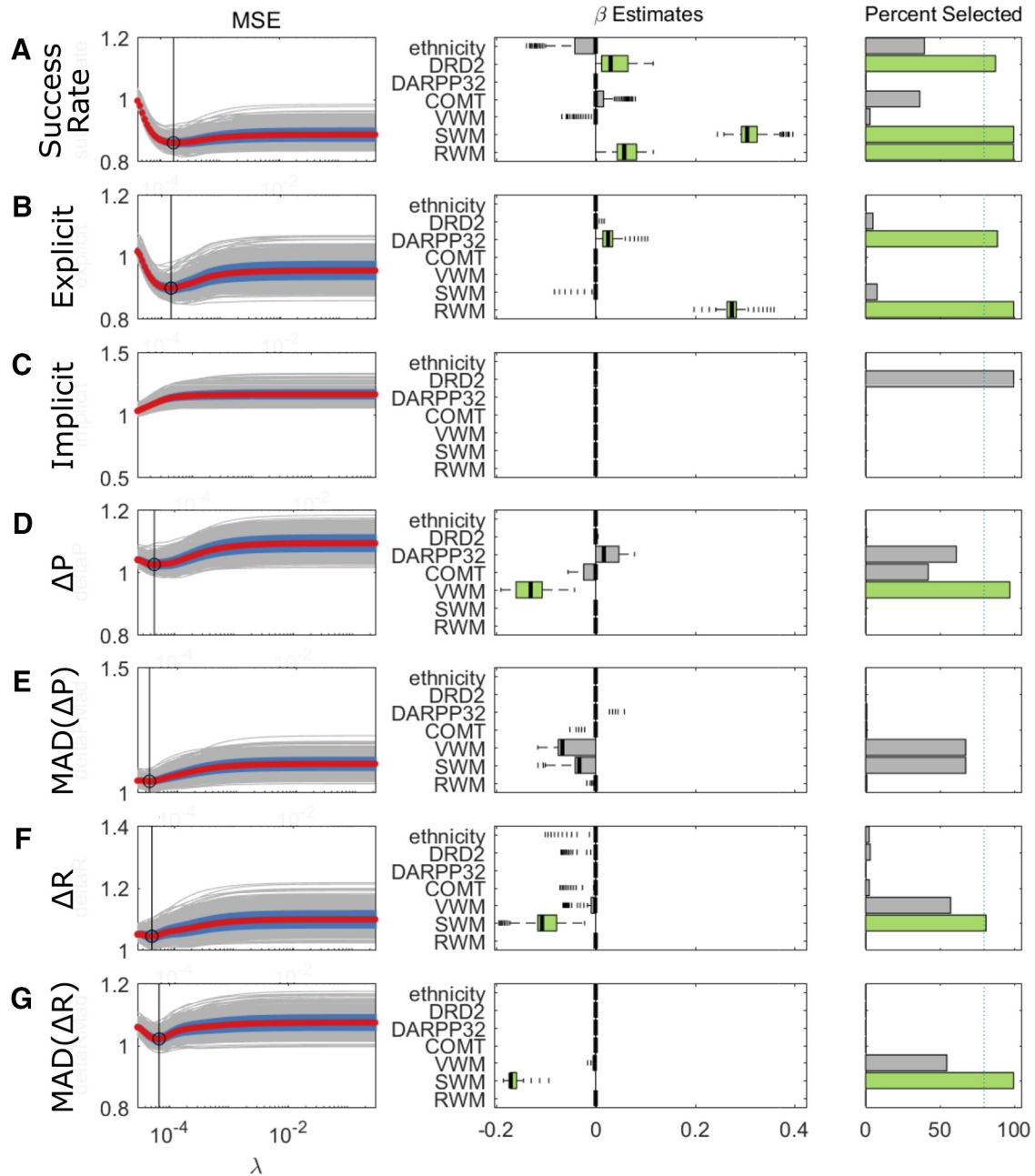


Figure 4. Lasso regression results for the Acquire task. **A–G**, Each row represents the results from one outcome variable. Left column represents the MSE as a function of changing the shrinkage parameter λ , with larger values of λ representing greater penalization and sparser models. A minimum in the MSE within its defined boundaries indicates the suitability of that choice of λ (indicated with a vertical line). Given the presence of a minimum, the values of the β for each predictor are taken. We performed 1000 repetitions of the lasso regression for each outcome variable. Middle, Box plots represent the distribution of the coefficient estimates. Rightmost column represents the percentage of times that the individual predictors were assigned nonzero coefficients. We used a threshold of 80% (indicated with a dashed vertical line) to signify that a particular predictor was robustly selected, and these variables are highlighted in green. Median absolute deviation of change in reach angle after rewarded (MAD[ΔR]) and unrewarded (MAD[ΔP]) trials.

SWM enables the maintenance of rewarding reach angles. It was also found that changes in reach angle following unrewarded trials (ΔP) were negatively associated with VWM ($\beta = -0.13$, Figs. 4D, 5C). This result was unexpected as it suggests that greater WMC predicts smaller changes following unrewarded trials, whereas previous results suggest a positive relationship between the amplitude of these changes and overall task success. Although the difference may be due to the domain of WM under consideration, it is unclear as to the reason for this relationship. Another important aspect of the analysis of trial-by-trial changes to control for is that the numbers of trials analyzed and their

phase in the experiment differs between successful and unsuccessful subjects. Therefore, we repeated the Lasso regression while only including successful subjects. The predictors that were selected were identical to those selected when using the full participant pool.

Overall, WM (in particular RWM and SWM) successfully predicted various aspects of performance in the Acquire task, while genetic predictors generally failed to do so. Specifically, greater SWM predicted smaller and less variable changes after correct trials. This suggests that SWM underlies one’s capacity to preserve and consistently express an acquired reach direction to ob-

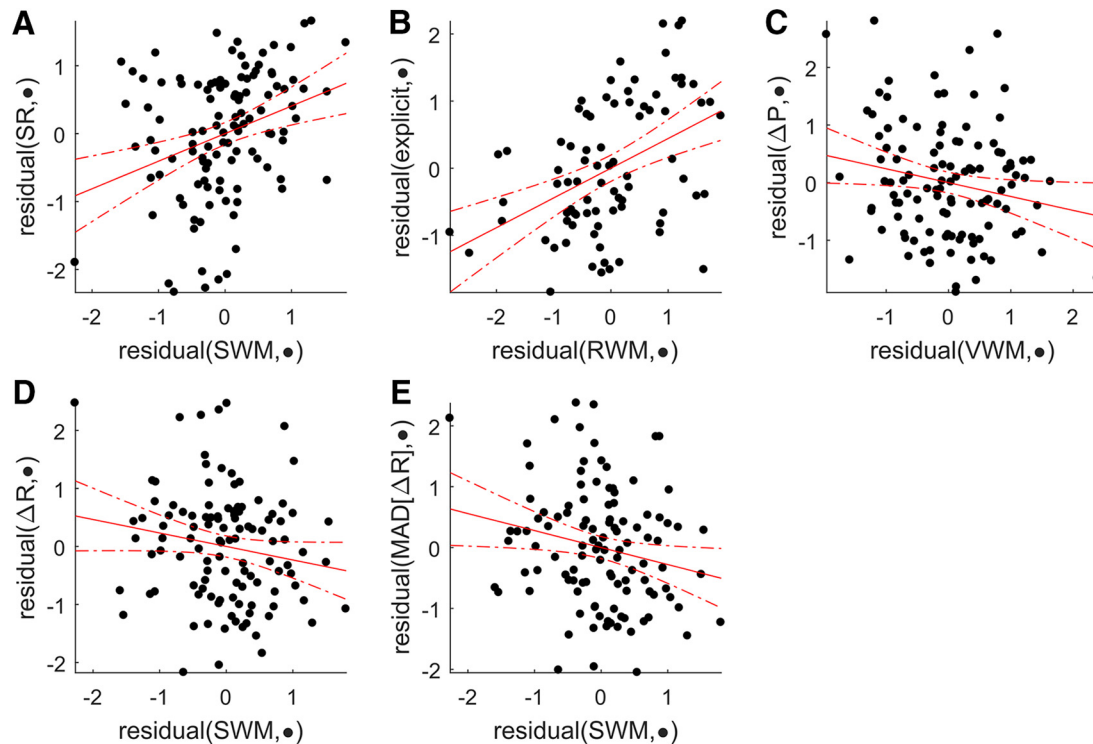


Figure 5. Added variable plots for selected predictors in the Acquire task. **A–E**, Each plot represents the relationship between the strongest predictor selected by the lasso regression (x axis) and the corresponding outcome variable (y axis). Added variable plots represent the residuals of regressing the response variable with all remaining independent variables, and the residuals of the regression of the selected predictor to the remaining predictors. The resulting relationship corresponds to the effect of the selected predictor on the outcome measure after controlling for the remaining predictors. MAD(ΔR), Median absolute deviation of change in reach angle after rewarded trials.

tain reward. Furthermore, SWM also directly predicted SR. In addition, greater RWM was a strong predictor of explicit control. The inverse relationship between VWM and the magnitude of changes after unrewarded trials was unexpected. However, one possible explanation is that participants with poorer WMC make larger errors, which require larger corrections.

Preserve task

In this task, we addressed how well participants can maintain a previously learnt adaptation after transitioning to binary feedback. As participants are unable to compensate for a large abrupt displacement of a hidden reward region (Manley et al., 2014; van der Kooij and Overvliet, 2016), participants first adapted to an abruptly introduced 20° clockwise rotation with full vision of the cursor available. Subsequently, visual feedback of the cursor position was replaced with binary feedback; participants were rewarded if they continued reaching toward the same angle that resulted in the cursor hitting the target during the adaptation phase. Overall, participants adapted to the visuomotor rotation successfully (Figs. 6, 7A–C) before transitioning to the binary feedback-based asymptote blocks. However, from the start of the asymptote blocks onward, participants exhibited very poor performance, expressing an average $45.0 \pm 24.2\%$ (SD) SR when considering all 200 asymptote trials (Figs. 6, 7A, D, E). We have previously shown, in Codol et al. (2018), that this drop in performance (Shmuelof et al., 2012) represents exploratory behavior that arises due to a lack of transfer of the cerebellar memory

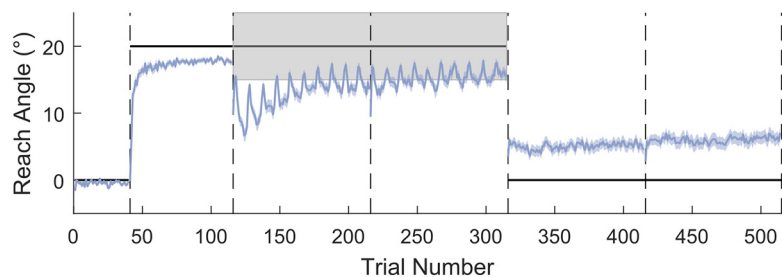


Figure 6. Reaching performance in the Preserve task. Gray shaded area represents the rewarded region. Thick black line indicates the perturbation. Vertical dashed lines indicate block limits. Blue line indicates mean reach angle for every trial. Blue shaded areas represent SEM. After successfully adapting to the visuomotor rotation, performance deteriorates at the onset of binary feedback; subsequently, SR increases toward the end of the asymptote blocks. Following the removal of all feedback, and the instruction to remove any strategy, a small amount of implicit retention remains. $N = 120$.

between the two contexts. Separating successful and unsuccessful participants (40% SR cutoff; Fig. 7A) revealed that successful participants expressed behavior greatly similar to that observed in Codol et al. (2018), in which unsuccessful participants were excluded, using the same cutoff (40% SR). The “spiking” behavior observed in reach angles during the asymptote blocks (Fig. 7A) is due to the presence of the “refresher” trials, with the large positive changes in reach angle corresponding to trials immediately following the refresher trials. This pattern of behavior is particularly pronounced in the unsuccessful participants. Finally, participants demonstrated at least a residual level of retention, even after being instructed to remove any strategy they had used ($t_{(69)} = 7.268$, $p = 3.345 \times 10^{-10}$, $d = 0.869$; Fig. 7A, F). Therefore, the results obtained in this sample ($N = 120$) replicate results from a previous study (Codol et al., 2018) ($N = 20$, binary feedback-Remove group) and provides further confirmation that

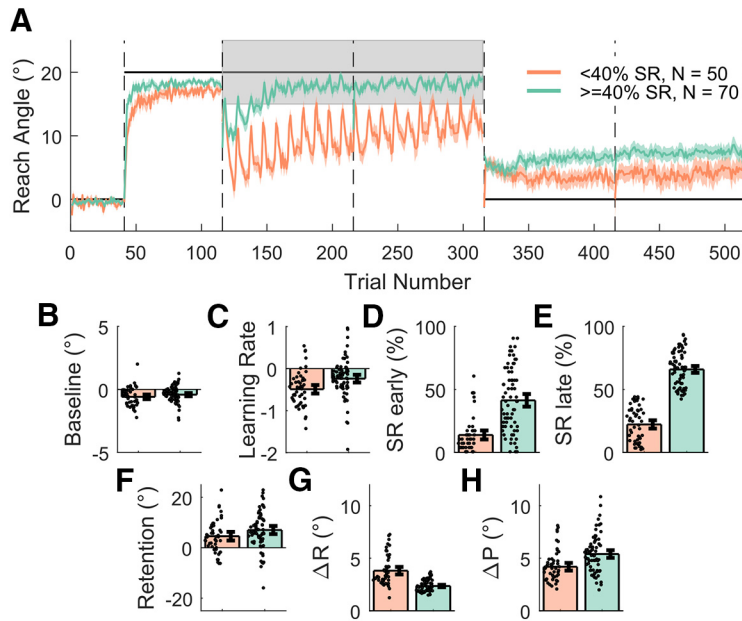


Figure 7. Preserve task split into two groups on the basis of SR. **A**, Shaded regions represent SEM. **B–H**, Derived variables, which acted as outcome variables for the regression analysis, for the two groups. Error bars indicate 95% bootstrapped CIs. Individual data points are displayed.

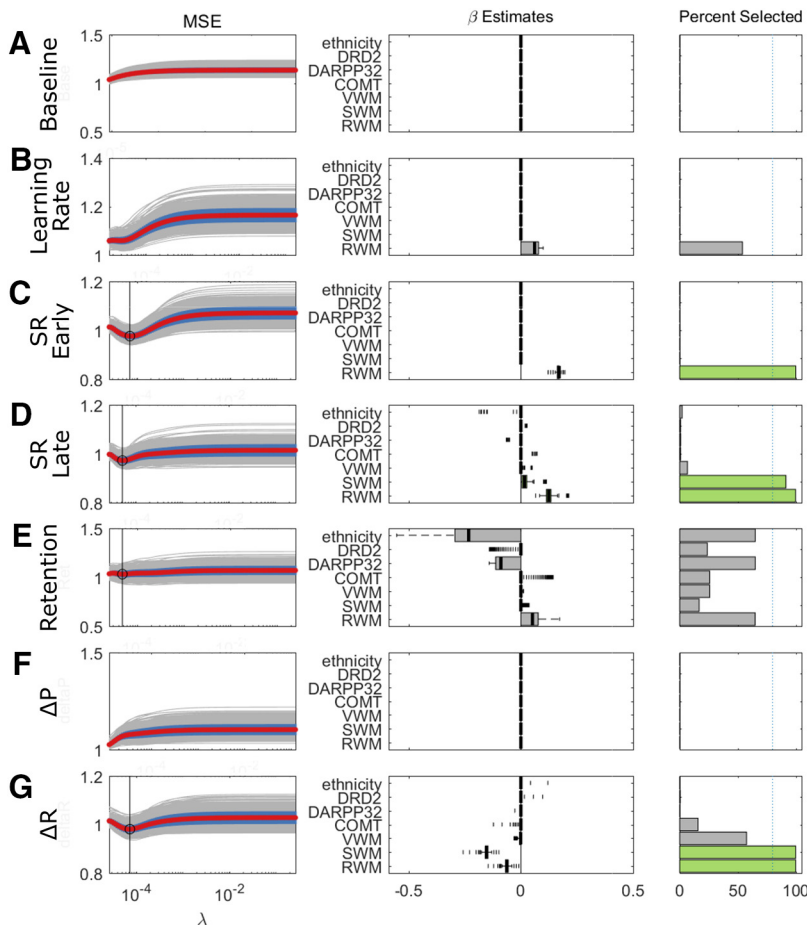


Figure 8. Results of the lasso regressions for the Preserve task. The format is identical to Figure 4, with each row (**A–G**) representing the predictors of a single outcome measure. Green represents selected predictors. Middle, β estimates. Right, Probability of each predictor being selected.

performance in this task is fundamentally explicitly driven. It should also be noted that the successful group displayed higher implicit retention than the unsuccessful participants. As with the Acquire task, successful participants displayed larger changes in angle after unrewarded trials than their unsuccessful counterparts ($t_{(117)} = 3.847, p = 1.952 \times 10^{-4}, d = 0.717$; Fig. 7H). However, in contrast to the Acquire task, successful participants also displayed smaller changes in angle after rewarded trials ($t_{(115)} = -7.534, p = 1.218 \times 10^{-11}, d = 1.421$; Fig. 7G).

As in the Acquire task, we examined whether performance in any of the WM tasks or genetic profile could predict participants' behavior in the reaching task. We performed separate lasso regressions for the following outcome variables: baseline reach direction as a control variable, learning rate in the adaptation block, early and late SR in the asymptote blocks (first 30 and last 170 trials) (Codol et al., 2018), retention in the no-feedback blocks, and ΔR and ΔP during the asymptote blocks. The most striking result was that both early and late SR could be reliably predicted by RWM (early: $\beta = 0.17$, late: $\beta = 0.12$; Figs. 8C,D, 9A,B), with greater RWM associated with increased SRs. An additional positive relationship was found between SWM and SR but only during the later period ($\beta = 0.02$; Fig. 8C).

Genetic profile did not predict any aspect of performance. In contrast, greater SWM successfully predicted reduced ΔR ($\beta = -0.15$; Figs. 8G, 9C) similarly to the Acquire task. Additionally, there was a weaker relationship between RWM and ΔR ($\beta = -0.06$; Fig. 8G), which was absent in the Acquire task. Despite the presence of a local minimum in the MSE for the regression involving retention, no individual predictor was consistently selected in $>80\%$ of repetitions (Fig. 8E).

Overall, the regression results across both tasks exhibited a similar pattern, with greater RWM predicting improved performance on the reaching task and greater SWM predicting smaller changes in reach angle after rewarded trials. The weak relationships found between genetic variables and performance measures in the Acquire task (DRD2-SR and DARPP32-Explicit retention) were not replicated in the Preserve task, questioning the reliability of these relationships.

Furthermore, we analyzed the data using group lasso (Yuan and Lin, 2006; Boyd, 2010) regression to check for the possibility that our analysis was insensitive to categorical predictors (the genetic

variables). The group lasso is an extension to lasso regression in which predictor variables can be assigned to groups. Although each member of a group can be assigned a different β , the group lasso applies the regularization penalty to all members of the group, leading to the removal of all members of the group from the model at the same value of λ . We used reference dummy variable coding for each genetic variable and treated the dummy variables representing each SNP as a group for the purposes of the group lasso; this ensures that the dummy variables representing each genetic factor are removed from the regression at the same time. The results of the group lasso analysis replicate those of the standard lasso; furthermore, no genetic predictors were found for any outcome variable in either task. The results obtained for both tasks via the lasso regression methods are similar to those obtained using a stepwise regression procedure. All data and code are available online, including the procedures, results, and significance tests of the lasso and stepwise regression analysis.

Relationships between predictors

In the full sample ($n = 241$), we assessed the relationship between the predictor variables. Despite the collinearity of the variables being within recommended values for use in regression (see Materials and Methods), we did find significant relationships between all three WM tasks. VWM and SWM were the most closely correlated ($r = 0.393$, $p = 3.153 \times 10^{-10}$), followed by SWM and RWM ($r = 0.384$, $p = 7.491 \times 10^{-10}$), and finally RWM and VWM ($r = 0.189$, $p = 0.003$). When examining the relationships between genetics and WM tasks, only one relationship was significant (DRD2 and SWM, $F_{(236,2)} = 3.927$, $p = 0.021$). However, this relationship did not survive correction for multiple comparisons.

Partial correlation analysis

To understand whether the RWM and SWM measures have separable effects on the outcome measures considered here, we performed a partial correlation analysis examining the relationships between RWM, SWM, and SR in both tasks. After controlling for the effect of RWM, SWM remained significantly correlated with success in both tasks (Preserve: $r = 0.343$, $p = 0.005$; Acquire: $r = 0.488$, $p = 6.823 \times 10^{-6}$). However, the partial correlation between RWM and SR was not significant for either task, indicating that, even in the Preserve task, SWM plays a dominant role in determining SR.

Exploratory analysis

As a relationship exists between SWM and ΔR in both the Acquire and Preserve paradigms, we ran exploratory regressions to assess the relationship between ΔR and SR across both tasks. Since ΔR and SR are conceptually strongly related variables, and measuring on the same dataset would render them nonindependent, we split each individual's reaching data into two sections and assessed whether ΔR or ΔP in the first section could reliably

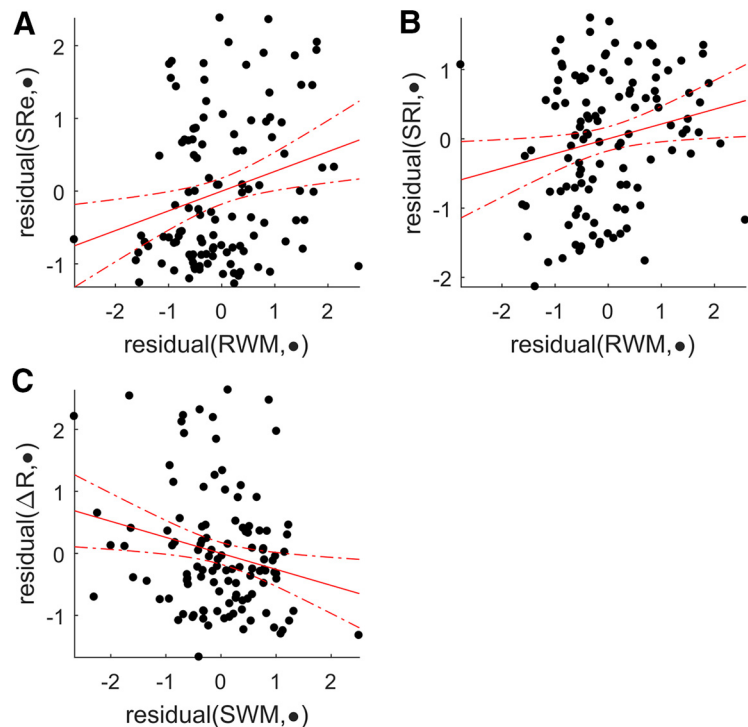


Figure 9. Added variable plots for selected predictors in the Preserve task. **A–C**, The effect of the considered predictor when accounting for the effect of all other predictors. Results are displayed for the strongest selected predictor for each outcome measure. SRe, Early SR; SRI, late SR.

predict SR in the second (for details, see Materials and Methods). Although we found no predictors of ΔP in our primary analysis, results here in combination with previous work (Holland et al., 2018) have demonstrated a link between ΔP and task success, with a greater ΔP indicative of greater success. Therefore, we also performed the same analysis for ΔP .

In the Acquire task, ΔR and ΔP in the first section of learning trials predicted SR in the final 20 trials, although ΔP appeared as the strongest predictor (ΔR : $\beta = -0.274$, $p = 0.015$; ΔP : $\beta = 0.581$, $p = 3.89 \times 10^{-6}$; Fig. 10A,B, yellow; Table 2). Similarly, for the Preserve task, ΔR and ΔP in the first half of asymptote trials predicted SR in the second half (ΔR : $\beta = -0.750$, $p = 1.07 \times 10^{-12}$; ΔP : $\beta = 0.229$, $p = 0.007$; Fig. 10A,B, pink; Table 2). In both tasks, the directions of these relationships were opposite; greater SR was predicted by smaller ΔR and greater ΔP . In summary, we found that, for both tasks, the magnitude of changes in behavior in response to rewarded and unrewarded trials early in learning was strongly predictive of future task success across both the Acquire and Preserve tasks.

Mediation analysis

Finally, to test whether the effect observed between SWM and SR was explained by an indirect effect through ΔR , we performed an exploratory mediation analysis on both tasks. For both the Acquire and Preserve tasks, the results indicate a significant proportion (median $p = 7.10 \times 10^{-4}$ and $p = 0.04$, respectively) of the relationship between SWM and SR can be explained by a mediation from SWM via ΔR to SR (Fig. 11). However, in the case of the Acquire task (Fig. 11A), a significant relationship between SWM and SR also remained, indicating that not all of the effect of SWM on SR could be explained by the indirect pathway. Of note, in the Preserve task (Fig. 11B), the SWM- ΔR relationship was weaker and was not significant on every repetition, occasionally leading

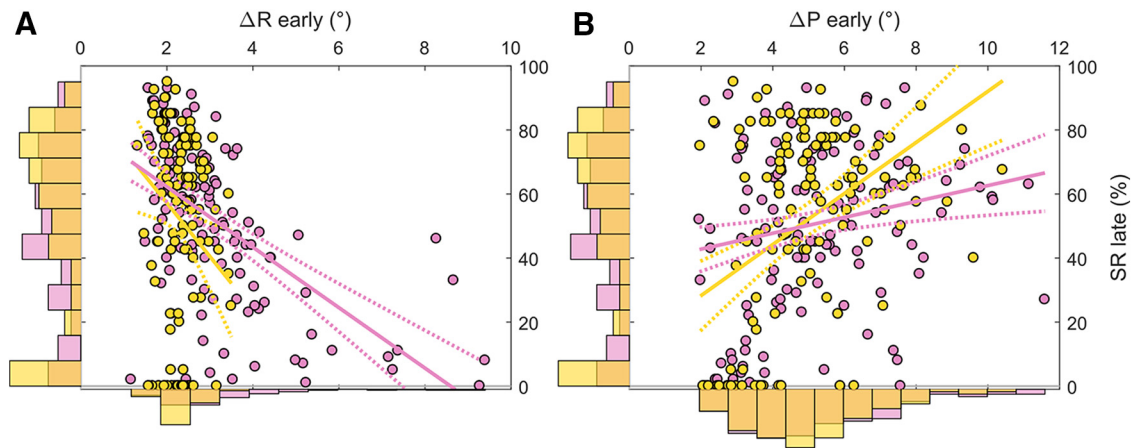


Figure 10. Slice plots showing regression results for prediction of late SR by changes in reach angle following rewarded (**A**) and unrewarded (**B**) trials during the early learning period. The central axis of each panel displays the individual data from the Acquire (yellow) and Preserve (pink) task. A histogram displaying the distribution of the data in each dimension is presented on the corresponding axis. Solid lines indicate the prediction of the regression model when the other predictor is held at its mean value.

Table 2. Regression results for split data for both the Acquire and Preserve tasks^a

	ΔR	ΔP	Model
Acquire			
β	−0.274	0.581	$F_{(115,2)} = 11.9$
SE	0.111	0.120	$p = 2.09 \times 10^{-5}$
p	0.015	3.89×10^{-6}	
Preserve			
β	−0.750	0.229	$F_{(112,2)} = 35.3$
SE	0.093	0.084	$p = 1.28 \times 10^{-12}$
p	1.07×10^{-12}	0.007	

^aOrdinary least squares linear regressions were performed with both ΔR and ΔP included as predictors. The regression coefficient, SE, and p value for each predictor are reported along with the significance of the comparison between the model and an intercept only model. In both tasks, there is an opposing relationship between ΔR and ΔP and success rate, with smaller changes after rewarded trials and larger changes after unrewarded trials predictive of success.

to an insignificant mediation effect, despite the median p value indicating an effect when considering all repetitions. We also examined an alternative possibility to the hypothesized model in which a relationship between SWM and ΔR was mediated by SR. We found that 31.20% of the total effect was mediated in the Acquire task using the hypothesized model, in contrast to only 0.17% in the alternative model. Similarly, in the Preserve task, the hypothesized model displayed a substantially larger mediation effect (44.77%) than the alternative model (5.02%). These results support the application of the hypothesized model.

Discussion

In this study, we sought to identify whether genetic background or specific domains of WMC could explain the variability observed in performance levels during reward-based motor learning tasks. We found that RWM and SWM predicted different aspects of the Acquire and Preserve tasks, whereas VWM only related to one performance measure (ΔP), but not consistently across tasks. Specifically, RWM predicted the explicit component of retention in the Acquire task and SR in the Preserve task, whereas SWM predicted SR in the Acquire task and the late period of the Preserve task. Furthermore, SWM negatively predicted ΔR in both tasks. Conversely, allelic variations of the three dopamine-related genes (DRD2, COMT, and DARPP32) did not consistently predict any behavioral variables across both tasks. This suggests that SWM predicts a participant’s capacity to reproduce a rewarded motor action, whereas RWM predicts a participant’s ability to express an explicit strategy when making large

behavioral adjustments. Therefore, we conclude that WMC plays a pivotal role in determining individual ability in reward-based motor learning.

Recently, Wong et al. (2019) described a positive relationship between SWM and the development of explicit strategies in visuomotor adaptation, complementing previous reports (Anguera et al., 2012; Christou et al., 2016; Vandevorde and Orban de Xivry, 2019). However, in contrast to the current findings, the previous experiments used relatively small sample sizes, which may render correlations unreliable. The large group sizes used here, and the confirmation of relationships across two tasks, provides strong evidence that these relationships are robust, replicable, and extend from visuomotor adaptation to reward-based motor learning. An interesting dichotomy was the reliance on SWM and RWM for the Acquire and Preserve task, respectively. Whereas the Preserve task required the maintenance of a large, abrupt behavioral change, the Acquire task required the gradual adjustment of behavior considering the outcomes of recent trials. Therefore, RWM may underscore one’s capacity to express a large correction consistently over trials with binary feedback, whereas SWM reflects one’s capacity to maintain a memory of previously rewarded actions and adjust behavior accordingly. Accordingly, McDougale and Taylor (2019) demonstrated that a mental rotation process is used in countering a visuomotor rotation, and Sidarta et al. (2018) reported that higher SWM is associated with reduced movement variability in a reward-based motor learning task. Here, the magnitude of ΔR was negatively related to SWM, but not RWM, in both tasks, suggesting that high SWM enables the maintenance of rewarding actions. Additionally, explicit retention, an element of the Acquire task requiring a large, sudden change in reach direction, was predicted by RWM rather than SWM. Notably, RWM and SWM were often selected as predictors simultaneously. The overlapping, but distinct, pattern of relationships between RWM, SWM, and outcome measures considered here supports the view that they share substrates but have different patterns of dependency on executive functions (Miyake et al., 2001).

A notable feature of the Preserve task is the “spiking” behavior observed immediately following “refresher” trials, suggesting a central role of refresher trials in binary feedback-based performance when included (Shmuelof et al., 2012; Codol et al., 2018). The transient nature of this decrease in error demonstrates this is insufficient to promote generalization to binary feedback trials, at

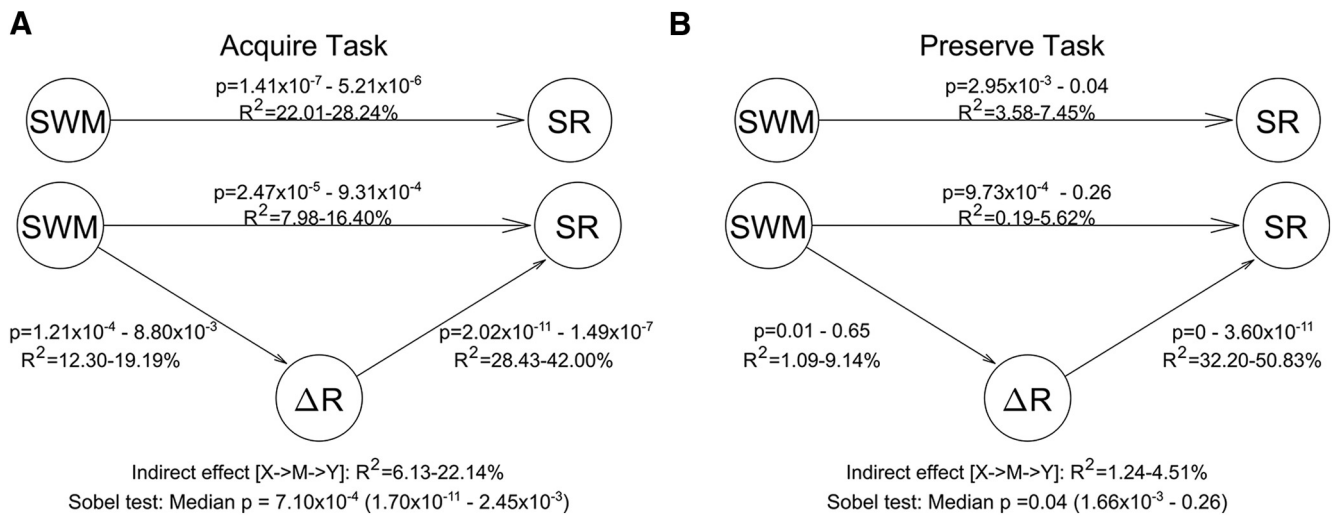


Figure 11. Mediation analysis for both the Acquire (**A**) and Preserve (**B**) tasks. The numbers associated with each arrow display the 95% CIs for each of the relationships (R^2 and p values) across the 1000 repetitions. Below the figure, the results of the Sobel test are displayed, indicating the amount of variance explained by the indirect pathway and the 95% CIs and median p value.

least in unsuccessful participants. It remains an open question whether superior performance of successful participants was partly due to a capacity to generalize information from “refresher” trials. McDougle and Taylor (2019) suggest that two separate strategies are used in visuomotor adaptation: response-caching and mental rotation. The balance between the two strategies is a function of task demands. The relationships between RWM and SWM to SR in the Preserve and Acquire tasks, respectively, may reflect a different balance of the use of these strategies. Visual feedback in “refresher” trials in the Preserve task may engage mental rotation processes, whereas the slow updating of behavior in the Acquire task engages the response-caching memory system. This would imply that response-caching is associated with SWM.

Surprisingly, although ΔP was a strong predictor of success in both tasks, it was not consistently predicted by any variable across both tasks. The lack of a consistent predictor of ΔP was unexpected given the importance of errors for the induction of structural learning in reinforcement learning (Sutton and Barto, 1998; Daw et al., 2011; Manley et al., 2014) and reward-based motor learning (Maxwell et al., 2001; Sidarta et al., 2018).

If RWM is important for explicit control and the main element predicting success in the Preserve task, it is worth considering whether gradual designs (as in the Acquire task) are more suitable to engage implicit reinforcement learning, at least initially. However, the Acquire task still bears a strong explicit component (Holland et al., 2018). How can these two views be reconciled? In reward-based motor learning tasks, it is observed that participants begin to reflect upon task structure and develop strategies upon encountering negative outcomes (Maxwell et al., 2001; Leow et al., 2016; Loonis et al., 2017), which occurs nearly immediately in the Preserve task after the introduction of binary feedback, due to a lack of generalization of cerebellar memory (Codol et al., 2018). In contrast, in the Acquire task, participants experience an early learning phase with mainly rewarding outcomes, possibly suppressing development of explicit control and allowing for this early window of implicit reward-based learning. Other studies have demonstrated that minor adjustments in reach direction under reward-based feedback can occur, although none has assessed their explicitness directly in the very early stages (Izawa and Shadmehr, 2011; Pekny et al., 2015; Ther-

rien et al., 2016). Notably, Izawa and Shadmehr (2011) observed that, after 8° shifts of a similarly sized reward region, participants indeed noticed the perturbation, but awareness was not assessed for smaller shifts.

In Holland et al. (2018), the addition of an RWM-like dual task was very effective in preventing explicit control, leading to participants invariably failing at the reaching task. Therefore, it may seem surprising that RWM does not predict SR in the Acquire task. A possible explanation is that RWM and SWM share the same memory buffer (Cohen et al., 1996; Jordan et al., 2001; Logie et al., 2005; Suchan et al., 2006; Anguera et al., 2010). Similarly, in force-field adaptation, the early component of adaptation, considered as bearing a strong explicit element, is selectively disrupted with a VWM dual task (Keisler and Shadmehr, 2010). However, we found no consistent relationship with VWM across our reward-based motor tasks. It may be possible that reward-based motor performance relies more on spatial instances of WM as opposed to tasks, such as force-field adaptation.

The absence of DA-related genetic relationships with behavior is a surprising result as a substantial body of literature points to a relationship between dopamine and performance in reward-based tasks, including those with motor components (Frank et al., 2007, 2009; Izawa and Shadmehr, 2011; Nakahara and Hikosaka, 2012; Deserno et al., 2015; Pekny et al., 2015; Doll et al., 2016; Therrien et al., 2016; Gershman and Schoenbaum, 2017). There is a growing appreciation of the links between decision-making and motor learning (Haith and Krakauer, 2013; Chen et al., 2017, 2018). However, the results presented here suggest that genetic predictors of exploration and exploitation in decision-making tasks are not also predictive of similar behaviors in reward-based motor learning.

Our sample sizes were defined *a priori* for 90% power based on previous work (see preregistrations) (Frank et al., 2009; Doll et al., 2016), and are unlikely to be underpowered. Another possibility is that we used the wrong variables to assess behavior. However, given the informative and coherent relationships between WM and motor learning, it could be that the SNPs we selected do not meaningfully relate to performance in reward-based motor tasks compared with WM. A similar claim was made in the decision-making literature (Collins and Frank, 2012). In line with this, a recent study showed that DA pharmacological ma-

nipulation did not alter reward effects in a visuomotor adaptation task (Quattrocchi et al., 2018). However, previous work has shown that Parkinson's disease patients show impaired reward-based motor performance (Pekny et al., 2015). It is possible that genetic variations may simply not impact reward-based motor learning significantly, especially compared with the wide depletion of dopaminergic neurons in Parkinson's disease. It is also important to note that, while we refer to both of our tasks as reward-based motor learning, they are both in essence visuomotor rotation paradigms. In the future, it is important to investigate whether these findings extend to more complex reward-based motor learning paradigms.

In conclusion, despite using two distinct tasks and an independent participant pool on different devices, we find strikingly similar results in reward-based motor learning. Whereas SWM strongly predicted a participant's capacity to reproduce successful motor actions, RWM predicted a participant's ability to express an explicit strategy when required to make large behavioral adjustments. Surprisingly, no dopamine-related genotypes predicted performance. Therefore, WMC plays a pivotal role in determining individual ability in reward-based motor learning. This could have important implications when using reward-based feedback in applied settings as only a subset of the population may benefit.

References

- Anguera JA, Reuter-Lorenz PA, Willingham DT, Seidler RD (2010) Contributions of spatial working memory to visuomotor learning. *J Cogn Neurosci* 22:1917–1930.
- Anguera JA, Bernard JA, Jaeggi SM, Buschkuhl M, Benson BL, Jennett S, Humfleet J, Reuter-Lorenz PA, Jonides J, Seidler RD (2012) The effects of working memory resource depletion and training on sensorimotor adaptation. *Behav Brain Res* 228:107–115.
- Auton A, Brooks LD, Durbin RM, Garrison EP, Kang HM, Korbel JO, Marchini JL, McCarthy S, McVean GA, Abecasis GR, 1000 Genomes Project Consortium (2015) A global reference for human genetic variation. *Nature* 526:68–74.
- Baron RM, Kenny DA (1986) The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. *J Pers Soc Psychol* 51:1173–1182.
- Bond KM, Taylor JA (2015) Flexible explicit but rigid implicit learning in a visuomotor adaptation task. *J Neurophysiol* 113:3836–3849.
- Boyd S (2010) Distributed optimization and statistical learning via the alternating direction method of multipliers. *FNT Machine Learning* 3:1–122.
- Brainard DH (1997) The psychophysics toolbox. *Spat Vis* 10:433–436.
- Buszard T, Masters RS (2018) Adapting, correcting and sequencing movements: does working-memory capacity play a role? *Int Rev Sports Exerc Psychol* 11:258–278.
- Cashback JG, McGregor HR, Mohatarem A, Gribble PL (2017) Dissociating error-based and reinforcement-based loss functions during sensorimotor learning. *PLoS Comput Biol* 13:e1005623.
- Cashback JG, Lao CK, Palidis DJ, Coltman SK, McGregor HR, Gribble PL (2019) The gradient of the reinforcement landscape influences sensorimotor learning. *PLoS Comput Biol* 15:e1006839.
- Chen X, Holland P, Galea JM (2018) The effects of reward and punishment on motor skill learning. *Curr Opin Behav Sci* 20:83–88.
- Chen X, Mohr K, Galea JM (2017) Predicting explorative motor learning using decision-making and motor noise. *PLoS Comput Biol* 13:e1005503.
- Christou AI, Miall RC, McNab F, Galea JM (2016) Individual differences in explicit and implicit visuomotor learning and working memory capacity. *Sci Rep* 6:36633.
- Codol O, Holland PJ, Galea JM (2018) The relationship between reinforcement and explicit control during visuomotor adaptation. *Sci Rep* 8:9121.
- Cohen MS, Kosslyn SM, Breiter HC, DiGirolamo GJ, Thompson WL, Anderson AK, Brookheimer SY, Rosen BR, Belliveau JW (1996) Changes in cortical activity during mental rotation: A mapping study using functional MRI. *Brain* 119:89–100.
- Collins AG, Frank MJ (2012) How much of reinforcement learning is working memory, not reinforcement learning? A behavioral, computational, and neurogenetic analysis: working memory in reinforcement learning. *Eur J Neurosci* 35:1024–1035.
- Daw ND, Niv Y, Dayan P (2005) Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control. *Nat Neurosci* 8:1704–1711.
- Daw ND, Gershman SJ, Seymour B, Dayan P, Dolan RJ (2011) Model-based influences on humans' choices and striatal prediction errors. *Neuron* 69:1204–1215.
- Deserno L, Huys QJ, Boehme R, Buchert R, Heinze HJ, Grace AA, Dolan RJ, Heinz A, Schlagenhauf F (2015) Ventral striatal dopamine reflects behavioral and neural signatures of model-based control during sequential decision making. *Proc Natl Acad Sci U S A* 112:1595–1600.
- Diedrichsen J, White O, Newman D, Lally N (2010) Use-dependent and error-based learning of motor behaviors. *J Neurosci* 30:5159–5166.
- Doll BB, Duncan KD, Simon DA, Shohamy D, Daw ND (2015) Model-based choices involve prospective neural activity. *Nat Neurosci* 18:767–772.
- Doll BB, Bath KG, Daw ND, Frank MJ (2016) Variability in dopamine genes dissociates model-based and model-free reinforcement learning. *J Neurosci* 36:1211–1222.
- Egan MF, Goldberg TE, Kolachana BS, Callicott JH, Mattay CM, Straub RE, Goldman D, Weinberger DR (2001) Effect of COMT Val108/158 met genotype on frontal lobe function and risk for schizophrenia. *Proc Natl Acad Sci U S A* 98:6917–6922.
- Frank MJ, Moustafa AA, Haughey HM, Curran T, Hutchison KE (2007) Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning. *Proc Natl Acad Sci U S A* 104:16311–16316.
- Frank MJ, Doll BB, Oas-Terpstra J, Moreno F (2009) Prefrontal and striatal dopaminergic genes predict individual differences in exploration and exploitation. *Nat Neurosci* 12:1062–1068.
- Gershman SJ, Schoenbaum G (2017) Rethinking dopamine prediction errors. *bioRxiv*. Available at <https://www.biorxiv.org/content/10.1101/239731v1>.
- Goldberg TE, Egan MF, Gscheidle T, Coppola R, Weickert T, Kolachana BS, Goldman D, Weinberger DR (2003) Executive subprocesses in working memory: relationship to catechol-O-methyltransferase Val158Met genotype and schizophrenia. *Arch Gen Psychiatry* 60:889–896.
- Haith AM, Krakauer JW (2013) Model-based and model-free mechanisms of human motor learning. In: *Progress in motor control* (Richardson MJ, Riley MA, Shockley K, eds), pp 1–21. New York: Springer.
- Hastie T, Tibshirani R, Wainwright M (2015) Statistical inference: the bootstrap. In: *Statistical learning with sparsity: the lasso and generalizations*. Boca Raton, FL: Chapman and Hall/CRC.
- Holland P, Codol O, Galea JM (2018) Contribution of explicit processes to reinforcement-based motor learning. *J Neurophysiol* 119:2241–2255.
- Huang VS, Haith A, Mazzoni P, Krakauer JW (2011) Rethinking motor learning and savings in adaptation paradigms: model-free memory for successful actions combines with internal models. *Neuron* 70:787–801.
- Izawa J, Shadmehr R (2011) Learning from sensory and reward prediction errors during motor adaptation. *PLoS Comput Biol* 7:e1002012.
- Jordan K, Heinze HJ, Lutz K, Kanowski M, Jäncke L (2001) Cortical activations during the mental rotation of different visual objects. *Neuroimage* 13:143–152.
- Keisler A, Shadmehr R (2010) A shared resource between declarative memory and motor memory. *J Neurosci* 30:14817–14823.
- Leow LA, de Rugy A, Marinovic W, Riek S, Carroll TJ (2016) Savings for visuomotor adaptation require prior history of error, not prior repetition of successful actions. *J Neurophysiol* 116:1603–1614.
- Logie RH, Della Sala S, Beschin N, Denis M (2005) Dissociating mental transformations and visuo-spatial storage in working memory: evidence from representational neglect. *Memory* 13:430–434.
- Loonis RF, Brincat SL, Antzoulatos EG, Miller EK (2017) A meta-analysis suggests different neural correlates for implicit and explicit learning. *Neuron* 96:521–534.e7.
- Mackinnon DP, Dwyer JH (1993) Estimating mediated effects in prevention studies. *Eval Rev* 17:144–158.
- Manley H, Dayan P, Diedrichsen J (2014) When money is not enough: awareness, success, and variability in motor learning. *PLoS One* 9:e86580.
- Maxwell JP, Masters RS, Kerr E, Weedon E (2001) The implicit benefit of learning without errors. *Q J Exp Psychol A* 54:1049–1068.

- McDougle SD, Taylor JA (2019) Dissociable cognitive strategies for sensorimotor learning. *Nat Commun* 10:40.
- McNab F, Klingberg T (2008) Prefrontal cortex and basal ganglia control access to working memory. *Nat Neurosci* 11:103–107.
- Miyake A, Friedman NP, Rettinger DA, Shah P, Hegarty M (2001) How are visuospatial working memory, executive functioning, and spatial abilities related? A latent-variable analysis. *J Exp Psychol Gen* 130:621–640.
- Modchalingam S, Vachon CM, Hart BM, Henriques DY (2019) The effects of awareness of the perturbation during motor adaptation on hand localization. *PLoS One* 14:20.
- Morehead JR, Qasim SE, Crossley MJ, Ivry R (2015) Savings upon re-aiming in visuomotor adaptation. *J Neurosci* 35:14386–14396.
- Morehead JR, Taylor JA, Parvin DE, Ivry RB (2017) Characteristics of implicit sensorimotor adaptation revealed by task-irrelevant clamped feedback. *J Cogn Neurosci* 29:1061–1074.
- Nakahara H, Hikosaka O (2012) Learning to represent reward structure: a key to adapting to complex environments. *Neurosci Res* 74:177–183.
- Pearson-Fuhrhop KM, Minton B, Acevedo D, Shahbaba B, Cramer SC (2013) Genetic variation in the human brain dopamine system influences motor learning and its modulation by L-dopa. *PLoS One* 8:e61197.
- Pekny SE, Izawa J, Shadmehr R (2015) Reward-dependent modulation of movement variability. *J Neurosci* 35:4015–4024.
- Peters M, Battista C (2008) Applications of mental rotation figures of the Shepard and Metzler type and description of a mental rotation stimulus library. *Brain Cogn* 66:260–264.
- Quattrocchi G, Monaco J, Ho A, Irmen F, Strube W, Ruge D, Bestmann S, Galea JM (2018) Pharmacological dopamine manipulation does not alter reward-based improvements in memory retention during a visuomotor adaptation task. *eNeuro* 5:ENEURO.0453–17.2018.
- Saijo N, Gomi H (2010) Multiple motor learning strategies in visuomotor rotation. *PLoS One* 5:e9399.
- Shepard RN, Metzler J (1971) Mental rotation of three-dimensional objects. *Science* 171:701–703.
- Shmuelof L, Huang VS, Haith AM, Delnicki RJ, Mazzoni P, Krakauer JW (2012) Overcoming motor “forgetting” through reinforcement of learned actions. *J Neurosci* 32:14617–14621.
- Sidarta A, van Vugt FT, Ostry DJ (2018) Somatosensory working memory in human reinforcement-based motor learning. *J Neurophysiol* 120:3275–3286.
- Sobel ME (1986) Some new results on indirect effects and their standard errors in covariance structure models. *Sociol Methodol* 16:159.
- Suchan B, Botko R, Gizewski E, Forsting M, Daum I (2006) Neural substrates of manipulation in visuospatial working memory. *Neuroscience* 139:351–357.
- Sutton RS, Barto A (1998) Reinforcement learning: an introduction. Cambridge, MA: Bradford.
- Taylor JA, Ivry RB (2011) Flexible cognitive strategies during motor learning. *PLoS Comput Biol* 7:e1001096.
- Taylor JA, Ivry RB (2014) Cerebellar and prefrontal cortex contributions to adaptation, strategies, and reinforcement learning. *Prog Brain Res* 210:217–253.
- Therrien AS, Wolpert DM, Bastian AJ (2016) Effective reinforcement learning following cerebellar damage requires a balance between exploration and motor noise. *Brain* 139:101–114.
- Therrien AS, Wolpert DM, Bastian AJ (2018) Increasing motor noise impairs reinforcement learning in healthy individuals. *eNeuro* 5:ENEURO.0050–18.2018.
- Tibshirani R (1996) Regression shrinkage and selection via the lasso. *J R Stat Soc B* 58:267–288.
- Tseng YW, Diedrichsen J, Krakauer JW, Shadmehr R, Bastian AJ (2007) Sensory prediction errors drive cerebellum-dependent adaptation of reaching. *J Neurophysiol* 98:54–62.
- van der Kooij K, Overvliet KE (2016) Rewarding imperfect motor performance reduces adaptive changes. *Exp Brain Res* 234:1441–1450.
- Vandevoorde K, Orban de Xivry JJ (2019) Internal model recalibration does not deteriorate with age while motor adaptation does. *Neurobiol Aging* 80:138–153.
- Vindras P, Desmurget M, Prablanc C, Viviani P (1998) Pointing errors reflect biases in the perception of the initial hand position. *J Neurophysiol* 79:3290–3294.
- Werner S, van Aken BC, Hulst T, Frens MA, van der Geest JN, Strüder HK, Donchin O (2015) Awareness of sensorimotor adaptation to visual rotations of different size. *PLoS One* 10:e0123321.
- Wong AL, Marvel CL, Taylor JA, Krakauer JW (2019) Can patients with cerebellar disease switch learning mechanisms to reduce their adaptation deficits? *Brain* 142:662–673.
- Yuan M, Lin Y (2006) Model selection and estimation in regression with grouped variables. *J R Stat Soc B* 68:49–67.