

Domain-specific working memory, but not dopamine-related genetic variability, shapes reward-based motor learning

Holland, Peter; Codol, Olivier; Oxley, Elizabeth; Taylor, Madison ; Hamshere, Elizabeth; Joseph, Shadiq; Huffer, Laura; Galea, Joseph

DOI:

[10.1101/524900](https://doi.org/10.1101/524900)

[10.1523/JNEUROSCI.0583-19.2019](https://doi.org/10.1523/JNEUROSCI.0583-19.2019)

License:

None: All rights reserved

Document Version

Peer reviewed version

Citation for published version (Harvard):

Holland, P, Codol, O, Oxley, E, Taylor, M, Hamshere, E, Joseph, S, Huffer, L & Galea, J 2019, 'Domain-specific working memory, but not dopamine-related genetic variability, shapes reward-based motor learning: Mechanisms of reward-based motor learning', *The Journal of Neuroscience*, vol. 39, no. 47, pp. 9383-9396. <https://doi.org/10.1101/524900>, <https://doi.org/10.1523/JNEUROSCI.0583-19.2019>

[Link to publication on Research at Birmingham portal](#)

General rights

Unless a licence is specified above, all rights (including copyright and moral rights) in this document are retained by the authors and/or the copyright holders. The express permission of the copyright holder must be obtained for any use of this material other than for purposes permitted by law.

- Users may freely distribute the URL that is used to identify this publication.
- Users may download and/or print one copy of the publication from the University of Birmingham research portal for the purpose of private study or non-commercial research.
- User may use extracts from the document in line with the concept of 'fair dealing' under the Copyright, Designs and Patents Act 1988 (?)
- Users may not further distribute the material nor use it for the purposes of commercial gain.

Where a licence is displayed above, please note the terms and conditions of the licence govern your use of this document.

When citing, please reference the published version.

Take down policy

While the University of Birmingham exercises care and attention in making items available there are rare occasions when an item has been uploaded in error or has been deemed to be commercially or otherwise sensitive.

If you believe that this is the case for this document, please contact UBIRA@lists.bham.ac.uk providing details and we will remove access to the work immediately and investigate.

1 **Domain-specific working memory, but not dopamine-related**
2 **genetic variability, shapes reward-based motor learning**

3
4 Running title: Mechanisms of reward-based motor learning

5
6 **Peter Holland*¹, Olivier Codol*¹, Elizabeth Oxley¹, Madison Taylor¹, Elizabeth**
7 **Hamshere¹, Shadiq Joseph¹, Laura Huffer¹, Joseph M. Galea¹**

8
9 ¹*School of Psychology and Centre for Human Brain Health, University of Birmingham,*
10 *Edgbaston, Birmingham, B15 2TT, UK*

11
12 * = These authors provided equal contribution to this work

13
14 Number of Figures: 11

15 Number of Tables: 2

16 Number of Pages: 47

17 Words: Abstract = 249 Introduction = 650 Discussion = 1494

18 Keywords: Explicit control, Motor adaptation, Reinforcement, Retention

19 Conflict of Interest: The authors declare no competing financial interests

20 Acknowledgements: This work was supported by the European Research Council starting
21 grant: MotMotLearn (637488)

22 Correspondence should be addressed to:

23 Peter Holland

24 Email: P.J.Holland@bham.ac.uk

25 **Abstract**

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

46

47 **Significance statement**

48 Reward-based motor learning tasks have repeatedly been shown to lead to idiosyncratic
49 behaviours that cause varying degrees of task success. Yet, the factors determining an

50 individual's capacity to use reward-based feedback are unclear. Here, we assessed a wide range
51 of possible candidate predictors, and demonstrate that domain-specific working memory plays
52 an essential role in determining individual capacity to use reward-based feedback. Surprisingly,
53 genetic variations in dopamine availability were not found to play a role. This is in stark
54 contrast with seminal work in the reinforcement and decision-making literature, which show
55 strong and replicated effects of the same dopaminergic genes in decision-making. Therefore,
56 our results provide novel insights into reward-based motor learning, highlighting a key role for
57 domain-specific working memory capacity.

58

59 **Introduction**

60 When performing motor tasks under altered environmental conditions, adaptation to the new
61 constraints occurs through the recruitment of a variety of systems (Taylor and Ivry, 2014).
62 Arguably the most studied of those systems is cerebellum-dependent adaptation, which consists
63 of the implicit and automatic recalibration of mappings between actual and expected outcomes
64 through sensory prediction errors (Morehead et al., 2017; Tseng et al., 2007). Besides
65 cerebellar adaptation, other work has demonstrated the involvement of a cognitive, deliberative
66 process whereby motor plans are adjusted based on structural understanding of the task (Bond
67 and Taylor, 2015; Taylor and Ivry, 2011). We label this process 'explicit control' (Codol et al.,
68 2018; Holland et al., 2018), although it has also been referred to as strategy (Taylor and Ivry,
69 2011) or explicit re-aiming (Morehead et al., 2015). Recently it has been proposed that
70 reinforcement learning, whereby the memory of successful or unsuccessful actions are
71 strengthened or weakened, respectively, may also play a role (Huang et al., 2011; Izawa and
72 Shadmehr, 2011; Shmuelof et al., 2012). Such reward-based reinforcement has been assumed
73 to be an implicit and automatic process (Haith and Krakauer, 2013). However, recent evidence
74 suggests that phenomena attributed to reinforcement-based learning during visuomotor rotation

75 tasks can largely be explained through explicit processes (Codol et al., 2018; Holland et al.,
76 2018).

77 One outstanding feature of reinforcement-based motor learning is the great variability
78 expressed across individuals (Codol et al., 2018; Holland et al., 2018; Therrien et al., 2016,
79 2018). What factors underlie such variability is unclear. If reinforcement is explicitly grounded,
80 it could be argued that individual working memory capacity (WMC), which is reliably related
81 to the propensity to employ explicit control in classical motor adaptation tasks (Anguera et al.,
82 2010, 2012; Christou et al., 2016; Holland et al., 2018; Sidarta et al., 2018), would also predict
83 performance in reinforcement-based motor learning. Anguera et al. (2010) demonstrated that
84 mental rotation WMC (RWM), unlike other forms of working memory such as verbal WMC
85 (VWM), correlates with explicit control. Recently, Christou et al. (2016) reported similar
86 results with spatial WMC (SWM). If this extends to reward based motor learning, this would
87 strengthen the proposal that it bears a strong explicit component.

88 Another potential contributor to this variability is genetic profile. In previous work (Codol et
89 al., 2018; Holland et al., 2018), we argue that reinforcement-based motor learning performance
90 relies on a balance between exploration and exploitation of the task space, a feature reminiscent
91 of structural learning and reinforcement-based decision-making (Daw et al., 2005; Frank et al.,
92 2009; Sutton and Barto, 1998). A series of studies from Frank and colleagues suggests that
93 individual tendencies to express explorative/exploitative behaviour can be predicted based on
94 dopamine-related genetic profile (Doll et al., 2016; Frank et al., 2007, 2009). Reinforcement
95 has consistently been linked to dopaminergic function in a variety of paradigms, and thus, such
96 a relationship could also be expected in reward-based motor learning (Pekny et al., 2015).
97 Specifically, Frank and colleagues focused on Catecholamine-O-Methyl-Transferase (COMT),
98 Dopamine- and cAMP-Regulated neuronal Phosphoprotein (DARPP32) and Dopamine

99 Receptor D2 (DRD2), and suggest a distinction between COMT-modulated exploration and
100 DARPP32- and DRD2-modulated exploitation (Frank et al., 2009).
101 Consequently, we investigated the influence of WMC (RWM, SWM, and VWM) and genetic
102 variations in dopamine metabolism (DRD2, DARPP32, and COMT) on individuals' ability to
103 perform reward-based motor learning. We examined this using two established reward-based
104 motor learning tasks. First, a task analogous to a gradually introduced rotation (Holland et al.,
105 2018) required participants to learn to adjust the angle at which they reached to a slowly and
106 secretly shifting reward region (Acquire); second, a task with an abruptly introduced rotation
107 (Codol et al., 2018; Shmuelof et al., 2012) required participants to preserve performance with
108 reward-based feedback after adapting to a visuomotor rotation (Preserve). The use of these two
109 tasks enabled us to examine whether similar predictors of performance explained participant's
110 capacity to acquire and preserve behaviour with reward-based feedback.

111

112 **Methods**

113 Prior to the start of data collection, the sample size, variables of interest and analysis method
114 were pre-registered. The pre-registered information, data and analysis code can be found online
115 at <https://osf.io/j5v2s/> and <https://osf.io/rmwc2/> for the Preserve and Acquire tasks,
116 respectively.

117

118 **Participants**

119 121 (30 male, mean age: 21.06, range: 18-32) and 120 (16 male, mean age: 19.24, range: 18-
120 32) participants were recruited for the Acquire and Preserve tasks, respectively. All participants
121 provided informed consent and were remunerated with either course credit or money
122 (£7.50/hour). All participants were free of psychological, cognitive, motor or uncorrected

123 visual impairment. The study was approved by and performed in accordance with the local
124 research ethics committee of the University of Birmingham, UK.

125

126 **Experimental design**

127 Participants were seated before a horizontally fixed mirror reflecting a screen placed above, on
128 which visual stimuli were presented. This arrangement resulted in the stimuli appearing at the
129 level on which participants performed their reaching movements. The Acquire (gradual) and
130 Preserve (abrupt) tasks were performed on two different stations, with a KINARM (BKIN
131 Technology, London, Ontario; sampling rate: 1000Hz) and a Polhemus 3SPACE Fastrak
132 tracking device (Colchester, Vermont; sampling rate: 120Hz), employed respectively. The
133 Acquire task was run using Simulink (The Mathworks, Natwick, MA) and Dexterit-E (BKIN
134 Technology), while the Preserve task was run using Matlab (The Mathworks, Natwick, MA)
135 and Psychophysics toolbox (Brainard, 1997). The Acquire task employed the same paradigm
136 and equipment as in Holland et al. (2018), with the exception of the maximum reaction time
137 (RT), which was increased from 0.6s to 1s, and the maximum movement time, which was
138 reduced from 1s to 0.6s. The Preserve task used the same setup and display as in Codol et al.
139 (2018); however, the number of ‘refresher’ trials during the binary feedback (BF) blocks was
140 increased from one to two in every 10 trials. The designs were kept as close as possible to their
141 respective original publications to promote replication and comparability across studies. In
142 both tasks reaching movements were made with the dominant arm. Both the Acquire and
143 Preserve tasks have previously been examined in isolation from each other (Acquire Task:
144 Cashaback et al., 2017, 2019; Holland et al., 2018; Therrien et al., 2016, 2018; Preserve: Codol
145 et al., 2018; Shmuelof et al., 2012) and we maintain this distinction here. However, it should
146 be noted that the two tasks are essentially visuomotor rotation tasks. One of the aims of this
147 study was to determine if similar mechanisms underly the use of binary feedback in both the

148 learning of a gradual rotation and maintenance of a previously learnt abrupt rotation. Therefore,
149 despite the similarities we analyse the results of each task in isolation in addition to comparing
150 the results across tasks.

151

152 **Reaching tasks**

153 *Acquire task.* Participants performed 670 trials, each of which followed a stereotyped timeline.
154 The starting position for each trial was in a consistent position roughly 30cm in front of the
155 midline and was indicated by a red circle (1cm radius). After holding the position of the handle
156 within the starting position, a target (red circle, 1cm radius) appeared directly in front of the
157 starting position at a distance of 10cm. Participants were instructed to make a rapid ‘shooting’
158 movement that passed through the target. If the cursor position at a radial distance of 10cm was
159 within a reward region ($\pm 5.67^\circ$, initially centred on the visible target; grey region in Figure 1a)
160 the target changed colour from red to green and a green tick was displayed just above the target
161 position, informing participants of the success of their movement. However, if the cursor did
162 not pass through the reward region, the target disappeared from view and no tick was displayed,
163 signalling failure (binary feedback). After each movement, the robot returned to the starting
164 position and participants were instructed to passively allow this.

165 For the first 10 trials, the position of the robotic handle was displayed as a white cursor (0.5
166 cm radius) on screen. Following this practice block, the cursor was extinguished for the
167 remainder of the experiment and participants only received binary feedback. The baseline block
168 consisted of the first 40 trials under binary feedback. During this period the reward region
169 remained centred on the visible target. Subsequently, unbeknownst to the participant, the
170 reward region rotated in steps of 1° every 20 trials; the direction of rotation was
171 counterbalanced across participants. After reaching a rotation of 25° , the reward region was
172 held constant for an additional 20 trials. Performance during these last 20 trials was used to

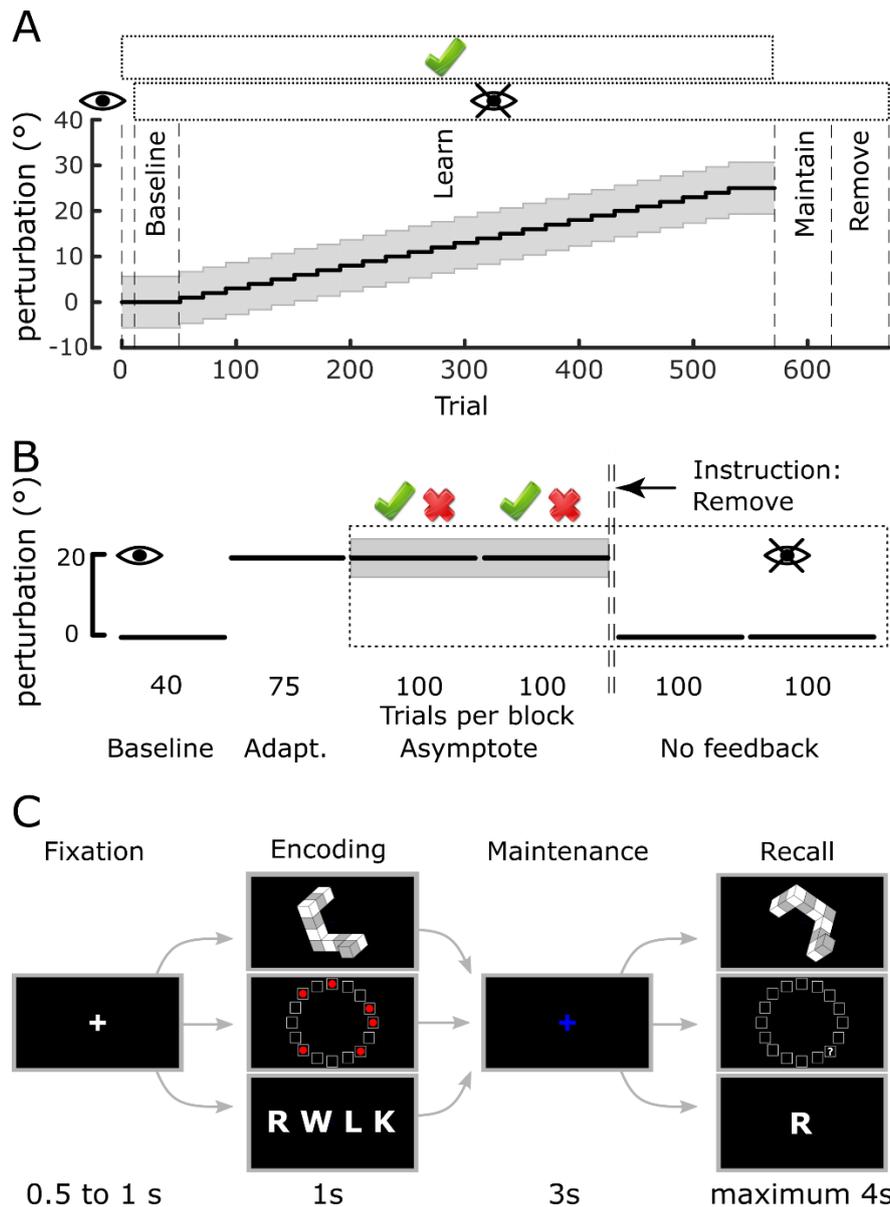
173 determine overall task success. Subsequently, binary feedback was removed, and participants
174 were instructed to continue reaching as they were (maintain block) for the following 50 trials.
175 Following this, participants were then informed that the reward region shifted during the
176 experiment but not of the magnitude or the direction of the shift. They were then instructed to
177 return to reaching in the same manner as they were at the start of the experiment (remove block,
178 50 trials). During the learning phase of the task participants were given a 1-minute rest after
179 trials 190 and 340.

180 *Preserve task.* Participants performed 515 trials in total. On each trial participants were
181 instructed to make a rapid ‘shooting’ movement that passed through a target (white circle,
182 radius: 0.125cm) visible on the screen. The starting position for each trial was indicated by a
183 white square (width: 1cm) roughly 30cm in front of the midline and the target was located at
184 angle of 45° from the perpendicular in a counter clockwise direction at a distance of 8cm. The
185 position of the tracking device attached to the fingertip was displayed as a cursor (green circle,
186 radius: 0.125cm). When the radial distance of the cursor from the starting position exceeded
187 8cm, the cursor feedback disappeared, and the end position was displayed instead.

188 First, participants performed a baseline period of 40 trials, during which the position of the
189 cursor was visible, and the cursor accurately reflected the position of the fingertip. In the
190 adaptation block (75 trials), participants were exposed to an abruptly introduced 20° clockwise
191 visuomotor rotation of the cursor feedback (Figure 1b). Subsequently, all visual feedback of
192 the cursor was removed, and participants received only binary feedback. If the end position of
193 the movement fell within a reward region, the trial was considered successful and a tick was
194 displayed; otherwise a cross was displayed. The reward region was centred at a clockwise
195 rotation of 20° with respect to the visual target with a width of 10° , that is, it was centred on
196 the direction that successfully accounted for the previously experienced visuomotor rotation.
197 Binary feedback was provided for 200 trials divided into 2 blocks of 100 trials (asymptote

198 blocks). Furthermore, participants experienced 2 ‘refresher’ trials for every 10 trials, where
199 rotated visual feedback of the cursor position was again accessible (Codol et al., 2018;
200 Shmuelof et al., 2012). This represents an increase compared to Codol et al. (2018) because
201 participants in this study tended to have poorer performance under binary feedback, possibly
202 due to a fatigue effect following the WM tasks (Anguera et al., 2012; see discussion). Finally,
203 two blocks (100 trials each) with no performance feedback were employed in order to assess
204 retention of the perturbation (no-feedback blocks). Before the first of those two blocks,
205 participants were informed of the visuomotor rotation, asked to stop accounting for it through
206 aiming off target and to aim straight at the target.

207



208

209

210 **Figure 1: Experimental design.** **A:** Time course of the Acquire task with the different
 211 experimental periods labelled. The grey region represents the reward region, which gradually
 212 rotated away from the visual target after the initial baseline period. The rectangle enclosing the
 213 green tick above the axes represents trials in which reward was available, and the rectangle
 214 with the ‘eye’ symbol indicates when vision was not available. **B:** Time course of the Preserve
 215 task. After adapting to an initial rotation with vision available, vision was removed (eye symbol)
 216 and reward-based feedback was introduced (tick and cross above the axes). Prior to the no-
 217 feedback blocks participants were instructed to remove any strategy they had been using. **C:**
 218 WMC tasks, the three tasks followed a stereotyped timeline with only the items to be
 219 remembered differing. Each trial consisted of 4 phases (Fixation, Encoding, Maintenance, and
 220 Recall) with the time allocated to each displayed below.

221

222

223 **Working memory tasks**

224 Participants performed three WM tasks, all of which followed the same design with the
225 exception of the nature of the items to be remembered (Figure 1c). All WM tasks were run
226 using Matlab (The Mathworks, Natwick, MA) and Psychophysics toolbox (Brainard, 1997).
227 At the start of each trial, a white fixation cross was displayed in the centre of the screen for a
228 period of 0.5 to 1s randomly generated from a uniform distribution (fixation period in Figure
229 1c). In the encoding period, the stimuli to be remembered was displayed for 1s and then
230 subsequently replaced with a blue fixation cross for the maintenance period which persisted
231 for 3s. Finally, during the recall period, participants were given a maximum of 4s to respond
232 by pressing one of three keys on a keyboard with their dominant hand. The '1' key indicated
233 that the stimuli presented in the recall period was a 'match' to that presented in the encoding
234 period, the '2' key indicated a 'non-match', and '3' indicated that the participant was unsure
235 as to the correct answer. Each WM task contained 5 levels of difficulty with the 12 trials
236 presented for each; 6 of which were trials in which 'match' was the correct answer and 6 in
237 which 'non-match' was the correct answer. Consequently, each WM task consisted of 60 trials
238 and the order in which the tasks were performed was pseudorandomised across participants.
239 Prior to the start of each task participants performed 10 practice trials to familiarise themselves
240 with the task and instructions. For both the Acquire and Preserve tasks, the WM tasks were
241 performed in the same experimental session as the reaching. However, in the case of the
242 Acquire task the WM tasks were performed after the reaching task whereas for the Preserve
243 task the WM tasks were performed first.

244 In the RWM task (Figure 1c, top row), the stimuli consisted of six 2D representations of 3D
245 shapes drawn from an electronic library of the Shepard and Metzler type stimuli (Peters and
246 Battista, 2008). The shape presented in the recall period was always the same 3D shape
247 presented in the encoding period after undergoing a screen-plane rotation of 60°, 120°, 180°, 247

248 240° or 300°. In ‘match’ trials, the only transform applied was the rotation; however, in ‘non-
249 match’ trials an additional vertical-axis mirroring was also applied. The difficulty of mental
250 rotation has been demonstrated to increase with larger angles of rotation (Shepard and Metzler,
251 1971) and therefore the different degrees of rotation corresponded to the 5 levels of difficulty.
252 However, given the symmetry of two pairs of rotations (60 and 300, 120 and 240), these 5
253 levels were collapsed to 3 for analysis.

254 In the SWM task (Figure 1c, middle row), stimuli in the encoding period consisted of a variable
255 number of red circles placed within 16 squares arranged in a circular array (McNab and
256 Klingberg, 2008). In the recall period, the array of squares was presented without the red circles
257 and instead a question mark appeared in one of the squares. Participants then answered to the
258 question ‘*Was there a red dot in the square marked by a question mark?*’ by pressing a
259 corresponding key. In ‘match’ trials the question mark appeared in one of the squares
260 previously containing a red circle and in ‘non-match’ trials it appeared in a square that was
261 previously empty. Difficulty was scaled by varying the number of red circles (i.e. the number
262 of locations to remember) from 3 to 7.

263 In the VWM task (Figure 1c, bottom row), participants were presented with a list of a variable
264 number of consonants during the encoding period. In the recall period a single consonant was
265 presented, and participants answered to the question ‘*Was this letter included in the previous*
266 *array?*’. Thus, the letter could either be drawn from the previous list (‘match’ trials) or have
267 been absent from the previous list (‘non-match’ trials). Difficulty in this task was determined
268 by the length of the list to be remembered, ranging from 5 to 9.

269 Both the SWM and RWM tasks have been suggested to fall under the general umbrella term
270 of spatial ability (Buszard and Masters, 2018). However, Miyake et al. (2001) suggest that
271 although both RWM and short term storage of spatial information (i.e. SWM) are within the
272 spatial domain, RWM appears to rely more heavily on executive function and SWM on basic

273 short term storage of spatial information. Furthermore, previous studies have found
274 relationships between motor learning and this SWM task (Christou et al., 2016; Vandevoorde
275 and Orban de Xivry, 2019) and tasks similar to our RWM task (Anguera et al., 2010). Therefore,
276 we included both tasks to investigate if there was any severability in their relationships with
277 reaching performance and leveraged our use of two separate reaching tasks and large cohorts
278 to probe if this was due to specific task parameters.

279

280 **Genetic sample collection and profiling**

281 COMT is thought to affect DA function mainly in the prefrontal cortex (Egan et al., 2001;
282 Goldberg et al., 2003), a region known for its involvement in WM and strategic planning
283 (Anguera et al., 2010; Doll et al., 2015), whereas DARPP32 and DRD2 act mainly in the basal
284 ganglia to promote exploitative behaviour, possibly by promoting selection of the action to be
285 performed (Frank et al., 2009). Consequently, we focused here on single-nucleotide
286 polymorphisms (SNP) related to those genes: RS4680 (COMT) and RS907094 (DARPP32).
287 Regarding DRD2, there are two potential SNPs available, RS6277 and RS1800497. Although
288 previous studies focusing on exploration and exploitation have assessed RS6277 expression
289 (Doll et al., 2016; Frank et al., 2007, 2009), it should be noted that this SNP varies greatly
290 across ethnic groups, with some allelic variations being nearly completely absent in non-
291 Caucasian-European groups (e.g. see RS6277 in 1000 Genomes Project (The 1000 Genomes
292 Project Consortium et al., [2015])). This has likely been inconsequential in previous work, as
293 Caucasian-European individual represented the majority of sampled groups; here however, this
294 represents a critical shortcoming, as we aim at investigating a larger and more representative
295 population including other ethnic groups. Consequently, we based our analysis on the
296 RS1800497 allele of the DRD2 gene (Pearson-Fuhrhop et al., 2013).

297 At the end of the task, participants were asked to produce a saliva sample which was collected,
298 stabilized and transported using Oragene.DNA saliva collection kits (OG-500, DNAGENOTEK,
299 Ontario, Canada). Participants were requested not to eat or drink anything except water for at
300 least two hours before sample collection. Once data collection was completed across all
301 participants, the saliva samples were sent to LGC (Hoddeson, Hertfordshire;
302 <https://www.lgcgroup.com/>) for DNA extraction (per Oragene protocols:
303 <https://www.dnagenotek.com/>) and genotyping. SNP genotyping was performed using the
304 KASP SNP genotyping system. KASP is a competitive allele-specific PCR incorporating a
305 FRET quencher cassette. Specifically, the SNP-specific KASP assay mix (containing two
306 different, allele specific, competing forward primers) and the universal KASP master mix
307 (containing FRET cassette plus Taq polymerase in an optimised buffer solution) were added to
308 DNA samples and a thermal cycling reaction performed, followed by an end-point fluorescent
309 read according to the manufacturer's protocol. All assays were tested on in-house validation
310 DNA prior to being run on project samples. No-template controls were used, and 5% of the
311 samples had duplicates included on each plate to enable the detection of contamination or non-
312 specific amplification. All assays had over 90% call rates. Following completion of the PCR,
313 all genotyping reaction plates were read on a BMG PHERAStar plate reader. The plates were
314 recycled until a laboratory operator was satisfied that the PCR reaction had reached its endpoint.
315 In-house Kraken software then automatically called the genotypes for each sample, with these
316 results being confirmed independently by two laboratory operators. Furthermore, the duplicate
317 saliva samples collected from 5% of participants were checked for consistency with the primary
318 sample. No discrepancies between primary samples and duplicates were discovered.

319

320

321

322 **Data analysis**

323 *Acquire task:* Reach trials containing movement times over 0.6s or less than 0.2s were removed
324 from analysis (6.9% of trials). The end point angle of each movement was defined at the time
325 when the radial distance of the cursor exceeded 10cm. This angle was defined in relation to the
326 visible target with positive angles indicating clockwise rotations. End point angles and target
327 angles for participants who experienced the counter clockwise rotations were sign-transformed.
328 The explicit component of retention was defined as the difference between the mean reach
329 angle of the maintain block and the remove block, while the implicit component was the
330 difference between the mean reach angle of the remove block and baseline (Werner et al., 2015).
331 Participants that achieved a mean reach angle within the reward region during the final 20 trials
332 before the maintain block were considered '*successful*' in learning the rotation; otherwise they
333 were considered '*unsuccessful*'. As in Holland et al. (2018), for unsuccessful participants, the
334 largest angle of rotation at which the mean reach angle fell within the reward region was taken
335 as the end of successful performance, and only trials prior to this point were included for further
336 analysis. Success rate was defined as the percentage of trials during the learning blocks in
337 which the end point angle was within the reward region. In order to examine the effect of
338 reward on the change in end point angle on the subsequent trial, we examined the magnitude
339 and variability of changes in end point angle between consecutive trials (Holland et al., 2018;
340 Sidarta et al., 2018; Therrien et al., 2016, 2018). To calculate the median absolute change
341 following rewarded (ΔR) and unrewarded (ΔP) trials we extracted the changes in reach angle
342 following each trial type and calculated the median of the absolute values of these changes for
343 each participant. These measures therefore represent the median of the magnitude of changes
344 in reach angle, regardless of direction. Furthermore, in order to examine the variability of trial-
345 by-trial adjustments ($MAD[\Delta R]$ and $MAD[\Delta P]$ for rewarded and unrewarded trials,
346 respectively) we calculated the median absolute deviation of the changes in reach angle. It is

347 important to note that ΔR and ΔP are calculated from the absolute magnitude of the changes in
348 reach angle, whereas, $MAD[\Delta R]$ and $MAD[\Delta P]$ are calculated from the non-absolute values
349 (including the direction of change).

350 *Preserve task:* Reach trials containing movement times over 1s were removed from analysis
351 (2.38% of trials). The end point angle for each movement was defined at the time that the radial
352 distance of the cursor from the start position exceeded 8cm. Trials in which the error was
353 greater than 80° were excluded from further analysis (0.94% of trials). As in Codol et al. (2018),
354 learning rate was calculated by fitting an exponential function to the angular error between
355 cursor and target for trials in the adaptation block, with the β value taken as the learning rate
356 (mean $R^2=0.34\pm 0.15$). The β estimates attained from all fits were first sign transformed and
357 then log-transformed to counteract skewness prior to entering the regression analysis. Using
358 this method, a value close to 0 indicated faster learning, whereas more negative values
359 indicated slower learning. Similar to Codol et al (2018), success rate, corresponding to
360 percentage of rewarded trials, was measured separately in the first 30 and last 170 trials of the
361 asymptote blocks and labelled early and late success rate, respectively. This reflects a
362 dichotomy between a dominantly exploration-driven early phase and a later exploitation-driven
363 phase. The analysis of changes in reach angle (ΔR and ΔP) was confined to the last 170 trials
364 of the asymptote blocks. Implicit retention was defined as the difference between the average
365 baseline reach direction and the mean reach direction of the last 20 trials of the last no-feedback
366 block (Codol et al., 2018). Analysis of changes in reach angle following rewarded trials were
367 not pre-registered but were included *post-hoc*.

368 *Exploratory analysis of reaching data:* In order to understand which outcome variables in the
369 reaching tasks were predictive of overall task success, we split the learning period into two
370 sections for every participant. We assessed trial-by-trial changes in end point angle in the first
371 section and compared them to success rate in the second section. For the Acquire task, we

372 assessed trial-by-trial adjustments during the learning block, excluding the final 20 trials, and
373 compared them to success rate in the last 20 trials of the learning block. In the Preserve task,
374 we measured adjustments in the first 100 trials of the asymptote blocks and compared them to
375 success rate in the last 100 trials of the asymptote blocks.

376 *WM tasks:* WM performance was defined as the average percentage of correct responses across
377 the 3 highest levels of difficulty for each task. In the case of the RWM task, the symmetrical
378 arrangement of the angles of rotation in effect produced three levels of difficulty and therefore
379 all trials were analysed.

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

388

389

390

391

392

393

394

395

396

397

398

399

400

401

402

403

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

404
405
406
407
408

Table 1: Coding for SNPs. The name of the SNP is provided along with the code assigned to each allele. The numbers represent the counts for the specific allele in the two tasks (Preserve, Acquire).

409 **Statistical analysis**

410 Regressions were performed using the linear Lasso method (Tibshirani, 1996; *lasso* function
411 in MatLab’s *Statistics and Machine Learning* Toolbox). Lasso regression employs a shrinkage
412 method that allows for some predictors to be shrunk to a value of 0, effectively removing them
413 from the regression model. Therefore, the method acts as a selection method for predictors in
414 an analogous way to stepwise regression. We used a 10-fold cross validation approach to
415 calculate the Mean Squared Error (MSE) over a range of values of a penalty term λ . Specifically,
416 as λ increases, the shrinkage of predictor values increases. For $\lambda=0$, the model reduced to a
417 standard linear regression, as all predictors were included without any shrinkage. Cross
418 validation protects against the problem of over-fitting by calculating the MSE on data ‘unseen’
419 by the model during fitting. For any given outcome variable, if its $MSE(\lambda)$ function exhibited
420 a minimum value within its defined boundaries, the model associated with that minimum value
421 was considered selected. If no minimum was observed, this signified that an empty model was
422 a better fit than any other possible model. If such minimum was detected in the $MSE(\lambda)$
423 function, the β estimates from that model (i.e. at that value of λ) were taken. We repeated this

424 procedure 1000 times to obtain the distribution of the true β from the estimates (Hastie et al.,
425 2015). In order for a potential variable to be considered a selected predictor, that predictor
426 should be selected (i.e. $\beta \neq 0$) in at least 80% of the repetitions. The threshold of 80% was chosen
427 as to maintain sufficient sensitivity whilst still returning relatively sparse models. We report
428 the median β estimate in the text for all selected predictors.

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

439 Prior to the regression analysis, all predictors and predicted variables were standardised (z-
440 scored). For all non-ordinal variables, individual data were considered outliers if further than
441 3 standard deviations from the mean and were removed prior to standardisation.
442 Multicollinearity of predictors was also assessed before regression with Belsley Collinearity
443 Diagnostics (*collintest* function in MatLab's *Econometrics Toolbox*) and no predictors were
444 found to exhibit condition indexes over 30, indicating acceptable levels of collinearity. When
445 considering retention for both tasks, unsuccessful participants were removed from the
446 regression analysis. We further characterised the relationships between predictor variables by
447 combining the data for the two tasks for the working memory (WM) tasks and the genetic codes

448 (N=241). We analysed relationships between the WM tasks with correlations and between
449 genetics and WM tasks with one-way ANOVAs.

450 *Exploratory mediation analysis:* We performed a mediation analysis to test if the relationship
451 between SWM and SR was mediated by ΔR . Our hypothesis was that higher SWM enables
452 smaller changes after correct trials (ΔR) and this then explains the relationship between SWM
453 and SR. To ensure that separate trials were used in the calculation of ΔR and SR, we split the
454 trials into two equally sized folds. The SR was then calculated for one-fold as a percentage of
455 correct trials, and ΔR was calculated as the median absolute change of reach angle after correct
456 trials in the other fold. For the Acquire task only successful subjects were included in the
457 mediation analysis. We employed Baron & Kenny's three step mediation analysis (Baron and
458 Kenny, 1986): first regress SR on SWM, then regress ΔR on SWM, and finally regress SR on
459 both SWM and ΔR . Subsequently, we performed a Sobel test to determine if there was a
460 significant reduction in the relationship between SWM and SR when including ΔR . The Sobel
461 test examines if the amount of variance in SR explained by SWM is significantly reduced by
462 including the mediator (Sobel, 1986). For a significant effect to be found, SWM must be a
463 significant predictor of ΔR and ΔR must also be a significant predictor of SR after controlling
464 for the effect of SWM on SR. We repeated this procedure 1000 times with the allocation of
465 trials to each fold randomised on each repetition. We present results in terms of the 95%
466 confidence intervals for the R^2 values for each of the regressions and the median p-value of the
467 Sobel test, along with the associated 95% confidence intervals. An alternative possibility to the
468 hypothesized model is that the relationship between SWM and ΔR is mediated by SR. In order
469 to compare the size of the mediation effect for these alternate models, we follow the Mackinnon
470 and Dwyer (1993) procedure and normalize the size of the indirect effect by dividing it by the
471 sum of the direct and indirect effects. This analysis allows to express the mediation effect in

472 terms of percentage of the total effect. We present the median of the normalized value for the
473 1000 repetitions on both the hypothesized and alternate models.

474

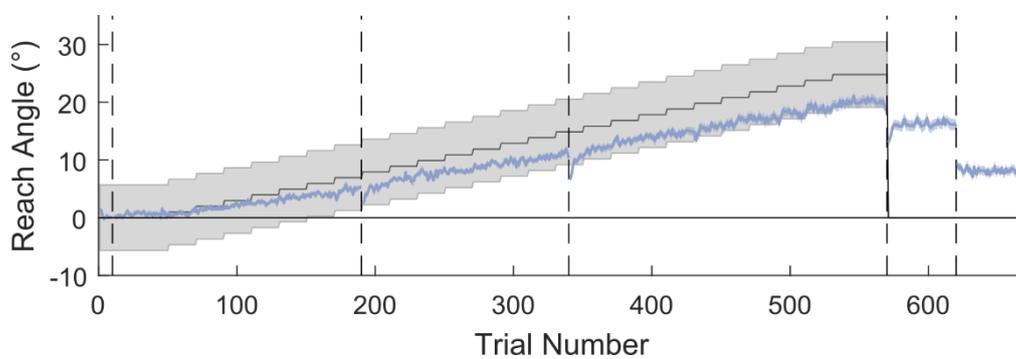
475 **Results**

476 **Acquire task**

477 In the Acquire task, participants had to learn to compensate for a secretly shifting reward region
478 in order to obtain successful feedback (Figure 2, 3). As in Holland et al. (2018), about a quarter
479 (28.1%) of participants failed to learn to compensate for the full extent of the rotation (Figure
480 3a). The inability of a significant proportion of participants to learn the full extent of the
481 rotation is also consistent with previous reports in reward-based motor learning paradigms
482 (Cashaback et al., 2019; Codol et al., 2018; Saijo and Gomi, 2010; Therrien et al., 2016, 2018).
483 Successful participants retained most of the learnt rotation (mean $80.7\% \pm 28.0\%$ SD) in the
484 maintain block. This level of retention is in accordance with that reported previously in similar
485 paradigms (Holland et al., 2018; Therrien et al., 2016). However, upon being asked to remove
486 any strategy they had been employing, their performance returned to near-baseline levels.
487 Consequently, a large explicit component to retention was found for successful participants
488 (Figure 3b), whereas both successful and unsuccessful participants manifest a small but non-
489 zero implicit component ($t(86)=9.90$, $p=7.43 \times 10^{-16}$, $d=1.061$ and $t(33)=4.53$, $p=7.39 \times 10^{-5}$,
490 $d=0.776$, respectively; Figure 3c). The persistent implicit retention is a common finding of
491 retention periods in which no visual feedback is provided and may reflect a combination of
492 implicit reinforcement (Shmuelof et al., 2012), use-dependent plasticity (Diedrichsen et al.,
493 2010), perceptual bias (Vindras et al., 1998), or perceptual recalibration (Modchalingam et al.,
494 2019). Furthermore, in accordance with Holland et al (2018), we found that participants made
495 larger ($t(120)=15.80$, $p=4.32 \times 10^{-31}$, $d=1.900$) and more variable changes in reach angle
496 following unrewarded trials ($t(120)=14.54$, $p=3.144 \times 10^{-28}$, $d=1.667$; Figure 3d-g). However, in

497 participants who would go on to fail, the post-error adjustments were smaller than in successful
498 participants ($t(119)=3.33$, $p=0.001$, $d=0.672$; Figure 3d). Changes following rewarded trials
499 were similar between the groups ($t(119)=0.71$, $p=0.48$, $d=0.143$; Figure 3f,g). The results
500 obtained in this sample ($N=121$) therefore replicate results from a previous study ($N=30$) and
501 provides further confirmation that performance in this task is fundamentally explicitly driven
502 (Holland et al., 2018).

503

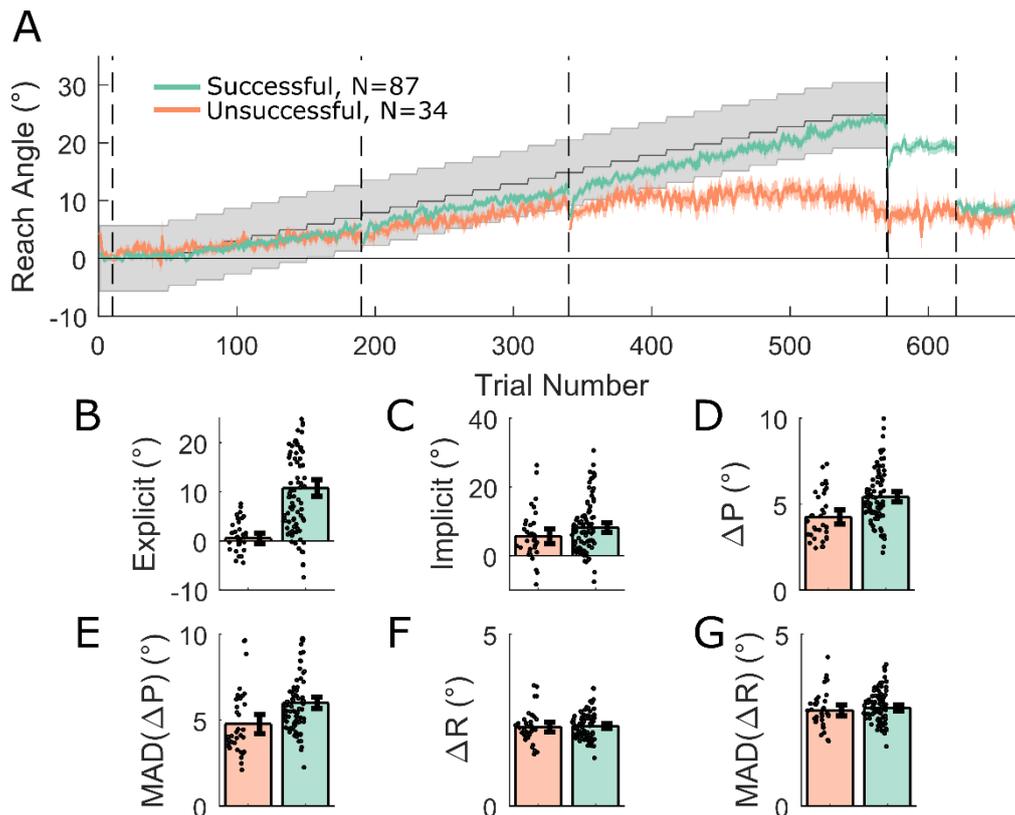


504

505 **Figure 2: Reaching performance in the Acquire task.** The grey region represents the
506 gradually rotating rewarded region, the blue line represents mean reach angle for each trial,
507 and the shaded blue region represent SEM. Vertical dashed lines represent different experiment
508 blocks or breaks. Average performance for the full cohort falls within the reward region and
509 demonstrates a clear reduction in reach angle when asked to remove strategy. $N=121$.
510

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

515



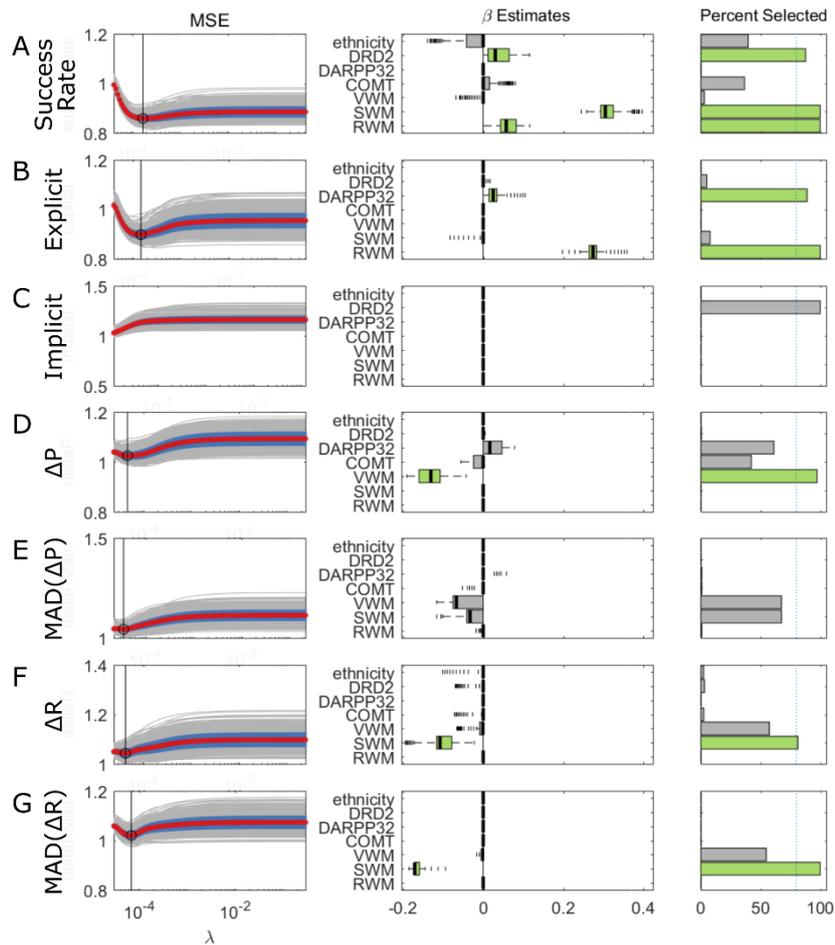
516

517 **Figure 3: Acquire task split by success at final angle.** **A:** Average reach angle for the
 518 successful (green) and unsuccessful (orange) groups, shaded regions represent SEM and grey
 519 shaded region represents the rewarded region. Despite similar initial performance, a clear
 520 divergence can be seen between the two groups and an explicit component to retention is only
 521 visible in the successful group, whereas implicit retention is similar between groups. **B-G:**
 522 subplots displaying derived measures, which acted as outcome variables for the regression
 523 analysis, separated into successful and unsuccessful participants overlaid with individual data
 524 points. Error bars represent 95% bootstrapped confidence intervals. ΔR and ΔP refer to changes
 525 made in reach angle after rewarded and unrewarded trials respectively. The bar plots in panels
 526 D and F display the median absolute change and panels E and G display the median absolute
 527 deviation of the changes in angle after each trial type.
 528

529 For success rate, SWM, RWM and DRD2 were selected as predictors (median $\beta=0.31$, 0.06,
 530 and 0.03, respectively; Figure 4a), with the strongest predictor being SWM. Figure 5 displays
 531 the effect of the strongest predictor selected for each outcome variable and shows that there
 532 was a positive relationship between SWM and success rate (Figure 5a). To ensure that the
 533 relationship between SWM and success rate was not due to failure at a later point in the task,
 534 success rate was measured during the initial period in which subjects who could not fully

535 account for the displacement are still successful; for those who could, the full learning block
 536 was included.

537

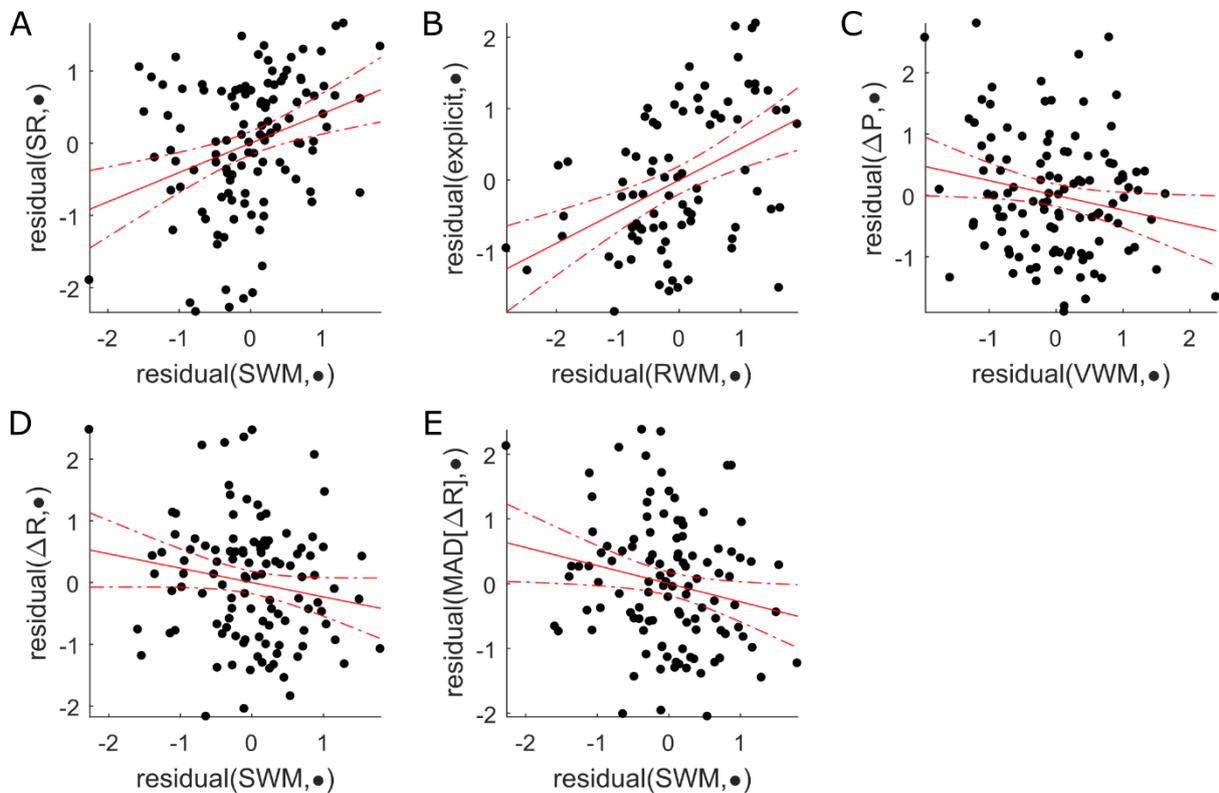


538

539 **Figure 4: Lasso regression results for the Acquire task.** Each row (A-G) represents the
 540 results from one outcome variable. The left column indicates the MSE as a function of changing
 541 the shrinkage parameter λ , with larger values of λ representing greater penalization and sparser
 542 models. A minimum in the MSE within its defined boundaries indicates the suitability of that
 543 choice of λ and is indicated with a vertical line. Given the presence of a minimum, the values
 544 of the β for each predictor are taken. We performed 1000 repetitions of the lasso regression for
 545 each outcome variable and box plots indicating the distribution of the coefficient estimates are
 546 displayed in the middle panel. The rightmost column indicates the percentage of times that the
 547 individual predictors were assigned non-zero coefficients. We employed a threshold of 80%
 548 (indicated with a dashed vertical line) to signify that a particular predictor was robustly selected,
 549 and these variables are highlighted in green. Median absolute change in reach angle after
 550 rewarded (ΔR) and unrewarded (ΔP) trials. Median absolute deviation of change in reach angle
 551 after rewarded ($MAD[\Delta R]$) and unrewarded ($MAD[\Delta P]$) trials.

552

553 Next, retention was assessed by splitting up the explicit and implicit components such as in
 554 Holland et al. (2018). No predictor was related to the implicit component, but the explicit
 555 component was strongly and positively associated with RWM ($\beta=0.27$; Figure 4b, 5b) with a
 556 weaker association between DARPP32 and explicit retention ($\beta=0.03$). These results suggest
 557 positive relationships for both RWM and SWM with task performance: greater RWM predicts
 558 a greater contribution of explicit processes to learning, whereas greater SWM predicts a greater
 559 percentage of correct trials.
 560



561
 562 **Figure 5: Added variable plots for selected predictors in the Acquire task.** Each plot
 563 displays the relationship between the strongest predictor selected by the lasso regression (x-
 564 axis), and the corresponding outcome variable (y-axis). Added variable plots display the
 565 residuals of regressing the response variable with all remaining independent variables, and the
 566 residuals of the regression of the selected predictor to the remaining predictors. The resulting
 567 relationship corresponds to the effect of the selected predictor on the outcome measure after
 568 controlling for the remaining predictors. SR: Success Rate. Median absolute change in reach
 569 angle after rewarded (ΔR) and unrewarded (ΔP) trials. MAD(ΔR): Median absolute deviation
 570 of change in reach angle after rewarded trials.
 571

572 In Holland et al (2018), the amplitude of the changes in reach angle participants made following
573 unrewarded trials was found to be predictive of task success, that is, greater ΔP was predictive
574 of an increased chance of overall task success. Thus, it could be that RWM and SWM, that are
575 shown to associate with performance in this study, are themselves predictors of changes in
576 reach angle. Conformingly, the regression results demonstrated that a large ΔR was inversely
577 related to SWM ($\beta=-0.11$; Figure 4f, 5d), as was $MAD[\Delta R]$ ($\beta=-0.17$; Figure 4g, 5e). The
578 results indicate that greater SWM was predictive of smaller and less variable changes in reach
579 angle after successful trials, suggesting high SWM enables the maintenance of rewarding reach
580 angles. It was also found that changes in reach angle following unrewarded trials (ΔP) were
581 negatively associated with VWM ($\beta=-0.13$, Figure 4d, 5c). This result was unexpected as it
582 suggests that greater WMC predicts smaller changes following unrewarded trials, whereas
583 previous results suggest a positive relationship between the amplitude of these changes and
584 overall task success. Although the difference may be due to the domain of WM under
585 consideration, it is unclear as to the reason for this relationship. Another important aspect of
586 the analysis of trial-to-trial changes to control for is that the numbers of trials analysed and
587 their phase in the experiment differs between successful and unsuccessful subjects. Therefore,
588 we repeated the Lasso regression while only including successful subjects. The predictors that
589 were selected were identical to those selected when using the full participant pool.

590 Overall, WM (in particular RWM and SWM) successfully predicted various aspects of
591 performance in the Acquire task, while genetic predictors generally failed to do so. Specifically,
592 greater SWM predicted smaller and less variable changes after correct trials. This suggests that
593 SWM underlies one's capacity to preserve and consistently express an acquired reach direction
594 to obtain reward. Furthermore, SWM also directly predicted success rate. In addition, greater
595 RWM was a strong predictor of explicit control. The inverse relationship between VWM and
596 the magnitude of changes after unrewarded trials was unexpected. However, one possible

597 explanation is that participants with poorer WMC make larger errors which require larger
598 corrections.

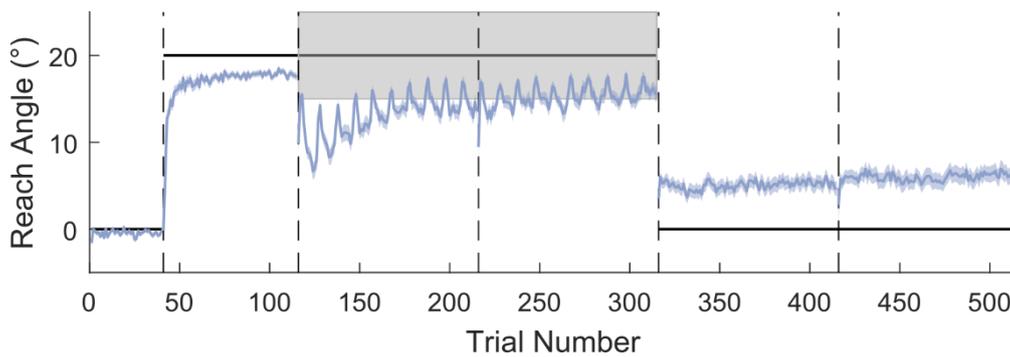
599

600 **Preserve task**

601 In this task, we addressed how well participants can maintain a previously learnt adaptation
602 after transitioning to binary feedback. As participants are unable to compensate for a large
603 abrupt displacement of a hidden reward region (van der Kooij and Overvliet, 2016; Manley et
604 al., 2014), participants first adapted to an abruptly introduced 20° clockwise rotation with full
605 vision of the cursor available. Subsequently, visual feedback of the cursor position was
606 replaced with binary feedback; participants were rewarded if they continued reaching towards
607 the same angle that resulted in the cursor hitting the target during the adaptation phase. Overall,
608 participants adapted to the visuomotor rotation successfully (Figure 6, 7a-c) before
609 transitioning to the binary feedback-based asymptote blocks. However, from the start of the
610 asymptote blocks onward, participants exhibited very poor performance, expressing an average
611 45.0 ± 24.2 SD% success rate when considering all 200 asymptote trials (Figure 6, 7a, d,e). We
612 have previously shown in (Codol et al., 2018) that this drop in performance (Shmuelof et al.,
613 2012) represents exploratory behaviour that arises due to a lack of transfer of the cerebellar
614 memory between the two contexts. Separating successful and unsuccessful participants (40%
615 success rate cut-off; Figure 7a) revealed that successful participants expressed behaviour
616 greatly similar to that observed in Codol et al. (2018), in which unsuccessful participants were
617 excluded, using the same cut-off (40% success rate). The ‘spiking’ behaviour observed in reach
618 angles during the asymptote blocks (Figure 7a) is due to the presence of the ‘refresher’ trials,
619 with the large positive changes in reach angle corresponding to trials immediately following
620 the refresher trials. This pattern of behaviour is particularly pronounced in the unsuccessful
621 participants. Finally, participants demonstrated at least a residual level of retention even after

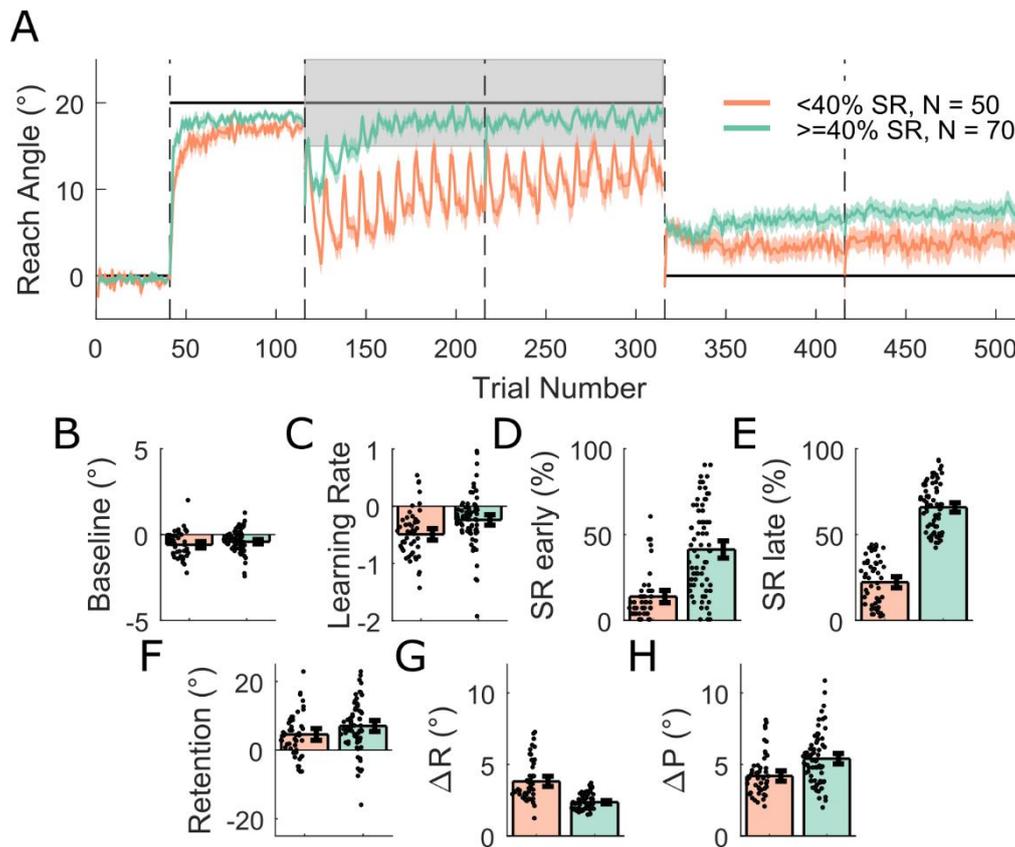
622 being instructed to remove any strategy they had employed ($t(69)=7.268$, $p=3.345 \times 10^{-10}$,
623 $d=0.869$; Figure 7a,f). Therefore, the results obtained in this sample ($N=120$) replicate results
624 from a previous study (Codol et al., 2018; $N=20$, BF-Remove group) and provides further
625 confirmation that performance in this task is fundamentally explicitly driven. It should also be
626 noted that the successful group displayed higher implicit retention than the unsuccessful
627 participants. As with the Acquire task successful participants displayed larger changes in angle
628 after unrewarded trials than their unsuccessful counterparts ($t(117)=3.847$, $p=1.952 \times 10^{-4}$,
629 $d=0.717$; Figure 7h). However, in contrast to the Acquire task, successful participants also
630 displayed smaller changes in angle after rewarded trials ($t(115)=-7.534$, $p=1.218 \times 10^{-11}$,
631 $d=1.421$; Figure 7g).

632



633

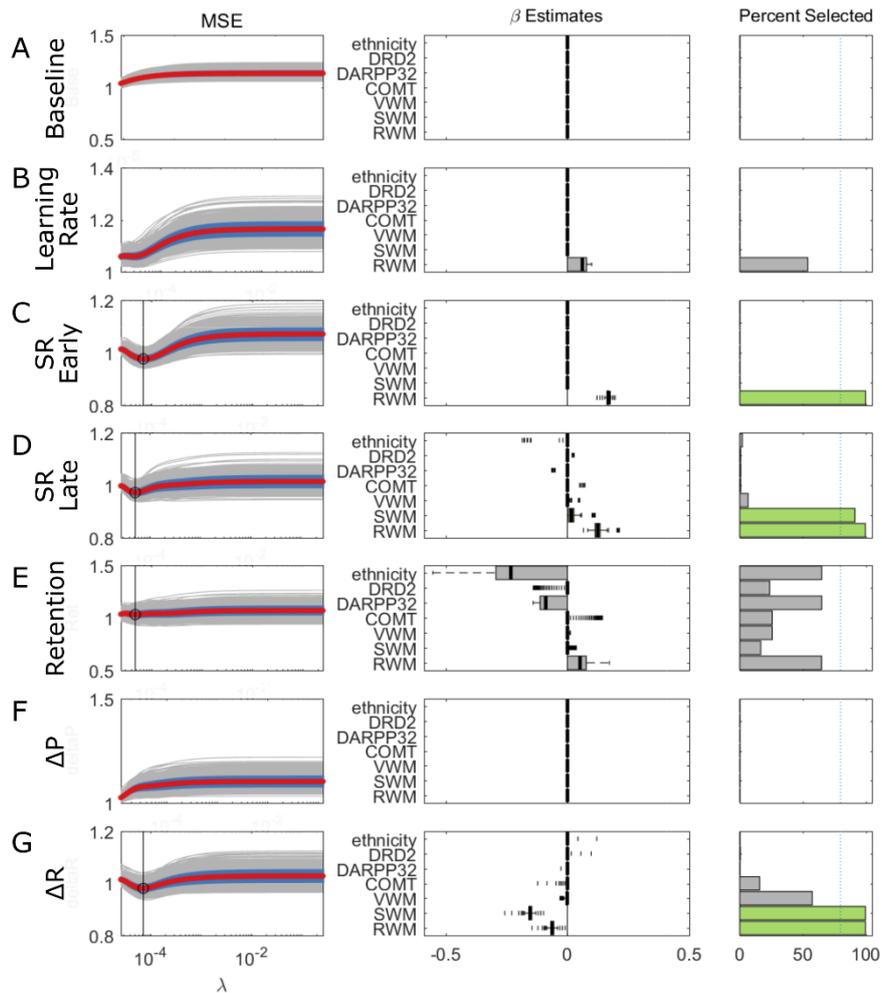
634 **Figure 6: Reaching performance in the Preserve task.** The grey shaded area represents the
635 rewarded region, and the thick black line represents the perturbation. The vertical dashed lines
636 represent block limits. The blue line indicates mean reach angle for every trial and blue shaded
637 areas represent SEM. After successfully adapting to the visuomotor rotation performance
638 deteriorates at the onset of binary feedback, subsequently success rate increases towards the
639 end of the asymptote blocks. Following the removal of all feedback, and the instruction to
640 remove any strategy, a small amount of implicit retention remains. $N=120$.



641

642 **Figure 7: Preserve task split into two groups on the basis of success rate.** A: Shaded regions
 643 represent SEM. B-H: Derived variables, which acted as outcome variables for the regression
 644 analysis, for the two groups, error bars on the bars represent 95% bootstrapped confidence
 645 intervals and individual data points are displayed. SR: Success Rate. Median absolute change
 646 in reach angle after rewarded (ΔR) and unrewarded (ΔP) trials.
 647

648 As in the Acquire task, we examined if performance in any of the WM tasks or genetic profile
 649 could predict participants' behaviour in the reaching task. We performed separate lasso
 650 regressions for the following outcome variables: baseline reach direction as a control variable,
 651 learning rate in the adaptation block, early and late success rate in the asymptote blocks (first
 652 30 and last 170 trials; Codol et al., 2018), retention in the no-feedback blocks, and ΔR and ΔP
 653 during the asymptote blocks. The most striking result was that both early and late success rate
 654 could be reliably predicted by RWM (early: $\beta=0.17$, late: $\beta=0.12$; Figure 8c,d, and 9a,b), with
 655 greater RWM associated with increased success rates. An additional positive relationship was
 656 found between SWM and success rate but only during the later period ($\beta=0.02$; Figure 8c).

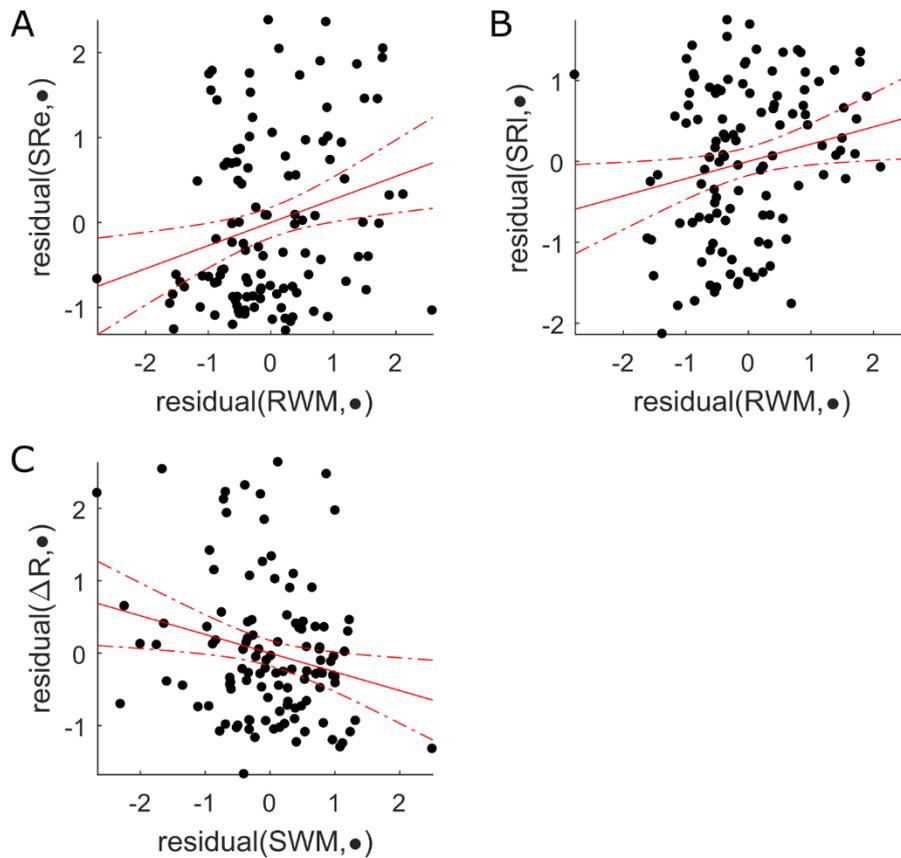


657

658 **Figure 8: Results of the Lasso regressions for the Preserve task.** The format is identical to
 659 Figure 4 with each row (A-G) representing the predictors of a single outcome measure. Selected
 660 predictors are highlighted in green, with the middle panels displaying the β estimates and the
 661 right panels displaying the probability of each predictor being selected. SR: Success Rate.
 662 Median absolute change in reach angle after rewarded (ΔR) and unrewarded (ΔP) trials.
 663

664 Genetic profile did not predict any aspect of performance. In contrast, greater SWM
 665 successfully predicted reduced ΔR ($\beta=-0.15$; Figure 8g, 9c) similarly to the Acquire task.
 666 Additionally, there was a weaker relationship between RWM and ΔR ($\beta=-0.06$; Figure 8g)
 667 which was absent in the Acquire task. Despite the presence of a local minimum in the MSE for
 668 the regression involving retention, no individual predictor was consistently selected in more
 669 than 80% of repetitions (Figure 8e).

670



671

672 **Figure 9: Added variable plots for selected predictors in the Preserve task.** Each panel (A-
 673 C) displays the effect of the considered predictor when accounting for the effect of all other
 674 predictors. Results are displayed for the strongest selected predictor for each outcome measure.
 675 SRe: Early Success Rate. SRI: Late Success Rate. ΔR : Median absolute change in reach angle
 676 after rewarded trials.

677

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

684 Furthermore, we analysed the data using group lasso (Boyd, 2010; Yuan and Lin, 2006)
 685 regression in order to check for the possibility that our analysis was insensitive to categorical
 686 predictors (the genetic variables). The group lasso is an extension to lasso regression in which
 687 predictor variables can be assigned to groups. Although each member of a group can be

688 assigned a different β , the group lasso applies the regularisation penalty to all members of the
689 group, leading to the removal of all members of the group from the model at the same value of
690 λ . We employed reference dummy variable coding for each genetic variable and treated the
691 dummy variables representing each SNP as a group for the purposes of the group lasso; this
692 ensures that the dummy variables representing each genetic factor are removed from the
693 regression at the same time. The results of the group lasso analysis replicate those of the
694 standard lasso and furthermore no genetic predictors were found for any outcome variable in
695 either task. The results obtained for both tasks via the lasso regression methods are similar to
696 those obtained using a stepwise regression procedure. All data and code are available online,
697 including the procedures, results, and significance tests of the lasso and stepwise regression
698 analysis.

699

700 **Relationships between predictors**

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

709

710 **Partial Correlation Analysis**

711 In order to understand if the RWM and SWM measures have separable effects on the outcome
712 measures considered here, we performed a partial correlation analysis examining the

713 relationships between RWM, SWM, and success rate in both tasks. After controlling for the
714 effect of RWM, SWM remained significantly correlated with success in both tasks ($r=0.343$,
715 $p=0.005$ Preserve, $r=0.488$, $p=6.823 \times 10^{-6}$ Acquire). However, the partial correlation between
716 RWM and success rate was not significant for either task, indicating that even in the Preserve
717 task SWM plays a dominant role in determining success rate.

718

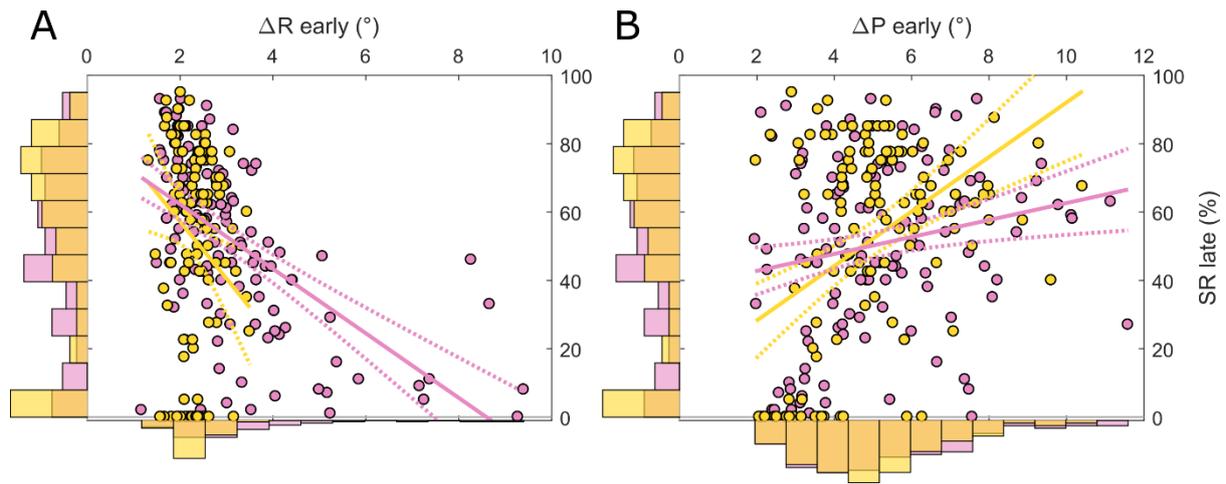
719 **Exploratory analysis**

720 As a relationship exists between SWM and ΔR in both the Acquire and Preserve paradigms,
721 we ran exploratory regressions to assess the relationship between ΔR and success rate across
722 both tasks. Since ΔR and success rate are conceptually strongly related variables, and
723 measuring on the same data set would render them non-independent, we split each individual's
724 reaching data into two sections and assessed whether ΔR or ΔP in the first section could reliably
725 predict success rate in the second (see methods for details). Although we found no predictors
726 of ΔP in our primary analysis, results here in combination with previous work (Holland et al.,
727 2018) has demonstrated a link between ΔP and task success, with a greater ΔP indicative of
728 greater success. Therefore, we also performed the same analysis for ΔP .

729

730

731



732

733 **Figure 10: Slice plots showing regression results for prediction of late success rate (SR)**
 734 **by changes in reach angle following rewarded (A) and unrewarded (B) trials during the**
 735 **early learning period.** The central axis of each panel displays the individual data from the
 736 Acquire (yellow) and Preserve (pink) task, a histogram displaying the distribution of the data
 737 in each dimension is presented on the corresponding axis. Solid lines represent the prediction
 738 of the regression model when the other predictor is held at its mean value. SR: Success Rate.
 739 Median absolute change in reach angle after rewarded (ΔR) and unrewarded (ΔP) trials.
 740

741 In the Acquire task, ΔR and ΔP in the first section of learning trials predicted success rate in
 742 the final twenty trials, though ΔP appeared as the strongest predictor (ΔR : $\beta=-0.274$, $p=0.015$;
 743 ΔP : $\beta=0.581$, $p=3.89 \times 10^{-6}$; Figure 10a,b, yellow; Table 2). Similarly, for the Preserve task, ΔR
 744 and ΔP in the first half of asymptote trials predicted success rate in the second half (ΔR : $\beta=-$
 745 0.750 , $p=1.07 \times 10^{-12}$; ΔP : $\beta=0.229$, $p=0.007$; Figure 10a,b, pink; Table 2). In both tasks, the
 746 directions of these relationships were opposite; greater success rate was predicted by smaller
 747 ΔR and greater ΔP . In summary, we found that for both tasks the magnitude of changes in
 748 behaviour in response to rewarded and unrewarded trials early in learning were strongly
 749 predictive of future task success across both the Acquire and Preserve tasks.

750

751

752

753

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

754

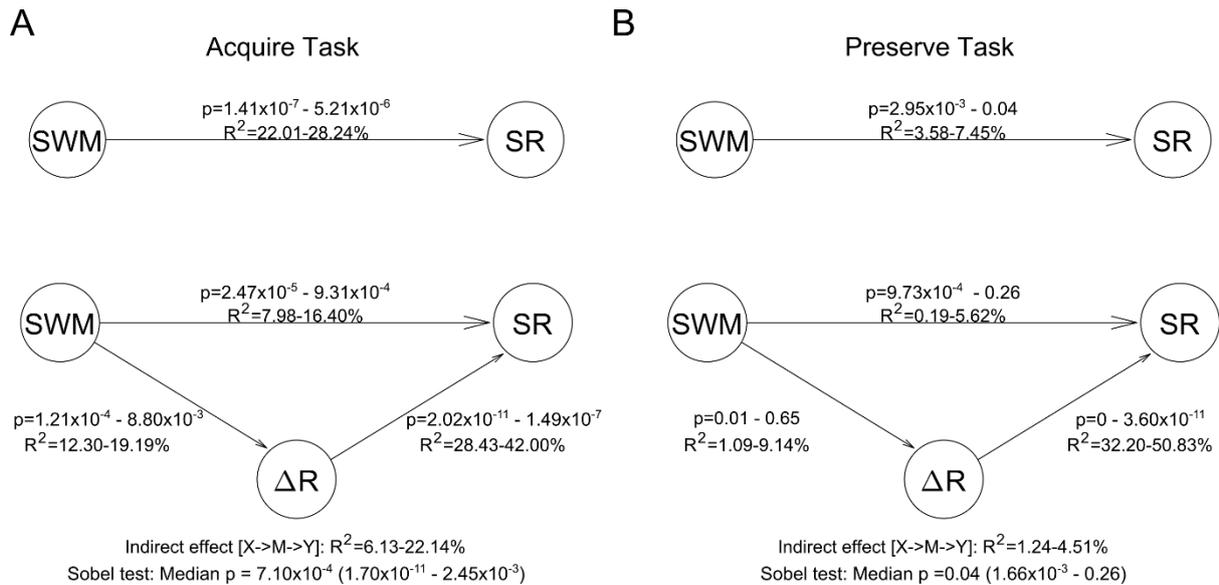
755 **Table 2: Regression results for split data for both the Acquire and Preserve tasks.**
756 Ordinary least squares linear regressions were performed with both ΔR and ΔP included as
757 predictors. The regression coefficient, standard error and p value for each predictor are reported
758 along with the significance of the comparison between the model and an intercept only model.
759 In both tasks there is an opposing relationship between ΔR and ΔP and success rate, with
760 smaller changes after rewarded trials and larger changes after unrewarded trials predictive of
761 success. SR: Success Rate. Median absolute change in reach angle after rewarded (ΔR) and
762 unrewarded (ΔP) trials.
763

764 **Mediation analysis**

765 Finally, to test whether the effect observed between SWM and SR was explained by an indirect
766 effect through ΔR , we performed an exploratory mediation analysis on both tasks. For both the
767 Acquire and Preserve tasks, the results indicate a significant proportion (median p=7.10×10⁻⁴
768 and p=0.04 respectively) of the relationship between SWM and SR can be explained by a
769 mediation from SWM via ΔR to SR (Figure 11). However, in the case of the Acquire task
770 (Figure 11a), a significant relationship between SWM and SR also remained, indicating that not
771 all of the effect of SWM on SR could be explained by the indirect pathway. Of note, in the
772 Preserve task (Figure 11b) the SWM- ΔR relationship was weaker and was not significant on
773 every repetition, occasionally leading to an insignificant mediation effect, despite the median
774 p-value indicating an effect when considering all repetitions. We also examined an alternative
775 possibility to the hypothesized model in which relationship between SWM and ΔR is mediated
776 by SR. We found that 31.20% of the total effect is mediated in the Acquire task using the

777 hypothesized model, in contrast to only 0.17% in the alternative model. Similarly, in the
 778 Preserve task the hypothesized model displayed a substantially larger mediation effect
 779 (44.77%) than the alternative model (5.02%). These results support the application of the
 780 hypothesized model.

781



782

783 **Figure 11: Mediation Analysis for both the Acquire (A) and Preserve (B) tasks.** The
 784 numbers associated with each arrow display the 95% confidence intervals for each of the
 785 relationships (R^2 and p -values) across the 1000 repetitions. Below the figure, the results of
 786 the Sobel test are displayed indicating the amount of variance explained by the indirect
 787 pathway and the 95% confidence intervals and median p -value. SR: Success Rate. ΔR :
 788 Median absolute change in reach angle after rewarded trials.

789

790 Discussion

791 In this study, we sought to identify if genetic background or specific domains of WMC could
 792 explain the variability observed in performance levels during reward-based motor learning
 793 tasks. We found that RWM and SWM predicted different aspects of the Acquire and Preserve
 794 tasks, whereas VWM only related to one performance measure (ΔP), but not consistently across
 795 tasks. Specifically, RWM predicted the explicit component of retention in the Acquire task and

796 success rate in the Preserve task, whereas SWM predicted success rate in the Acquire task and
797 the late period of the Preserve task. Furthermore SWM negatively predicted ΔR in both tasks.
798 Conversely, allelic variations of the three dopamine-related genes (DRD2, COMT and
799 DARPP32) did not consistently predict any behavioural variables across both tasks. This
800 suggests that SWM predicts a participant's capacity to reproduce a rewarded motor action,
801 while RWM predicts a participant's ability to express an explicit strategy when making large
802 behavioural adjustments. Therefore, we conclude that WMC plays a pivotal role in determining
803 individual ability in reward-based motor learning.

804 Recently, Wong et al. (2019) described a positive relationship between SWM and the
805 development of explicit strategies in visuomotor adaptation, complementing previous reports
806 (Anguera et al., 2012; Christou et al., 2016; Vandevorde and Orban de Xivry, 2019). However,
807 in contrast to the current findings the previous experiments employed relatively small sample
808 sizes, which may render correlations unreliable. The large group sizes employed here, and the
809 confirmation of relationships across two tasks, provides strong evidence that these relationships
810 are robust, replicable, and extend from visuomotor adaptation to reward-based motor learning.

811 An interesting dichotomy was the reliance on SWM and RWM for the Acquire and Preserve
812 task, respectively. While the Preserve task required the maintenance of a large, abrupt
813 behavioural change, the Acquire task required the gradual adjustment of behaviour considering
814 the outcomes of recent trials. Therefore, RWM may underscore one's capacity to express a
815 large correction consistently over trials with binary feedback, whereas SWM reflects one's
816 capacity to maintain a memory of previously rewarded actions and adjust behaviour
817 accordingly. Accordingly, McDougle and Taylor (2019) demonstrated a mental rotation
818 process is employed in countering a visuomotor rotation, and Sidarta et al. (2018) reported that
819 higher SWM is associated with reduced movement variability in a reward-based motor learning
820 task. Here, the magnitude of ΔR was negatively related to SWM but not RWM in both tasks,

821 suggesting high SWM enables the maintenance of rewarding actions. Additionally, explicit
822 retention, an element of the Acquire task requiring a large, sudden change in reach direction,
823 was predicted by RWM rather than SWM. Notably, RWM and SWM were often selected as
824 predictors simultaneously. The overlapping but distinct pattern of relationships between RWM,
825 SWM, and outcome measures considered here supports the view that they share substrates but
826 have different patterns of dependency on executive functions (Miyake et al., 2001).

827 A notable feature of the Preserve task is the ‘spiking’ behaviour observed immediately
828 following ‘refresher’ trials, suggesting a central role of refresher trials in binary feedback-based
829 performance when included (Codol et al., 2018; Shmuelof et al., 2012). The transient nature of
830 this decrease in error demonstrates this is insufficient to promote generalisation to binary
831 feedback trials, at least in unsuccessful participants. It remains an open question whether
832 superior performance of successful participants was partly due to a capacity to generalise
833 information from ‘refresher’ trials. McDougle and Taylor (2019) suggest that two separate
834 strategies are employed in visuomotor adaptation: response-caching and mental rotation. The
835 balance between the two strategies is a function of task demands. The relationships between
836 RWM and SWM to success rate in the Preserve and Acquire tasks respectively may reflect a
837 different balance of the use of these strategies. Visual feedback in ‘refresher’ trials in the
838 Preserve task may engage mental rotation processes, whereas the slow updating of behaviour
839 in the Acquire task engages the response-caching memory system. This would imply that
840 response-caching is associated with SWM.

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

846 If RWM is important for explicit control and the main element predicting success in the
847 Preserve task, it is worth considering whether gradual designs (as in the Acquire task) are more
848 suitable to engage implicit reinforcement learning, at least initially. However, the Acquire task
849 still bears a strong explicit component (Holland et al., 2018). How can these two views be
850 reconciled? In reward-based motor learning tasks, it is observed that participants begin to
851 reflect upon task structure and develop strategies upon encountering negative outcomes (Leow
852 et al., 2016; Loonis et al., 2017; Maxwell et al., 2001), which occurs nearly immediately in the
853 Preserve task after the introduction of binary feedback, due to a lack of generalisation of
854 cerebellar memory (Codol et al., 2018). In contrast, in the Acquire task, participants experience
855 an early learning phase with mainly rewarding outcomes, possibly suppressing development of
856 explicit control and allowing for this early window of implicit reward-based learning. Other
857 studies have demonstrated that minor adjustments in reach direction under reward-based
858 feedback can occur, though none has assessed their explicitness directly in the very early stages
859 (Izawa and Shadmehr, 2011; Pekny et al., 2015; Therrien et al., 2016). Notably, Izawa and
860 Shadmehr, (2011) observed that after 8° shifts of a similarly-sized reward region, participants
861 indeed noticed the perturbation, but awareness was not assessed for smaller shifts.

862 In Holland et al., (2018), the addition of a RWM-like dual-task was very effective in preventing
863 explicit control, leading to participants invariably failing at the reaching task. Therefore, it may
864 seem surprising that RWM does not predict success rate in the Acquire task. A possible
865 explanation is that RWM and SWM share the same memory buffer (Anguera et al., 2010;
866 Beschin et al., 2005; Cohen et al., 1996; Jordan et al., 2001; Suchan et al., 2006). Similarly, in
867 force-field adaptation the early component of adaptation – considered as bearing a strong
868 explicit element – is selectively disrupted with a VWM dual-task (Keisler and Shadmehr, 2010).

869 However, we found no consistent relationship with VWM across our reward-based motor tasks.

870 It may be possible that reward-based motor performance relies more on spatial instances of
871 WM as opposed to tasks such as force-field adaptation.

872 The absence of DA-related genetic relationships with behaviour is a surprising result as a
873 substantial body of literature points to a relationship between dopamine and performance in
874 reward-based tasks, including those with motor components (Deserno et al., 2015; Doll et al.,
875 2016; Frank et al., 2007, 2009; Gershman and Schoenbaum, 2017; Izawa and Shadmehr, 2011;
876 Nakahara and Hikosaka, 2012; Pekny et al., 2015; Therrien et al., 2016). There is a growing
877 appreciation of the links between decision-making and motor learning (Chen et al., 2017, 2018;
878 Haith and Krakauer, 2013). However, the results presented here suggest that genetic predictors
879 of exploration and exploitation in decision-making tasks are not also predictive of similar
880 behaviours in reward-based motor learning.

881 Our sample sizes were defined *a priori* for 90% power based on previous work (Doll et al.,
882 2016; Frank et al., 2009; see pre-registrations), and are unlikely to be underpowered. Another
883 possibility is that we employed the wrong variables to assess behaviour. However, given the
884 informative and coherent relationships between WM and motor learning, it could be that the
885 SNPs we selected do not meaningfully relate to performance in reward-based motor tasks
886 compared to WM. A similar claim was made in the decision-making literature (Collins and
887 Frank, 2012). In line with this, a recent study showed that DA pharmacological manipulation
888 did not alter reward effects in a visuomotor adaptation task (Quattrocchi et al., 2018). However,
889 previous work has shown that Parkinson's disease patients show impaired reward-based motor
890 performance (Pekny et al., 2015). It is possible that genetic variations may simply not impact
891 reward-based motor learning significantly, especially compared to the wide depletion of
892 dopaminergic neurons in Parkinson's disease. It is also important to note that while we refer to
893 both of our tasks as reward-based motor learning, they are both in essence visuomotor rotation

894 paradigms. In future it is important to investigate if these findings extend to more complex
895 reward-based motor learning paradigms.

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

904

905 **References**

- 906 Anguera, J.A., Reuter-Lorenz, P.A., Willingham, D.T., and Seidler, R.D. (2010).
907 Contributions of spatial working memory to visuomotor learning. *Journal of Cognitive*
908 *Neuroscience* 22, 1917–1930.
- 909 Anguera, J.A., Bernard, J.A., Jaeggi, S.M., Buschkuhl, M., Benson, B.L., Jennett, S.,
910 Humfleet, J., Reuter-Lorenz, P.A., Jonides, J., and Seidler, R.D. (2012). The effects of
911 working memory resource depletion and training on sensorimotor adaptation. *Behavioural*
912 *Brain Research* 228, 107–115.
- 913 Baron, R.M., and Kenny, D.A. (1986). The Moderator-Mediator Variable Distinction in
914 *Social Psychological Research: Conceptual, Strategic, and Statistical Considerations. Journal*
915 *of Personality and Social Psychology* 51, 1173–1182.

916 Beschin, N., Denis, M., Logie, R.H., and Della Sala, S. (2005). Dissociating mental
917 transformations and visuo-spatial storage in working memory: Evidence from
918 representational neglect. *Memory* *13*, 430–434.

919 Bond, K.M., and Taylor, J.A. (2015). Flexible explicit but rigid implicit learning in a
920 visuomotor adaptation task. *Journal of Neurophysiology* *113*, 3836–3849.

921 Boyd, S. (2010). Distributed Optimization and Statistical Learning via the Alternating
922 Direction Method of Multipliers. *FNT in Machine Learning* *3*, 1–122.

923 Brainard, D.H. (1997). The Psychophysics Toolbox. *Spat Vis* *10*, 433–436.

924 Buszard, T., and Masters, R.S.W. (2018). Adapting, correcting and sequencing movements:
925 does working-memory capacity play a role? *International Review of Sport and Exercise*
926 *Psychology* *11*, 258–278.

927 Cashaback, J.G.A., McGregor, H.R., Mohatarem, A., and Gribble, P.L. (2017). Dissociating
928 error-based and reinforcement-based loss functions during sensorimotor learning. *PLoS*
929 *Comput Biol* *13*, e1005623.

930 Cashaback, J.G.A., Lao, C.K., Palidis, D.J., Coltman, S.K., McGregor, H.R., and Gribble,
931 P.L. (2019). The gradient of the reinforcement landscape influences sensorimotor learning.
932 *PLoS Comput Biol* *15*, e1006839.

933 Chen, X., Mohr, K., and Galea, J.M. (2017). Predicting explorative motor learning using
934 decision-making and motor noise. *PLoS Comput Biol* *13*, e1005503.

935 Chen, X., Holland, P., and Galea, J.M. (2018). The effects of reward and punishment on
936 motor skill learning. *Current Opinion in Behavioral Sciences* *20*, 83–88.

937 Christou, A.I., Miall, R.C., McNab, F., and Galea, J.M. (2016). Individual differences in
938 explicit and implicit visuomotor learning and working memory capacity. *Scientific Reports* 6.

939 Codol, O., Holland, P.J., and Galea, J.M. (2018). The relationship between reinforcement and
940 explicit control during visuomotor adaptation. *Scientific Reports* 8.

941 Cohen, M.S., Kosslyn, S.M., Breiter, H.C., DiGirolamo, G.J., Thompson, W.L., Anderson,
942 A.K., Bookheimer, S.Y., Rosen, B.R., and Belliveau, J.W. (1996). Changes in cortical
943 activity during mental rotation A mapping study using functional MRI. *Brain* 119, 89–100.

944 Collins, A.G.E., and Frank, M.J. (2012). How much of reinforcement learning is working
945 memory, not reinforcement learning? A behavioral, computational, and neurogenetic
946 analysis: Working memory in reinforcement learning. *European Journal of Neuroscience* 35,
947 1024–1035.

948 Daw, N.D., Niv, Y., and Dayan, P. (2005). Uncertainty-based competition between prefrontal
949 and dorsolateral striatal systems for behavioral control. *Nat. Neurosci.* 8, 1704–1711.

950 Daw, N.D., Gershman, S.J., Seymour, B., Dayan, P., and Dolan, R.J. (2011). Model-Based
951 Influences on Humans' Choices and Striatal Prediction Errors. *Neuron* 69, 1204–1215.

952 Deserno, L., Huys, Q.J.M., Boehme, R., Buchert, R., Heinze, H.-J., Grace, A.A., Dolan, R.J.,
953 Heinz, A., and Schlagenhaut, F. (2015). Ventral striatal dopamine reflects behavioral and
954 neural signatures of model-based control during sequential decision making. *Proceedings of*
955 *the National Academy of Sciences* 112, 1595–1600.

956 Diedrichsen, J., White, O., Newman, D., and Lally, N. (2010). Use-Dependent and Error-
957 Based Learning of Motor Behaviors. *Journal of Neuroscience* 30, 5159–5166.

958 Doll, B.B., Duncan, K.D., Simon, D.A., Shohamy, D., and Daw, N.D. (2015). Model-based
959 choices involve prospective neural activity. *Nature Neuroscience* *18*, 767–772.

960 Doll, B.B., Bath, K.G., Daw, N.D., and Frank, M.J. (2016). Variability in Dopamine Genes
961 Dissociates Model-Based and Model-Free Reinforcement Learning. *Journal of Neuroscience*
962 *36*, 1211–1222.

963 Egan, M.F., Goldberg, T.E., Kolachana, B.S., Callicott, J.H., Mazzanti, C.M., Straub, R.E.,
964 Goldman, D., and Weinberger, D.R. (2001). Effect of COMT Val108/158 Met genotype on
965 frontal lobe function and risk for schizophrenia. *Proceedings of the National Academy of*
966 *Sciences* *98*, 6917–6922.

967 Frank, M.J., Moustafa, A.A., Haughey, H.M., Curran, T., and Hutchison, K.E. (2007).
968 Genetic triple dissociation reveals multiple roles for dopamine in reinforcement learning.
969 *Proceedings of the National Academy of Sciences* *104*, 16311–16316.

970 Frank, M.J., Doll, B.B., Oas-Terpstra, J., and Moreno, F. (2009). Prefrontal and striatal
971 dopaminergic genes predict individual differences in exploration and exploitation. *Nature*
972 *Neuroscience* *12*, 1062–1068.

973 Gershman, S.J., and Schoenbaum, G. (2017). Rethinking dopamine prediction errors.
974 *BioRxiv* 239731.

975 Goldberg, T.E., Egan, M.F., Gscheidle, T., Coppola, R., Weickert, T., Kolachana, B.S.,
976 Goldman, D., and Weinberger, D.R. (2003). Executive Subprocesses in Working Memory:
977 Relationship to Catechol-O-methyltransferase Val158Met Genotype and Schizophrenia. *Arch*
978 *Gen Psychiatry* *60*, 889–896.

979 Haith, A.M., and Krakauer, J.W. (2013). Model-Based and Model-Free Mechanisms of
980 Human Motor Learning. In *Progress in Motor Control*, M.J. Richardson, M.A. Riley, and K.
981 Shockley, eds. (New York, NY: Springer New York), pp. 1–21.

982 Hastie, T., Tibshirani, R., and Wainwright, M. (2015). Chapter 6. Statistical Inference: The
983 Bootstrap. In *Statistical Learning with Sparsity: The Lasso and Generalizations*, (Boca Raton:
984 Chapman and Hall/CRC), p.

985 Holland, P.J., Codol, O., and Galea, J.M. (2018). The contribution of explicit processes to
986 reinforcement-based motor learning. *Journal of Neurophysiology*.

987 Huang, V.S., Haith, A., Mazzoni, P., and Krakauer, J.W. (2011). Rethinking Motor Learning
988 and Savings in Adaptation Paradigms: Model-Free Memory for Successful Actions Combines
989 with Internal Models. *Neuron* 70, 787–801.

990 Izawa, J., and Shadmehr, R. (2011). Learning from Sensory and Reward Prediction Errors
991 during Motor Adaptation. *PLoS Computational Biology* 7, e1002012.

992 Jordan, K., Heinze, H.-J., Lutz, K., Kanowski, M., and Jäncke, L. (2001). Cortical
993 Activations during the Mental Rotation of Different Visual Objects. *NeuroImage* 13, 143–
994 152.

995 Keisler, A., and Shadmehr, R. (2010). A Shared Resource between Declarative Memory and
996 Motor Memory. *Journal of Neuroscience* 30, 14817–14823.

997 van der Kooij, K., and Overvliet, K.E. (2016). Rewarding imperfect motor performance
998 reduces adaptive changes. *Experimental Brain Research* 234, 1441–1450.

999 Leow, L.-A., de Rugy, A., Marinovic, W., Riek, S., and Carroll, T.J. (2016). Savings for
1000 visuomotor adaptation require prior history of error, not prior repetition of successful actions.
1001 *Journal of Neurophysiology* *116*, 1603–1614.

1002 Loonis, R.F., Brincat, S.L., Antzoulatos, E.G., and Miller, E.K. (2017). A Meta-Analysis
1003 Suggests Different Neural Correlates for Implicit and Explicit Learning. *Neuron* *96*, 521-
1004 534.e7.

1005 Mackinnon, D.P., and Dwyer, J.H. (1993). Estimating Mediated Effects in Prevention
1006 Studies. *Eval Rev* *17*, 144–158.

1007 Manley, H., Dayan, P., and Diedrichsen, J. (2014). When Money Is Not Enough: Awareness,
1008 Success, and Variability in Motor Learning. *PLoS ONE* *9*, e86580.

1009 Maxwell, J.P., Masters, R.S.W., Kerr, E., and Weedon, E. (2001). The implicit benefit of
1010 learning without errors. *The Quarterly Journal of Experimental Psychology A* *54*, 1049–1068.

1011 McDougle, S.D., and Taylor, J.A. (2019). Dissociable cognitive strategies for sensorimotor
1012 learning. *Nature Communications* *10*.

1013 McNab, F., and Klingberg, T. (2008). Prefrontal cortex and basal ganglia control access to
1014 working memory. *Nature Neuroscience* *11*, 103–107.

1015 Miyake, A., Friedman, N.P., Rettinger, D.A., Shah, P., and Hegarty, M. (2001). How are
1016 visuospatial working memory, executive functioning, and spatial abilities related? A latent-
1017 variable analysis. *J Exp Psychol Gen* *130*, 621–640.

1018 Modchalingam, S., Vachon, C.M., 't Hart, B.M., and Henriques, D.Y.P. (2019). The effects
1019 of awareness of the perturbation during motor adaptation on hand localization. *PLOS ONE*
1020 *14*, 20.

1021 Morehead, J.R., Qasim, S.E., Crossley, M.J., and Ivry, R. (2015). Savings upon Re-Aiming in
1022 Visuomotor Adaptation. *Journal of Neuroscience* 35, 14386–14396.

1023 Morehead, J.R., Taylor, J.A., Parvin, D., and Ivry, R.B. (2017). Characteristics of Implicit
1024 Sensorimotor Adaptation Revealed by Task-irrelevant Clamped Feedback. *Journal of*
1025 *Cognitive Neuroscience* 1–14.

1026 Nakahara, H., and Hikosaka, O. (2012). Learning to represent reward structure: A key to
1027 adapting to complex environments. *Neuroscience Research* 74, 177–183.

1028 Pearson-Fuhrhop, K.M., Minton, B., Acevedo, D., Shahbaba, B., and Cramer, S.C. (2013).
1029 Genetic Variation in the Human Brain Dopamine System Influences Motor Learning and Its
1030 Modulation by L-Dopa. *PLoS ONE* 8, e61197.

1031 Pekny, S.E., Izawa, J., and Shadmehr, R. (2015). Reward-Dependent Modulation of
1032 Movement Variability. *Journal of Neuroscience* 35, 4015–4024.

1033 Peters, M., and Battista, C. (2008). Applications of mental rotation figures of the Shepard and
1034 Metzler type and description of a mental rotation stimulus library. *Brain and Cognition* 66,
1035 260–264.

1036 Quattrocchi, G., Monaco, J., Ho, A., Irmen, F., Strube, W., Ruge, D., Bestmann, S., and
1037 Galea, J.M. (2018). Pharmacological Dopamine Manipulation Does Not Alter Reward-Based
1038 Improvements in Memory Retention during a Visuomotor Adaptation Task. *Eneuro* 5,
1039 ENEURO.0453-17.2018.

1040 Saijo, N., and Gomi, H. (2010). Multiple Motor Learning Strategies in Visuomotor Rotation.
1041 *PLoS ONE* 5, e9399.

1042 Shepard, R., N., and Metzler, J. (1971). Mental Rotation of Three-Dimensional Objects.
1043 *Science* 171, 701–703.

1044 Shmuelof, L., Huang, V.S., Haith, A.M., Delnicki, R.J., Mazzoni, P., and Krakauer, J.W.
1045 (2012). Overcoming Motor “Forgetting” Through Reinforcement Of Learned Actions.
1046 *Journal of Neuroscience* 32, 14617–14621a.

1047 Sidarta, A., van Vugt, F., and Ostry, D.J. (2018). Somatosensory Working Memory in Human
1048 Reinforcement-Based Motor Learning. *Journal of Neurophysiology* 41.

1049 Sobel, M.E. (1986). Some New Results on Indirect Effects and Their Standard Errors in
1050 Covariance Structure Models. *Sociological Methodology* 16, 159.

1051 Suchan, B., Botko, R., Gizewski, E., Forsting, M., and Daum, I. (2006). Neural substrates of
1052 manipulation in visuospatial working memory. *Neuroscience* 139, 351–357.

1053 Sutton, R.S., and Barto, A. (1998). *Reinforcement Learning: An Introduction* (Cambridge,
1054 Mass: A Bradford Book).

1055 Taylor, J.A., and Ivry, R.B. (2011). Flexible Cognitive Strategies during Motor Learning.
1056 *PLoS Computational Biology* 7, e1001096.

1057 Taylor, J.A., and Ivry, R.B. (2014). Cerebellar and Prefrontal Cortex Contributions to
1058 Adaptation, Strategies, and Reinforcement Learning. In *Progress in Brain Research*,
1059 (Elsevier), pp. 217–253.

1060 The 1000 Genomes Project Consortium, Gibbs, R.A., Boerwinkle, E., Doddapaneni, H., Han,
1061 Y., Korchina, V., Kovar, C., Lee, S., Muzny, D., Reid, J.G., et al. (2015). A global reference
1062 for human genetic variation. *Nature* 526, 68–74.

1063 Therrien, A.S., Wolpert, D.M., and Bastian, A.J. (2016). Effective reinforcement learning
1064 following cerebellar damage requires a balance between exploration and motor noise. *Brain*
1065 *139*, 101–114.

1066 Therrien, A.S., Wolpert, D.M., and Bastian, A.J. (2018). Increasing Motor Noise Impairs
1067 Reinforcement Learning in Healthy Individuals. *Eneuro* *5*, ENEURO.0050-18.2018.

1068 Tibshirani, R. (1996). Regression Shrinkage and Selection via the Lasso. *Journal of the Royal*
1069 *Statistical Society. Series B (Methodological)* *58*, 267–288.

1070 Tseng, Y. -w., Diedrichsen, J., Krakauer, J.W., Shadmehr, R., and Bastian, A.J. (2007).
1071 Sensory Prediction Errors Drive Cerebellum-Dependent Adaptation of Reaching. *Journal of*
1072 *Neurophysiology* *98*, 54–62.

1073 Vandevorde, K., and Orban de Xivry, J.-J. (2019). Internal model recalibration does not
1074 deteriorate with age while motor adaptation does. *Neurobiology of Aging* *80*, 138–153.

1075 Vindras, P., Desmurget, M., Prablanc, C., and Viviani, P. (1998). Pointing Errors Reflect
1076 Biases in the Perception of the Initial Hand Position. *Journal of Neurophysiology* *79*, 3290–
1077 3294.

1078 Werner, S., van Aken, B.C., Hulst, T., Frens, M.A., van der Geest, J.N., Strüder, H.K., and
1079 Donchin, O. (2015). Awareness of Sensorimotor Adaptation to Visual Rotations of Different
1080 Size. *PLOS ONE* *10*, e0123321.

1081 Wong, A.L., Marvel, C.L., Taylor, J.A., and Krakauer, J.W. (2019). Can patients with
1082 cerebellar disease switch learning mechanisms to reduce their adaptation deficits? *Brain* *142*,
1083 662–673.

1084 Yuan, M., and Lin, Y. (2006). Model selection and estimation in regression with grouped
1085 variables. *J Royal Statistical Soc B* 68, 49–67.

1086