

The contribution of explicit processes to reinforcement-based motor learning

Holland, Peter James; Codol, Olivier; Galea, Joseph M.

DOI:

[10.1152/jn.00901.2017](https://doi.org/10.1152/jn.00901.2017)

License:

None: All rights reserved

Document Version

Peer reviewed version

Citation for published version (Harvard):

Holland, PJ, Codol, O & Galea, JM 2018, 'The contribution of explicit processes to reinforcement-based motor learning', *Journal of Neurophysiology*. <https://doi.org/10.1152/jn.00901.2017>

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

Publisher Rights Statement:

Published in *Journal of Neurophysiology* on 14/03/2018

DOI: 10.1152/jn.00901.2017

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 *The contribution of explicit processes to reinforcement-based motor*
2 *learning*

3 **Peter Holland^{1*}, Olivier Codol¹ & Joseph M Galea¹**

4
5 ¹School of Psychology, University of Birmingham, UK

6 *Corresponding Author

7
8 **Author Contributions**

9 PH and JM designed the experiments. PH collected the data. PH and OC analyzed the data. PH, OC
10 and JM drafted the manuscript

11
12 **Running Head**

13 Explicit contribution to reinforcement-based motor learning

14
15 **Correspondence**

16 Peter Holland

17 School of Psychology

18 University of Birmingham, UK

19 Email: p.j.holland@bham.ac.uk

20 Phone: +44 (0) 121 414 2861

28 **Abstract**

29 Despite increasing interest in the role of reward in motor learning, the underlying mechanisms remain
30 ill-defined. In particular, the contribution of explicit processes to reward-based motor learning is
31 unclear. To address this, we examined subject's (n=30) ability to learn to compensate for a gradually
32 introduced 25° visuomotor rotation with only reward-based feedback (binary success/failure). Only
33 two-thirds of subjects (n=20) were successful at the maximum angle. The remaining subjects initially
34 followed the rotation but after a variable number of trials began to reach at an insufficiently large
35 angle and subsequently returned to near baseline performance (n=10). Furthermore, those that were
36 successful accomplished this largely via a large explicit component, evidenced by a reduction in reach
37 angle when asked to remove any strategy they employed. However, both groups displayed a small
38 degree of remaining retention even after the removal of this explicit component. All subjects made
39 greater and more variable changes in reach angle following incorrect (unrewarded) trials. However,
40 subjects who failed to learn showed decreased sensitivity to errors, even in the initial period in which
41 they followed the rotation, a pattern previously found in Parkinsonian patients. In a second
42 experiment, the addition of a secondary mental rotation task completely abolished learning (n=10),
43 whilst a control group replicated the results of the first experiment (n=10). These results emphasize a
44 pivotal role of explicit processes during reinforcement-based motor learning, and the susceptibility of
45 this form of learning to disruption has important implications for its potential therapeutic benefits.

46

47 **Keywords**

48 Motor Learning, Reward, Strategies, Visuomotor Adaptation

49

50 **New & Noteworthy**

51 We demonstrate that learning a visuomotor rotation with only reward-based feedback is principally
52 accomplished via the development of a large explicit component. Furthermore, this form of learning is
53 susceptible to disruption with a secondary task. The results suggest that future experiments utilizing
54 reward-based feedback should aim to dissect the roles of implicit and explicit reinforcement learning

55 systems. Therapeutic motor learning approaches based on reward should be aware of the sensitivity to
56 disruption.

57

58 **Introduction**

59 The motor system's ability to adapt to changes in the environment is essential for maintaining
60 accurate movements (Tseng et al., 2007). Such adaptive behavior is thought to involve several distinct
61 learning systems (Haith and Krakauer, 2013; Izawa and Shadmehr, 2011; Smith et al., 2006). For
62 example, the two-state model proposed by Smith et al. (2006) has been able to explain a range of
63 results in force-field adaptation paradigms in which a force is applied to perturb a reaching
64 movement. The model states that learning is accomplished via both 'fast' and 'slow' processes, the
65 'fast' process learns rapidly but has poor retention, whereas the 'slow' process learns more slowly but
66 retains this information over a longer timescale. Subsequently using a visuomotor rotation paradigm,
67 in which the visible direction of a cursor is rotated from the actual direction of hand movement, it has
68 been suggested that the 'fast' process resembles explicit re-aiming whereas the 'slow' process is
69 implicit (McDougle et al., 2015). The implicit aspect may be composed of several different processes
70 (McDougle et al., 2015), the first and most widely researched being cerebellar adaptation (Izawa et
71 al., 2012). However, additional processes such as use-dependent plasticity and reinforcement of
72 actions that lead to task success are required to fully explain experimental findings (Huang et al.,
73 2014). Haith and Krakauer (2013) have proposed a scheme based on these four processes that
74 attempts a synthesis between the principles of motor learning and the distinction between model-
75 based and model-free mechanisms proposed for reinforcement learning and decision-making (Doll et
76 al., 2016).

77

78 The addition of rewarding feedback has proven beneficial in increasing retention of adaptation (Galea
79 et al., 2015; Shmuelof et al., 2012; Therrien et al., 2016) and motor skills (Abe et al., 2011; Chen et
80 al., 2018; Dayan et al., 2014). Findings such as these have generated interest in the possibility that the
81 addition of reward to rehabilitation regimes may improve the length of time that adaptations are

82 maintained after training (Goodman et al., 2014; Quattrocchi et al., 2017; Shmuelof et al., 2012).
83 However, it is still unclear which of the multiple systems mediating motor learning reward may be
84 acting on. Motor learning via purely reward based feedback is also possible and has been applied in
85 two separate forms: binary and graded. Graded point based reward is often based on the distance of
86 the reaching movement from the target and provides information about the magnitude but not the
87 direction of the error (Manley et al., 2014; Nikooyan and Ahmed, 2015). Graded feedback has proved
88 sufficient for learning abrupt rotations (Nikooyan and Ahmed, 2015), however, in certain conditions
89 explicit awareness is required for successful learning (Manley et al., 2014). An alternative method is
90 to only provide binary feedback in which the reward signals task success, such as hitting a target
91 (Izawa and Shadmehr, 2011; Pekny et al., 2015; Therrien et al., 2016). In contrast to graded feedback,
92 only gradually introduced perturbations have successfully been learnt via binary feedback alone (van
93 der Kooij and Overvliet, 2016) and the contribution of explicit processes has yet to be examined.

94
95 In classical visuomotor adaptation, in which full visual feedback of the cursor is available, gradual
96 adaptation is considered to be largely implicit (Galea et al., 2010). However, this may not be the case
97 when only end-point feedback is provided (Saijo and Gomi, 2010). The question remains as to
98 whether learning a gradually introduced visuomotor rotation based on binary feedback also mainly
99 involves implicit processes. Various methods (Huberdeau et al., 2015) have been used to separate the
100 implicit and explicit components of learning such as asking subjects to verbally report aiming
101 directions (McDougle et al., 2015; Taylor et al., 2014) and forcing subjects to move at reduced
102 reaction times (Haith et al., 2015; Leow et al., 2017). In the current paradigm, we assessed the
103 contribution of explicit processes at the end of the learning period by removing all feedback but
104 asking subjects to maintain their performance. Subsequently, we asked subjects to remove any explicit
105 strategy they may have been using. Such an approach has previously been used to measure the relative
106 implicit and explicit components of adaptation to different sizes of visuomotor rotations (Werner et
107 al., 2015). It is important to note that here we define the explicit component to learning as the amount
108 that participants could remove on request. Such a definition maybe more akin to awareness (Werner et
109 al., 2015) or a form of cognitive control (Cavanagh et al., 2009), rather than an explicit strategy which

110 is often defined as a subject's ability to verbalize the strategy they have employed. Therefore, we do
111 not believe subjects had to be able to verbalize a strategy in order for learning to be defined as
112 explicit.

113

114 Our second approach to investigating the explicit contribution to learning based on binary feedback
115 was the introduction of a dual task in order to divide cognitive load and suppress the use of explicit
116 processes. Dual task designs have previously successfully been employed to disrupt explicit processes
117 in adaptation (Galea et al., 2010; Taylor and Thoroughman, 2007, 2008), sequence learning (Brown
118 and Robertson, 2007) and motor skill learning (Liao and Masters, 2001). Various forms of dual task
119 have been used such as counting auditory stimuli (Maxwell et al., 2001), repeating an auditory
120 stimulus (Galea et al., 2010) or recalling words from a memorized list (Keisler and Shadmehr, 2010).
121 We selected a mental rotation task based on using an electronic library of three-dimensional shapes
122 (Peters and Battista, 2008; Shepard and Metzler, 1971). This particular task was selected in order to
123 maximize the likelihood of interfering with the explicit re-aiming process. Indeed, it has previously
124 been shown that both spatial working memory and mental rotation ability correlate with performance
125 in the early 'fast' phase of adaptation (Anguera et al., 2009; Christou et al., 2016). Additionally,
126 depletion of spatial working memory resources prior to visuomotor adaptation is detrimental to
127 performance in the early phase (Anguera et al., 2012). Furthermore, the same prefrontal regions are
128 activated during the early phase of adaptation and during the performance of a mental rotation task
129 (Anguera et al., 2009). It has also been suggested that the explicit process of re-aiming in response to
130 visuomotor rotations may involve a mental rotation of the required movement direction
131 (Georgopoulos and Massey, 1987)

132

133 If the learning of a gradually introduced rotation via binary feedback is dominated by explicit
134 processes, this should be evidenced by a large change in performance when subjects are asked to
135 remove any strategy. Furthermore, the dual task should severely disrupt learning and could possibly
136 unmask any implicit process.

137

138 **Materials and Methods**

139 *Subjects*

140 Sixty healthy volunteers aged between 18 and 35 participated in the study. Forty subjects (thirty-seven
141 females, mean age = 19.9 years) completed experiment 1 and twenty (fifteen females, mean age =
142 21.6 years) in experiment 2. The number of subjects was selected to match the group size that is
143 commonly employed within the field of motor learning (Morehead et al., 2017; Shmuelof et al., 2012;
144 Therrien et al., 2016) and was not based on *a priori* power analysis. All subjects were right-handed
145 with no history of neurological or motor impairment and had normal or corrected-normal vision.
146 Volunteers were recruited from the undergraduate pool in the School of Psychology and wider student
147 population at the University of Birmingham and all gave written informed consent. Subjects were
148 remunerated with their choice of either course credits or money (£7.50/hour). The study was approved
149 by the local ethics committee of the University of Birmingham and performed in accordance with
150 those guidelines.

151

152 *Experimental Protocol*

153 A similar paradigm has previously been employed and the current protocol was designed to replicate
154 this as closely as possible (Therrien et al., 2016). In addition to the rotation of 15°, we extended this
155 paradigm to a 25° rotation. Subjects performed reaching movements with their right arm using a
156 KINARM (B-KIN Technologies), Figure 1A. Subjects were seated in front of a horizontally placed
157 mirror that reflected the visual stimuli presented on a screen above (60 Hz refresh rate). Reaching
158 movements were performed in the horizontal plane whilst subjects held the handle of a robotic
159 manipulandum, with the arm hidden from view by the mirror.

160

161 *Experiment 1*

162 Two different paradigms were employed in Experiment 1, both consisted of a gradually introduced
163 rotation of the required angle of reach for a trial to be considered successful. The maximal extent of
164 the rotation was either 15° (n=10) or 25° (n=30). The motivation for the use of the two different

165 magnitudes of rotation was first to replicate the results of Therrien et al. (2016) and subsequently to
166 investigate if subjects could successfully adapt to a larger angle (25°) than previously employed in
167 binary feedback based motor learning. Subjects were required to learn the rotation on the basis of only
168 binary feedback indicating if they had successfully hit the target region. After the rotation had reached
169 the maximal extent, all feedback was extinguished and two further blocks of trials were performed to
170 assay the level of retention and to what extent this was explicit in nature.

171

172 A total of 470 or 670 trials were performed for the 15° and 25° paradigms, respectively. Each trial
173 followed an identical sequence. Initially a starting position was displayed on screen (red colored
174 circle, 1cm radius), after subjects had moved the position of the cursor (white circle, 0.5cm radius)
175 into the starting position, the starting position changed color from red to green. After a small delay
176 (randomly generated, 500-700ms), in which subjects had to maintain the position of the cursor within
177 the starting circle, a target (red circle, 1cm radius) appeared directly in front of the starting circle at a
178 distance of 10cm. Subjects were instructed to make rapid ‘shooting’ movements that intercepted a
179 visual target, they were instructed that they did not have to attempt to terminate their movement in the
180 target but pass directly through it (Figure 1B). If the cursor intercepted a ‘reward region’ ($\pm 5.67^\circ$),
181 initially centered on the visible target, the movement was considered successful and the target
182 changed color from red to green and a large (8x8cm) green ‘tick’ was displayed at a distance of 20cm
183 directly in front of the starting position (Figure 1C). However, if the cursor did not intercept the
184 reward region the trial was considered unsuccessful and the visible target disappeared from view.
185 Movement times, defined as the time from leaving the starting circle to reaching a radial distance of
186 10cm, were constrained to a range of 200-1000ms. Movements outside of this range but at the correct
187 angle were counted as incorrect trials and no tick was displayed. As a visual cue, movements outside
188 of the acceptable duration were signaled with a change of the target color, blue for too slow and
189 yellow for too fast. After the completion of a reaching movement the robot returned the handle to the
190 start position and subjects were instructed to passively allow this whilst maintaining their grip on the
191 handle, during the passive movement subjects continued to receive no visual feedback of hand
192 position. Reaction times, defined as the difference in time between the appearance of the target and

193 the time at which the cursor left the starting circle, were limited to a maximum 600ms. If a movement
194 was not initiated before this time, the target disappeared and the next trial began after a small delay
195 and these trials were excluded from further analysis.

196

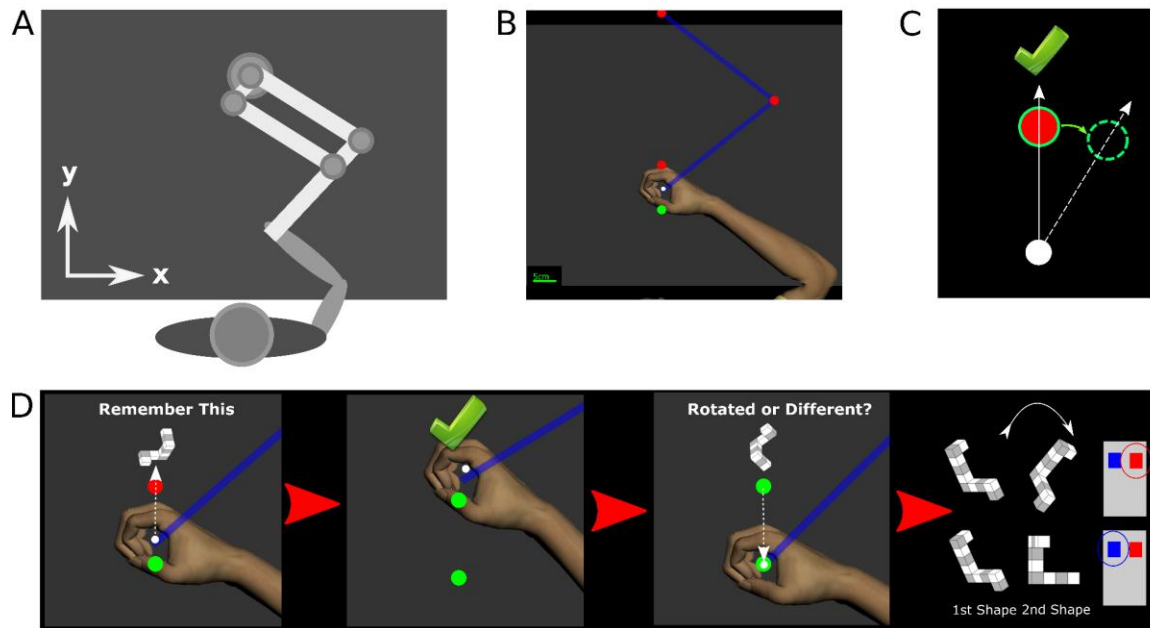
197 After an initial period of ten trials, in which the cursor position was constantly visible, for the
198 remainder of the experiment it was extinguished. The only feedback subjects received was a binary
199 (success/fail) signal indicating if the angle of reach was correct, in the form of a change of target color
200 and the appearance of the tick. For an initial period of forty trials, the reward region remained
201 centered on the position of the visual target, after this it was shifted in steps of 1° every twenty trials.

202 The number of trials within the initial period and the rate of introduction of the rotation were identical
203 in the 15° and 25° paradigms, only the total number of trials required to reach the maximum angle
204 differed. This manipulation ensured that for a reaching movement to be considered correct it must be
205 made at an increasingly rotated angle from the visual target (Figure 1C). Subjects were pseudo-
206 randomly assigned to groups that received either a clockwise or counter-clockwise rotation. Once the
207 reward region had reached the maximal angle, either 15° or 25° , it was held constant for an additional
208 twenty trials. Subsequently, subjects were informed that they would no longer receive any feedback
209 about their performance but that they should continue to perform in the same manner as before; this
210 'Maintain' block consisted of fifty trials. Following this, subjects were asked a series of simple
211 questions to assay their awareness of the rotation; answers were noted by the experimenter. Firstly,
212 subjects were asked 'Did you notice anything change during the course of the experiment?'.
213 Secondly, 'Did you deliberately change anything about how you were performing the task?'. If the
214 answer to the second question was affirmative they were asked a follow-up question 'What did you
215 do?'. Subsequently all subjects were told 'During the task we secretly moved the position of the target
216 that you had to hit. You will still not receive information on whether you hit the target or not but
217 please try to move as you did at the start of the experiment'. Throughout the text we refer to this
218 instruction as being asked to remove any strategy. Crucially subjects were not informed of the
219 direction or magnitude of the rotation they had experienced. The final 'Remove' block consisted of
220 fifty trials.

221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247

In order to test for any effects on retention due to the passage of time it took subjects to respond to the questions we performed a control experiment. The first 570 trials of the experiment were identical to the 25° paradigm previously described. However, at the end of the first block of fifty trials of no visual feedback (Maintain 1 block) subjects were asked to respond verbally to two questions from the BAS reward responsiveness section of the BIS/BAS questionnaire. These questions were selected on the basis of pilot experiments which demonstrated that they took approximately the same length of time to complete as the awareness related questions described previously. After subjects had responded to these questions they performed another block of fifty trials in which they received no feedback but were instructed to continue reaching in the same manner as before (Maintain 2 block). Subsequently, subjects were asked the task awareness questions, those that occurred in between Maintain and Remove blocks in the main experiment. The answers were noted down by the experimenter and subjects were then instructed to remove any strategy they had employed and then completed another fifty trials without visual or binary feedback (Remove block). For this experiment, we recruited an additional ten subjects who were successful in compensating for the final angle of rotation (fifteen in total recruited), the direction of the rotation was counterbalanced between subjects.

The position of the handle throughout the task was recorded at a sampling rate of 1 kHz and saved for offline analysis.



248

249 **Figure 1. Experimental design.** *A*, Subjects held the handle of robotic manipulandum with their right
 250 hand, the position of the arm and handle was hidden from sight and feedback was provided on a
 251 horizontal screen. **B**, Subjects made ‘shooting’ movements from a starting position (green circle)
 252 towards a target (red circle), after the initial practice trials the position of the cursor (white circle)
 253 was no longer visible at any point. **C**, Successful trials were indicated to the subject with the display
 254 of a green tick after the cursor had passed through a region centered on the target, over the course of
 255 the paradigm the position of the reward region gradually moved (solid green circle to dashed green
 256 circle) whilst the visible target (red circle) remained in the central location. By the end of the learning
 257 period a successful reach (dotted white line) was rotated by a maximum of either 15° or 25°. **D**, Time-
 258 course of Experiment 2, at the same time as the target appeared on screen a ‘shape’ was also
 259 displayed slightly above it, the subject was asked to memorize this shape. After the reach was
 260 completed and the hand returned to the starting position subjects used their left hand to respond with
 261 a button press as to whether they believed the new shape shown on screen was a rotated version of the
 262 shape or an entirely different shape.

263

264

265

266 ***Experiment 2***

267 Experiment 2 comprised of the same reaching task as Experiment 1 but with the addition of a mental
268 rotation dual task. The dual task required subjects to hold a three-dimensional shape in working
269 memory for the duration of the reaching movement (Figure 1D). Subjects had to respond with a
270 button press using their left hand to indicate if a shape displayed at the end of the reaching movement
271 was a rotated version of a shape displayed at the time of target presentation or a different shape.

272

273 Shapes had the form of a series of connected cubes, alternately colored grey and white, they were
274 selected from an electronic library designed on the basis of the Shepard and Metzler type stimuli
275 (Peters and Battista, 2008; Shepard and Metzler, 1971). All rotations were performed within the plane
276 of the screen, i.e. although the stimuli represented three-dimensional shapes all rotations were in two-
277 dimensions. A subset of 26 shapes were selected from the library for use in this experiment and are
278 available on <https://osf.io/vwr7c/>. The trial protocol was the same as that employed in Experiment 1
279 but at the time when the target circle appeared, a randomly selected shape from the subset was
280 displayed in an 8x8cm region at a position 20cm away from the starting position. Subjects were
281 instructed to commit this shape to memory. The shape remained visible on screen until the end of the
282 reaching movement, the point at which the radial amplitude of the cursor exceeded 10cm. The shape
283 was then extinguished and the same binary feedback as employed in Experiment 1 was displayed.
284 After the robot had guided the handle back to the starting position a second shape was displayed in the
285 same position as the first. In half of the trials this was an identical shape to the first one but had
286 undergone a rotation selected at random from a uniform distribution of 0-360°, in the other half of
287 trials it was a different shape selected at random from the library. The order of trials in which the
288 shape was either rotated or different was randomized and subjects had a maximum of 2s to respond.
289 Subjects in the Dual Task group (n=10) were instructed to press the right-sided button of two buttons
290 on a button box held in their left hand if they believed the second shape to be a rotated version of the
291 first one and the left-sided button if they believed it was a different shape. Importantly subjects were
292 given no feedback on their performance in the dual task but were informed prior to the experiment
293 that this would be monitored, the responses were recorded and analyzed offline. This design was

294 selected in order to avoid any interfering effects of rewarding feedback from the dual task with the
295 binary feedback in the reaching task. As a control, another group of subjects received identical visual
296 stimuli but were instructed to press a random button of the two on each trial. Subjects were pseudo-
297 randomly assigned to either the Control or Dual Task groups.

298

299 For Experiment 2, the familiarization period at the start of the experiment, in which the position of the
300 cursor was visible, was extended to twenty trials in order for subjects to have sufficient time to
301 acclimatize to the additional timing requirements of the button press. The paradigm subsequently
302 followed that of Experiment 1 with a maximal angular rotation of 25°.

303

304 *Data Analysis*

305 All data analysis was performed with custom written routines in MATLAB (The Mathworks) and
306 extracted data and all code required to reproduce the analysis and figures in this paper are freely
307 available on (<https://osf.io/vwr7c/>).

308

309 The end point angle of each reaching movement was calculated either at the time that the cursor
310 intercepted the reward region or in the case of incorrect trials when the cursor reached a radial
311 amplitude of 10cm. An angle of zero degrees was defined as a movement directly ahead, i.e. toward
312 the visible target position. A positive angle of rotation was defined as a clockwise shift of the reward
313 region, and reach angles and target positions for the counter-clockwise rotation were sign-transformed
314 to positive values for comparability. The ‘Baseline’ period was defined as the first forty trials without
315 visual feedback of the cursor, during which the reward region was centered on the visual target.
316 Subjects were considered to have successfully learnt the rotation if the mean end point angle of the
317 reaching movements fell within the reward region during the last twenty trials before the ‘Maintain’
318 period, a time at which the rotation was held constant at its maximal value.

319

320 During the retention phase of the experiment (last one hundred trials), we calculated the amount of
321 retention that could be accounted for by explicit and implicit processes. A subject’s implicit retention

322 was defined as the difference between the mean reach angle in the final fifty trials ('Remove' blocks),
323 after subjects had been instructed to remove any strategy they had been using, and mean reach angle
324 during the 'Baseline' blocks. A subject's explicit retention was defined as the difference between the
325 mean reach angle during the 'Maintain' blocks, the first fifty trials after removal of binary feedback in
326 which subjects were instructed to continue reaching as before, and the implicit retention.

327

328 In order to analyze the effect of reward on subjects behavior, we conducted trial-by-trial analysis in a
329 manner similar to one that has previously been employed for analysis of reaching performance in
330 response to binary feedback (Pekny et al., 2015). The change in reach angle following trial n , $\Delta u^{(n)}$,
331 was defined as the difference between consecutive trials:

332

$$\Delta u^{(n)} = u^{(n+1)} - u^n$$

333

334 Subsequently we examined the distributions of Δu following only rewarded (correct) or unrewarded
335 (wrong) trials. The resulting distributions of Δu were non-normal and therefore we analyzed and
336 report the median and median absolute deviation (MAD) of each subject's distributions. We also
337 examined the absolute change in reach angle $|\Delta u|$, i.e. the magnitude of change regardless of
338 direction.

339

340 In order to investigate the effects of a reward history spanning multiple trials we examined the $|\Delta u|$
341 following all possible combinations of success in the previous three trials. We first searched each
342 subject's responses for the occurrence of all eight possible sequences of reward and calculated the
343 mean change in reach angle following each. We then quantified this behavior using a model in which
344 $|\Delta u|$ was a function of the outcome of the previous three trials as well as variability (ε) that could not
345 be accounted for by the recent outcomes (Pekny et al., 2015):

346

$$|u(n)| = \alpha_0(1 - R(n)) + \alpha_1(1 - R(n - 1)) + \alpha_2(1 - R(n - 2)) + \varepsilon$$

347

348 In the above equation, R represents the presence of reward on a given trial with a value of 1 for a
349 correct trial, $R(n)$ therefore represents the presence of reward on the previous trial with $R(n - 1)$ and
350 $R(n - 2)$ the preceding two trials. The components α_0 , α_1 and α_2 represent the sensitivity to the
351 outcomes of these trials with higher values indicating subjects made larger changes in response to the
352 outcome of that trial. The values of these components were estimated using the least squares error
353 solution to the equation using the mean value of $|\Delta u|$ recorded for each sequence on a subject-by-
354 subject basis. We repeated this analysis using $|\Delta u|$ of every occurrence of a sequence (i.e. trial-by-
355 trial analysis rather than using a mean value), and obtained similar estimates for the components. The
356 model fits for both methods are reported as R^2 values in the results section.

357

358 The verbal responses to the questions asked before the start of the 'Remove' block was noted down by
359 the experimenter and analyzed offline. A subject's awareness of the perturbation and efforts to
360 deliberately counter it were rated on a scale of 0, 0.5 and 1, with 0 indicating no awareness and 1
361 indicating full awareness, including deliberately aiming at a rotated angle. A score of 0.5 was given
362 when subjects were aware of some change throughout the course of the experiment but could not
363 accurately state the nature of the perturbation or what they changed about their movement to counter
364 it.

365

366 *Statistical Analysis*

367 Statistical analysis was performed in MATLAB. In order to test for initial effects mixed design
368 ANOVAs were used, with Group (25RotSucces, 25RotFail etc.) as the between-subjects factor and
369 time-point (Baseline, 15° Block, Maintain etc.) or MeasuredVariable (Median Δu , Reward
370 Component etc.) as the within-subjects factor. The Greenhouse-Geiser correction was applied in cases
371 of violation of sphericity and corrected p-values and degrees of freedom are reported in the text. In
372 cases in which a significant interaction was found in the ANOVA, post-hoc tests were performed to
373 test for differences between groups at each TimePoint or MeasuredVariable. As data was often found

374 to be non-normally distributed using Kolmogorov-Smirnov tests, the non-parametric Kruskal-Wallis
375 test was applied throughout. In cases of a significant effect of group on an individual outcome
376 measure, further pairwise comparisons of mean group ranks were employed and Bonferroni corrected
377 p-values are reported in the text. For tests of a difference of a single group from zero, such as in
378 testing for implicit learning, Wilcoxon-Signed Rank tests were employed and Bonferroni corrected p-
379 values are reported in the text. A critical significance level of $\alpha=0.05$ was used to determine statistical
380 significance. The probability density estimates displayed as shaded regions in distribution plot figures
381 were estimated using a Gaussian kernel.

382

383 **Results**

384 *Experiment 1: Successfully learning to compensate for a 25° rotation includes a large explicit* 385 *component*

386 We first sought to investigate the size of a gradual introduced visuomotor rotation that subjects can
387 learn based on binary feedback. All subjects who experienced the 15° rotation (15Rot group) learnt to
388 fully compensate (Figure 2A). Successful compensation was defined as having a mean reach angle
389 within the reward region in the final twenty trials before the retention phase. However, for the 25°
390 group (25Rot, magenta group, Figure 2B), the average reach direction fell outside the reward region,
391 indicating incomplete learning. Underlying the mean performance was a split in behavior: some
392 subjects successfully learnt the full rotation, whereas one third of subjects did not. On the basis of this
393 behavior, they were categorized into two subgroups: 25RotSuccess (red group, N=20) and 25RotFail
394 (blue group, N=10), respectively.

395

396 Next, we compared reach angle for the three groups (15Rot, 25RotSuccess and 25RotFail) at specific
397 time points in order to gain an understanding at which stage the difference emerged (Figure 2C, D).
398 Despite no difference between groups at baseline ($H(2) = 4.03$, $p = 0.13$, Kruskal Wallis), a difference
399 had emerged at 15° ($H(2) = 9.63$, $p = 0.008$; Figure 2C). Specifically, reach angle for the 25RotFail
400 group was lower than both the 15Rot ($p = 0.022$) and the 25RotSuccess groups ($p = 0.014$). During

401 the 'Maintain' phase, when binary feedback had been removed but subjects were instructed to
402 continue reaching as before, there was a significant effect of group ($H(2) = 20.08$, $p < 0.001$; Figure
403 2B, C). Unsurprisingly, the 25RotSuccess group was greater than the 15Rot ($p = 0.002$) and the
404 25RotFail groups ($p < 0.001$). Crucially, after subjects were instructed to remove any strategy and
405 reach as they did at the beginning of the experiment, there was no difference between the groups
406 ($H(2) = 0.78$, $p = 0.68$; Figure 2B, C). Analysis of the reach angles during the paradigm revealed that
407 even at a rotation of 15° there was divergence between the 25RotFail and 25RotSuccess groups.
408 Furthermore, the instruction to remove any strategy resulted in a return to a similar level of
409 performance across all three groups.

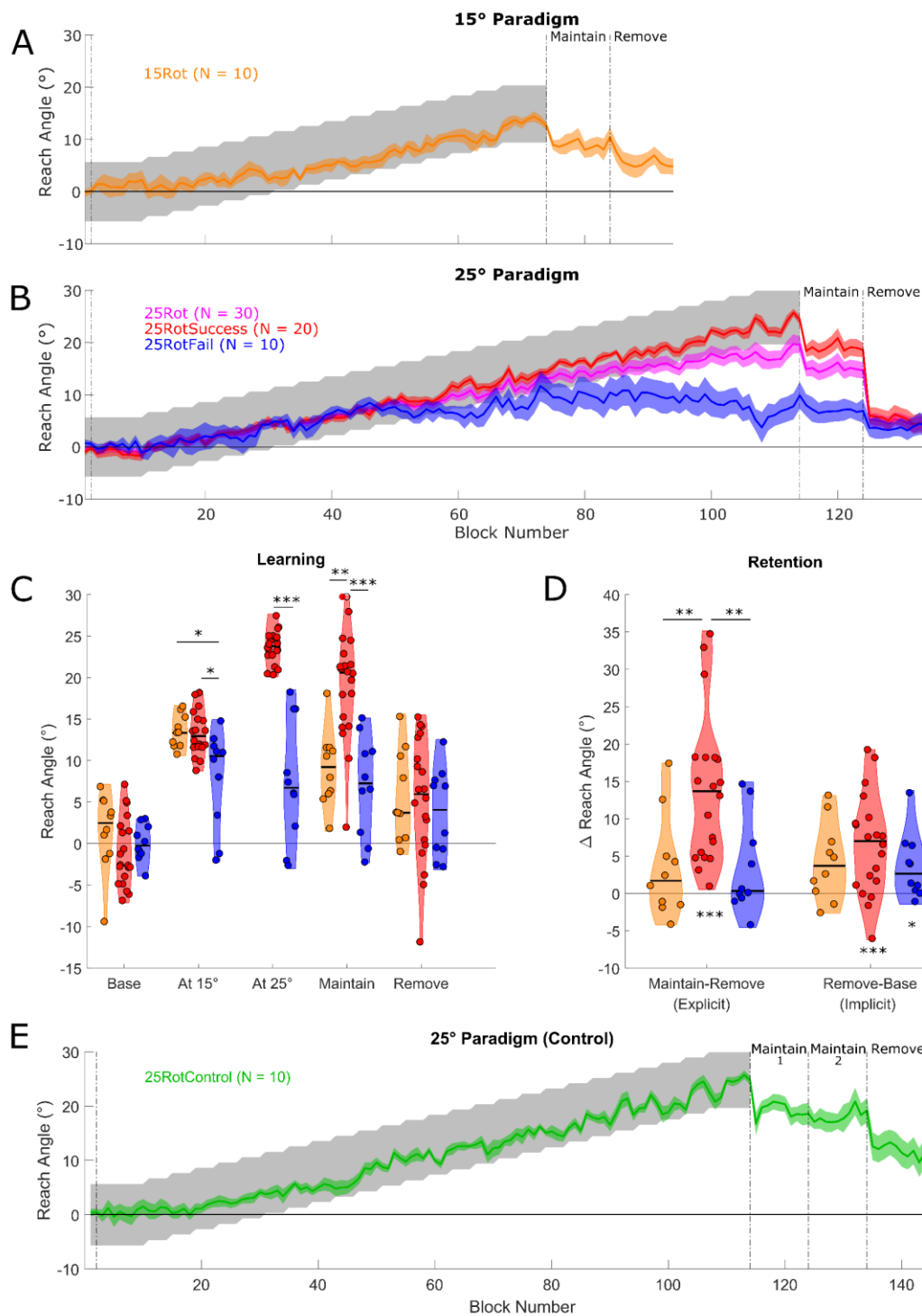
410

411 We probed the nature of learning by calculating the implicit and explicit components of retention
412 (Figure 2D). Implicit retention reflected the retention after removal of any strategies, whereas Explicit
413 retention represented the change in behavior accounted for by the removal of strategies. The Explicit
414 component of the 25RotSuccess group was greater than both 15Rot ($p = 0.006$) and 25RotFail ($p =$
415 0.006). Furthermore, only the 25RotSuccess ($Z = 210$, $p < 0.001$) group had a significant Explicit
416 component to their retention. Whilst there was no effect of Group on the Implicit component ($H(2) =$
417 1.84 , $p = 0.40$), both groups in the 25° paradigm showed a significant difference from 0
418 (25RotSuccess, $Z = 193$, $p = 0.001$; 25RotFail, $Z = 48$, $p = 0.014$), however, the 15Rot group was no
419 longer significant after correction for multiple comparisons ($Z = 48$, uncorrected $p = 0.037$, corrected
420 $p = 0.111$). Therefore, whilst all three groups showed a similar small level of implicit retention, only
421 the subjects who successfully learnt the 25° rotation showed evidence for explicit learning. Whilst at a
422 group level there was no evidence for an explicit component to retention in either the 15Rot or
423 25RotFail groups, there was variability within the groups with 2 subjects in each group displaying
424 Explicit components greater than 10° .

425
426

427 It is possible that the reduction in reach angle observed between the 'Maintain' and 'Remove' blocks
428 in the 25RotSuccess group could be accounted for by the decay of a labile memory during the time in

429 which the awareness questions were asked (Smith et al., 2006). In the 25Rot paradigm, the time
430 between the end of the ‘Maintain’ block and the start of the ‘Remove’ block was 37.16 ± 8.49 s. The
431 time taken for the two control questions between the ‘Maintain 1’ and ‘Maintain 2’ blocks for the ten
432 subjects in 25RotControl group was 49.48 ± 8.63 s, and for the awareness questions and instruction to
433 remove strategy between ‘Maintain 2’ and ‘Remove’ was 45.80 ± 13.38 s. There was no significant
434 difference between the length of time taken for either set of questions in the 25RotControl group and
435 those in the 25Rot group ($H(2) = 5.47$, $p = 0.065$; Figure 2E). Crucially, we observed no difference in
436 reach angle between ‘Maintain 1’ and ‘Maintain 2’ ($Z=36$, $p=0.432$). However, there was a clear
437 reduction in reach angle following the instruction to remove any strategy between ‘Maintain 2’ and
438 ‘Remove’ ($Z=52$, $p=0.010$). These results indicate that the passage of time is not the critical factor
439 causing the reduction in reach angle observed, but rather it is the instruction to remove any strategy
440 subjects had employed.



441

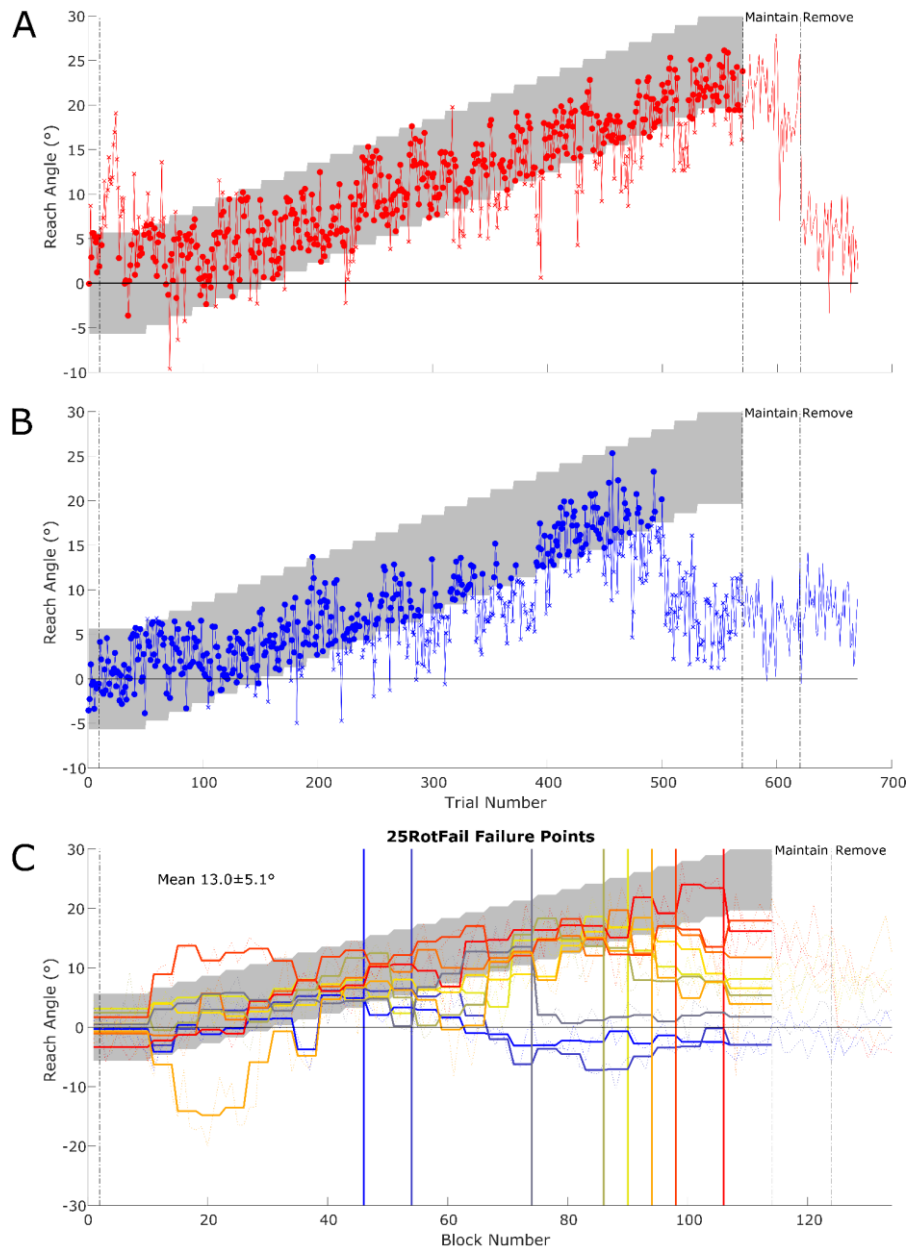
442 **Figure 2. Experiment 1: group performance.** *A*, Reach angle averaged over blocks of 5 trials, solid
 443 colored lines represent the mean of each group and the shaded region represents SEM. The average
 444 behavior of subjects in the 15Rot paradigm (Orange) fell consistently within the rewarded region
 445 (grey shaded region) indicating successful learning. *B*, Average reach angle over blocks for all
 446 subjects in the 25Rot paradigm (magenta) and also the same subjects split into two groups based on

447 *success at the final angle (25RotSuccess – red, 25RotFail – blue). C, Distribution plots displaying the*
448 *reach angles for subjects in the three groups at various timepoints throughout the experiment with*
449 *individual data points overlaid on an estimate of the distribution. Horizontal black line in the*
450 *distribution represents the group median. D, Distribution plots of the computed variables of Implicit*
451 *(‘Remove-Baseline’) and Explicit (‘Maintain-Implicit’) retention. Significance stars above horizontal*
452 *black bars indicate differences between the groups (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$).*
453 *Significance stars below the distributions represent a significant difference from zero. E, Reach angle*
454 *averaged over blocks of 5 trials for subjects in the 25RotControl group. There was no reduction in*
455 *reach angle during the time taken for the control questions between Maintain 1 and Maintain 2*
456 *blocks. However, when subjects were subsequently asked to remove their strategy, the period between*
457 *Maintain 2 and Remove blocks, a significant reduction in reach angle was observed.*

458

459 In order to understand the mechanism of learning, and how this might differ between the
460 25RotSuccess and 25RotFail groups, we examined trial-by-trial behavior. Two distinct types of
461 behavior were apparent (Figure 3). Behavior in those that failed (Figure 3B) was initially similar to
462 successful subjects (Figure 3A), but at some point subjects began to fail to reach at a sufficient angle.
463 Subsequently the angle of reach began to decline over further trials, despite a continued lack of
464 reward. However, given the length of the paradigm it unclear if this reduction was limited to the angle
465 of the last successful trial they experienced or would have continued to baseline levels given more
466 trials. The angles at which subjects in the 25RotFail group failed varied (mean=13.0±5.1°), but all
467 displayed the same pattern of return to baseline (Figure 3C). Given the apparently similar behavior in
468 the initial learning stage, it is important to know whether there are differences even at this early stage.
469 To this end, we only included trials in the initial successful period for the 25RotFail group in all
470 subsequent analysis of trial-by-trial behavior, i.e. trials on the left-hand side of the vertical colored
471 line for each subject (Figure 3C). For the 25RotSuccess and 15Rot groups all trials during the learning
472 period were analyzed. Crucially, there was no difference in the percentage of correct trials within this
473 period between the groups ($H(2) = 2.19$, $p = 0.33$).

474



475

476 **Figure 3. Experiment 1: trial-by-trial behavior.** Example of trial by trial reach angles from a subject
 477 who was successful at the final angle (A) and one who was unsuccessful (B). In each case rewarded
 478 trials are indicated with a circular marker and non-rewarded trials with a 'x'. The grey shaded
 479 region indicates the reward region. C, Failure points for subjects in the 25RotFail group, thick lines
 480 are the mean reach angle for each subject at each rotation angle, thin lines represent mean of each
 481 block (average of 5 trials), colors go from hot to cold matching failure angles ranging from high to
 482 low. Vertical lines represent the last angle at which mean reach fell within rewarded region for each
 483 subject. The mean and standard deviation of all angles of failure is displayed as text.

484

485 Next, we examined if changes in reach angle were affected by the outcome of the previous trial. A
486 similar analysis has been employed previously (Pekny et al., 2015). We examined the distributions of
487 Δu following only rewarded (Correct) or unrewarded (Wrong) trials. The resulting distributions of Δu
488 were non-normal and therefore we report the median and median absolute deviation from the median
489 (MAD). Whilst the median Δu was greater following unrewarded trials ($F(1,37) = 119.80$, $p < 0.001$;
490 Figure 4A), this effect was similar across groups ($F(2,37) = 1.18$, $p = 0.64$). Similarly, the MAD of Δu
491 was also greater following Wrong trials, indicating that not only did all groups make larger changes in
492 reach angle but also that there was greater variability in these changes (Figure 4B). Despite a
493 significant interaction with Group ($F(2,37) = 5.32$, $p = 0.019$), the trend for a higher MAD of Δu
494 following Wrong trials for the 25RotSuccess group (Figure 4B) did not reach significance after
495 correction for multiple comparisons ($H(2) = 5.63$, $p = 0.06$). Subsequently we repeated the analysis
496 but considered the absolute change in reach angle ($|\Delta u|$, Figure 4C, D). Here there was a significant
497 interaction with Group for both median $|\Delta u|$ ($F(2,37) = 7.89$, $p = 0.003$) and MAD of $|\Delta u|$ ($F(2,37) =$
498 7.39 , $p = 0.004$) following Wrong trials. Post-hoc tests revealed that the 25RotSuccess group
499 displayed a significantly greater median $|\Delta u|$ ($p = 0.024$) and MAD of $|\Delta u|$ ($p = 0.035$) than the
500 25RotFail group. There was no difference between the groups in the magnitude or variability of the
501 change in reach angle after correct trials. The analysis of the absolute changes in reach angle reveal
502 that even during the period in which they are successful, the 25RotFail group made smaller and less
503 variable changes following unrewarded trials.

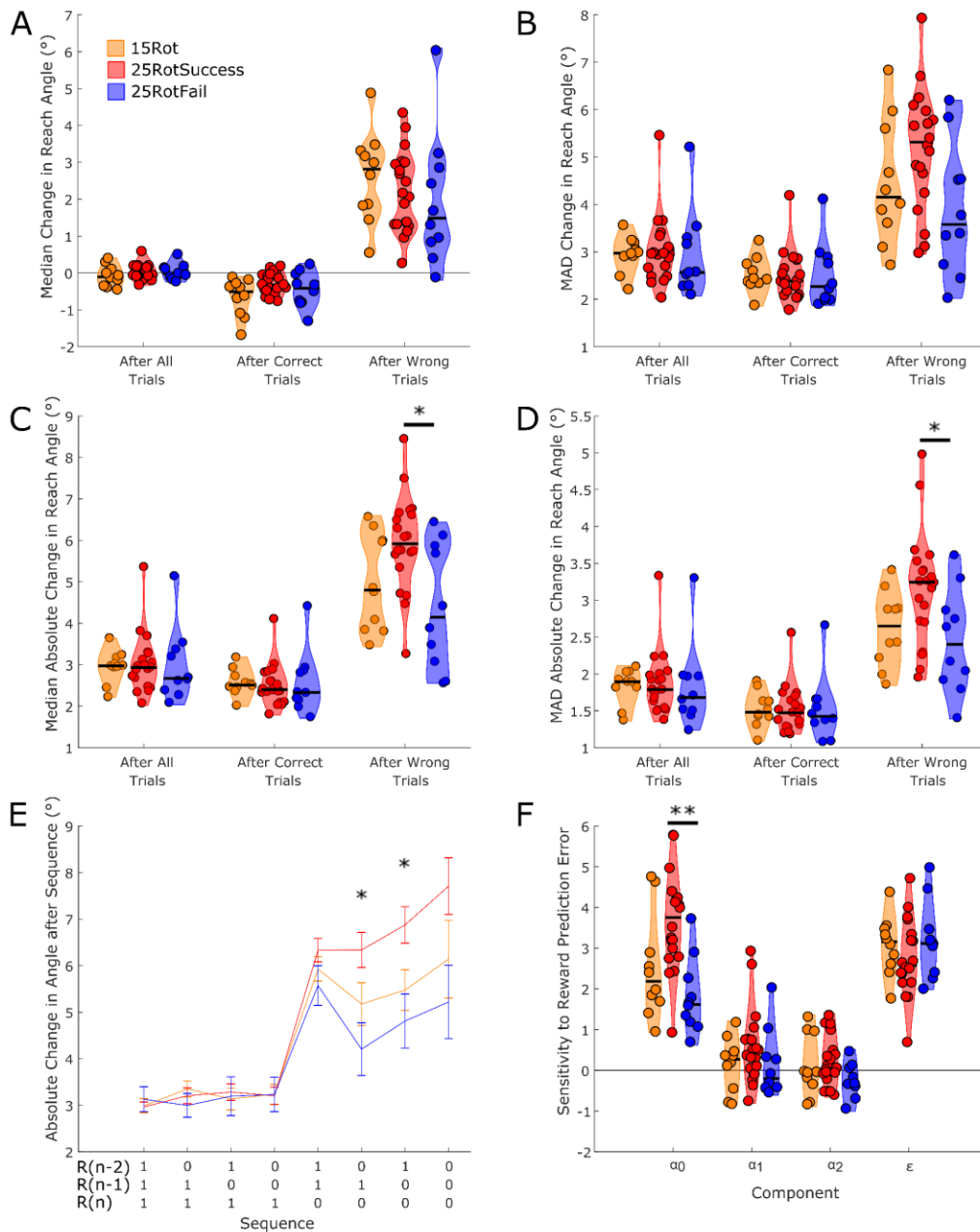
504

505 In addition to the effect of the previous trial it is possible that subjects were sensitive to a history of
506 outcomes spanning multiple previous trials (Pekny et al., 2015). In order to investigate the effects of
507 reward history, we examined the $|\Delta u|$ following all possible combinations of success in the previous
508 three trials (Figure 4E). We quantified this behavior using a model in which $|\Delta u|$ was a function of
509 the outcome of the previous three trials. The components α_0 , α_1 and α_2 represent the sensitivity to the
510 outcome of the last three trials with α_0 being the most recent (Figure 4F), ε represents variability that

511 could not be accounted for by the recent outcomes. There was an interaction between component and
512 group ($F(3.49,64.51) = 4.49, p = 0.004$). All groups were most sensitive to the most recent trial
513 outcome (α_0) with the 25RotSuccess group displaying significantly greater change than 25RotFail (p
514 $= 0.001$). There was no difference between groups for other components indicating that differences in
515 behavior were driven by the sensitivity to the outcome of the most recent trial. R^2 values for model
516 fits based on the mean $|\Delta u|$ of each sequence had a mean of 0.90 and a range of 0.67 to 0.99, model
517 fits based on a trial by trial basis had a mean R^2 of 0.39 and a range of 0.15 to 0.57. From these results
518 it becomes apparent that, even in the initial period of success, subjects who will go on to fail to learn
519 the full rotation show a decreased sensitivity to errors.

520

521 There was no difference between groups for either movement time ($H(2) = 4.82, p = 0.090$) or
522 reaction time ($H(2) = 4.01, p = 0.13$). The mean of the median movement times across subjects was
523 0.38 ± 0.08 s. Additionally, within the 25RotFail group reaction and movement times did not differ
524 before and after the point of failure ($Z = 28, p = 1$ and $Z = 40, p = 0.23$ respectively). In response to
525 the questions asked to probe awareness, we found no significant difference between the groups ($\chi^2(2)$
526 $= 3.75, p = 0.15$). However, within the 25RotSuccess group there was a significantly non-uniform
527 distribution of answers ($\chi^2(2) = 9.1, p = 0.005$) with 60% of participants reporting a specific strategy
528 to counter the rotation and only one reporting not to notice any change. The remainder of subjects
529 reported some awareness of a change (categorized as 0.5 on our scale), or an explicit effort to counter
530 it, but were often not confident in describing the change or could not easily verbalize their strategy.
531 There was no difference between the subjects reporting full or partial awareness in terms of the
532 quantified Explicit component to retention ($Z = 123, p = 0.837$).



533

534 **Figure 4. Experiment 1: performance after correct and incorrect trials.** Analysis of the effects of the
 535 success of the previous trial and reward history on trial by trial changes in reach angle for the three
 536 groups in Experiment 1 (15Rot – Orange, 25RotSuccess – Red, 25RotFail – Blue). Median (A) and
 537 MAD (B) of change in reach angle separated by the success of the previous trial. Median (C) and
 538 MAD (D) of the absolute change in reach angle separated by the success of the previous trial. E, The
 539 absolute change in reach angle following all combinations of trial success over the previous three
 540 trials. F, Sensitivity to the outcomes of each of the previous trials. Significance stars above horizontal
 541 black bars indicate differences between the groups (* $P < 0.05$, ** $P < 0.01$).

542

543 ***Experiment 2: Addition of a dual task prevents learning***

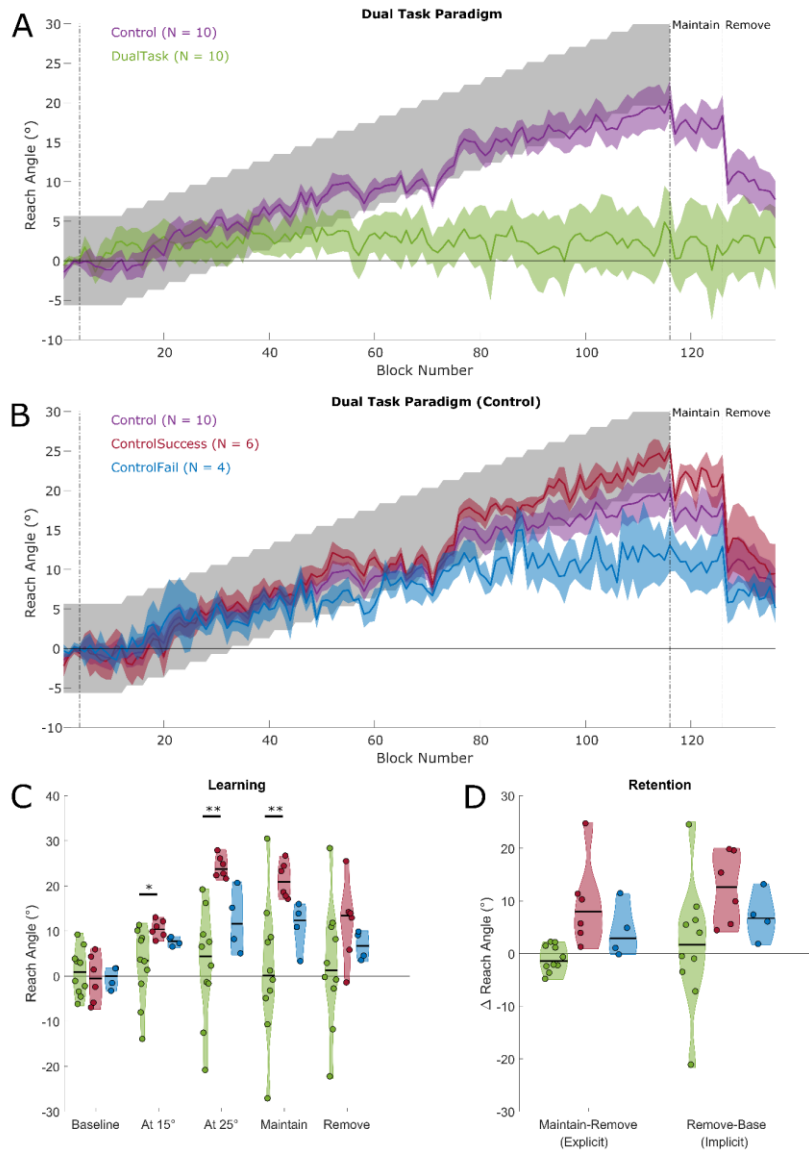
544 Following the finding of Experiment 1 that successful reinforcement-based motor learning involves a
545 strong explicit component, we sought to investigate if it was possible to disrupt learning by dividing
546 cognitive load. To this end, we required subjects to hold a shape in memory during the period of
547 movement (Figure 1D).

548

549 The DualTask (N=10) group displayed little learning and none successfully compensated for the
550 maximum rotation (Green group, Figure 5A). As in Experiment 1, the Control (N=10) group on
551 average fell short of complete learning (Purple group, Figure 5A, B), indicated by the mean reach
552 direction falling outside the reward region in the final learning blocks. However, the average of the
553 group obscures a similar split in behavior with only six subjects successfully learning the full rotation
554 and four failing to do so, which we will label (ControlSuccess and ControlFail, respectively; Figure
555 5B).

556

557 Examining performance in the same time periods as Experiment 1 (Figure 5C) revealed no difference
558 between the three groups at baseline ($H(2) = 0.38$, $p = 0.83$). However, by the time the angle of
559 rotation had increased to 15° a significant difference had already emerged ($H(2) = 6.88$, $p = 0.03$),
560 with the DualTask group displaying lower reach angle than ControlSuccess ($p = 0.011$).



561

562 **Figure 5. Experiment 2: group performance.** Change in reach angle over blocks (average of 5 trials)

563 during the dual task experiment. **A**, Group performance for the DualTask (Green) and Control

564 (Purple) task groups, the line indicates the mean and shaded region the SEM. The grey shaded region

565 represents the reward region. **B**, the split of the control task group into ControlSuccess (Dark Red)

566 and ControlFail (Blue). **C**, Distribution plots displaying the performance at different time points for

567 the dual task, and split control groups. The shaded region represents an estimation of the distribution

568 and is overlaid with data for each individual subject. **D**, Distribution plots of the difference in reach

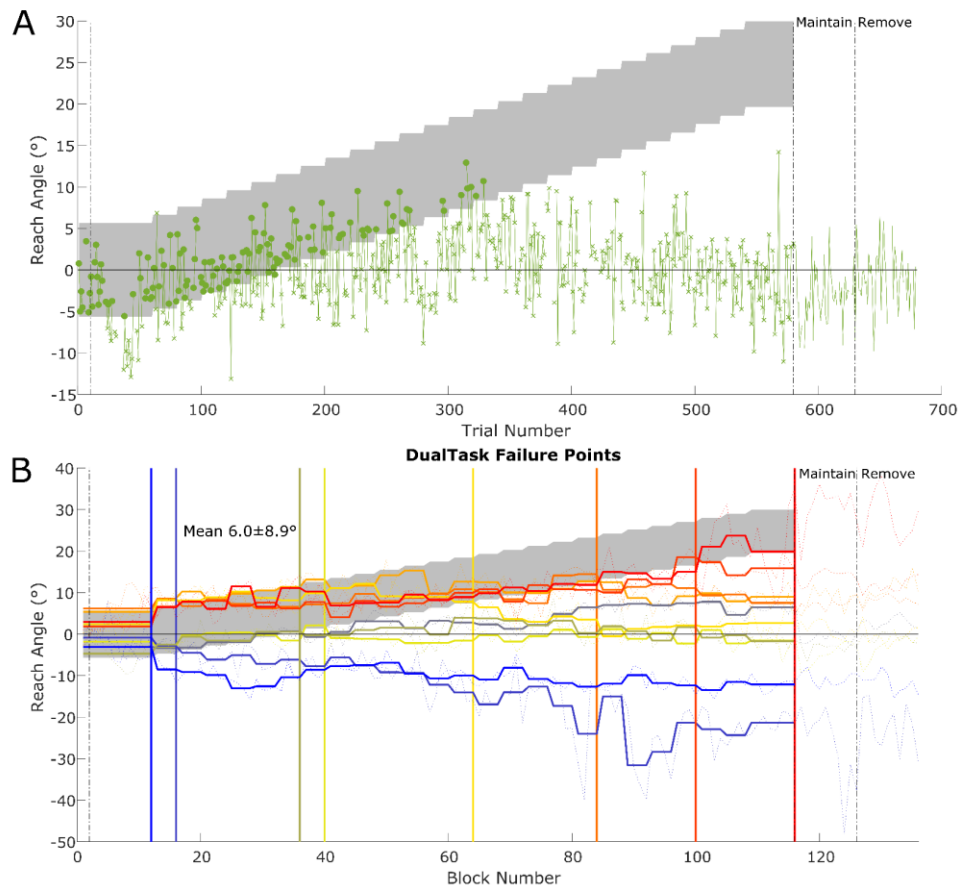
569 angle during retention phases indicating the implicit and explicit components of retention.

570 Significance stars above horizontal black bars indicate differences between the groups (* $P < 0.05$,

571 ** $P < 0.01$).

572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587

As can be seen from the performance of individuals in the DualTask group (Figure 6), there were very few correct trials (mean angle of failure 6.0°) rendering the analysis of trials within the successful period employed for Experiment 1 invalid. Despite this limitation for the DualTask group, the analysis could still elucidate differences between the ControlSuccess and ControlFail groups and reassuringly the mean angle of failure in ControlFail group is 13° , similar to Experiment 1. However, the small group numbers preclude statistical comparison between the ControlSuccess and ControlFail groups but the pattern of behavior was visually similar to that in Experiment 1 (Figure 7). Overall the analysis of sensitivity to reward history produced remarkably similar results to Experiment 1 with the primary difference between those who learn and those who fail to do so being the sensitivity to the outcome of the most recent trial (Figure 7F).



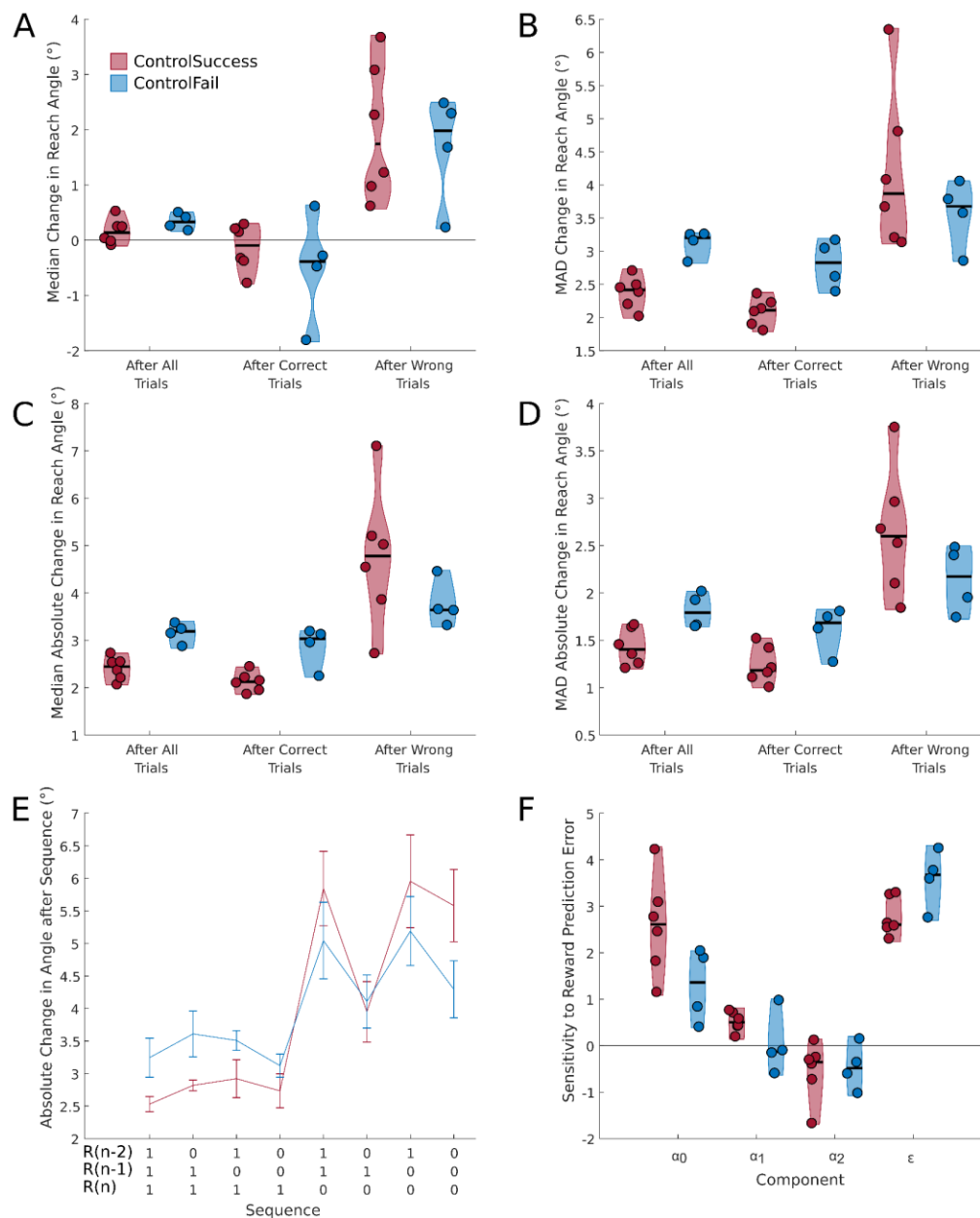
588

589 **Figure 6. Experiment 2: trial-by-trial behavior.** Example of trial by trial reach angles from a subject
 590 performing the dual task (A) rewarded trials are indicated with a circular marker and non-rewarded
 591 trials with a 'x'. The grey shaded region represents the reward region. B, Failure points for subjects
 592 in the DualTask group, thick lines are the mean reach angle for each subject at each rotation angle,
 593 thin lines represent mean of each block, colors go from hot to cold matching failure angles ranging
 594 from high to low. Vertical lines represent the last angle at which mean reach fell within rewarded
 595 region for each subject. The mean and standard deviation of the angle of failure is reported as text in
 596 the figure.

597

598 Finally, the DualTask subjects successfully engaged in the task mental rotation task as evidenced by a
 599 significant difference in percentage of correct button presses ($H(2) = 15.30, p < 0.001$). The DualTask
 600 group responded correctly ($67.21 \pm 3.60\%$) more than either the ControlSuccess ($p = 0.014$) and the
 601 ControlFail ($p = 0.002$) groups. Engagement in the DualTask increased reaction time when compared

602 to ControlSuccess ($p = 0.007$). There was no effect of Group on movement time ($H(2) = 0.33, p =$
 603 0.84).



604

605 **Figure 7. Experiment 2: performance after correct and incorrect trials.** Analysis of the effects of the
 606 success of the previous trial and reward history on trial by trial changes in reach angle for the two
 607 groups performing the control task in Experiment 2. Distribution plots for median (A) and MAD (B)
 608 of change in reach angle separated by the success of the previous trial. Median (C) and MAD (D) of
 609 the absolute change in reach angle separated by the success of the previous trial. E, the absolute

610 *change in reach angle following all combinations of trial success over the previous three trials. F,*
611 *sensitivity to the outcomes of each of the previous trials.*

612

613 **Discussion**

614

615 The role of explicit processes during reinforcement-based motor learning was previously unclear.
616 Here, we reveal that successfully learning to compensate for large, gradually introduced, rotations
617 based on binary (reinforcement-based) feedback involves the development of a strong explicit
618 component, and that not all subjects are able to do so. In both Experiment 1 and the Control group of
619 Experiment 2 only two thirds of subjects were able to successfully learn a large perturbation, and
620 those that did accomplished this principally via explicit processes. Analysis of the trial-by-trial
621 behavior indicated that subjects adjusted their motor commands mainly in response to incorrect trials,
622 and that they were most sensitive to errors made in the most recent trial. Subjects who would go on to
623 fail to learn the full rotation exhibited reduced sensitivity to errors, even in the initial period in which
624 they successfully followed the rotation. Further evidence for the explicit nature of the learning in this
625 task was provided by Experiment 2, where increasing cognitive load via the addition of a dual task
626 prevented learning.

627

628 Previous experiments investigating the learning of rotations based on binary feedback have employed
629 relatively small angles (Izawa and Shadmehr, 2011; Pekny et al., 2015; Therrien et al., 2016), with the
630 15° rotation used by Therrien et al. (2016) the largest reported to date. Indeed, when a rotation of 15°
631 was used in Experiment 1 all subjects were successful in fully compensating for the visual rotation.
632 Furthermore, there was no evidence for an explicit component to retention in the subjects who learnt
633 the 15° rotation. In contrast, successful subjects in both experiments with a 25° rotation demonstrated
634 a large explicit component to the learning, evidenced by a large reduction in the reach angle when
635 asked to remove any strategy. It could therefore be speculated that multiple mechanisms might be
636 available when learning from binary feedback, but that if the size of the perturbation exceeds a certain

637 magnitude an explicit process is required to compensate for it. Previously it has been suggested that
638 additional learning mechanisms are recruited in response to gradually introduced visuomotor rotations
639 when only end-point feedback is available, (Izawa and Shadmehr, 2011; Saijo and Gomi, 2010).
640 Indeed Saijo and Gomi (2010) suggest, on the basis of an increase in reaction times, that explicit
641 changes in motor planning occur in this paradigm. Furthermore, similarly to the results presented
642 here, the authors also find that not all subjects are able to accomplish this. However, none of the
643 previous studies investigating learning of rotations based on binary feedback (Izawa and Shadmehr,
644 2011; Pekny et al., 2015; Therrien et al., 2016) have attempted to dissect the role of implicit and
645 explicit processes. However, learning a rotation based on binary feedback was not accompanied by a
646 change in perceived hand position, as was found when learning was based on full visual feedback of
647 the cursor (Izawa and Shadmehr, 2011). This could be taken as evidence that the learning described
648 by the authors was also explicit in nature in contrast to the implicit, cerebellar-driven, adaptation.

649

650 There is increasing appreciation of the role of explicit processes in traditional visuomotor adaptation
651 paradigms, in which visibility of the cursor ensures that both direction and magnitude of the error are
652 available (Bond and Taylor, 2015, 2017). The use of an ‘error-clamp’ technique has estimated the
653 limit of implicit adaptation based on sensory prediction errors to be at around 15° (Morehead et al.,
654 2017). Such an estimate is roughly in accordance with other estimates obtained either by the use of
655 forcibly reduced movement preparation times (Haith et al., 2015; Leow et al., 2017), self-reporting of
656 aiming directions (Bond and Taylor, 2015) or the difference between trials with and without an
657 explicit component (Werner et al., 2015). It is important to note in our data that all groups, with the
658 exception of those performing the dual task, display a small amount of retention even after the
659 removal of the explicit component suggesting that there is some implicit aspect to the learning.
660 Presumably the implicit learning process triggered in the current study is distinct from the sensory
661 prediction error driven processes as here the error signal is binary in nature and provides no
662 information about direction or magnitude of error. However, it is interesting that such implicit
663 processes appear to be unable to compensate for rotations greater than $15\text{-}20^\circ$, with explicit
664 mechanisms required for greater angles. Haith and Krakauer (2013) have proposed a theoretical

665 framework in which model-based (strategic/explicit) and implicit model-free (reinforcement/use-
666 dependent) learning processes contribute to motor learning. Our findings suggest that in the current
667 paradigm these processes might be engaged but that implicit processes are limited in the size of
668 rotation they can learn. It remains to be seen if this is a limitation of magnitude, as with learning from
669 sensory prediction errors, or a limitation of speed. In other words, if the rotation was introduced more
670 gradually or held constant for a longer period, could this implicit process account for all learning? It is
671 unclear whether the implicit retention observed here reflects use-dependent learning, implicit
672 reinforcement learning or a combination of both (Diedrichsen et al., 2010). However, the current
673 experimental design does not allow us to dissociate between these possibilities. Interestingly, the
674 greatest amount of implicit retention was observed in the 25RotControl group who had received an
675 additional fifty no feedback trials. Given the lack of reward in these trials, this suggests that use-
676 dependent learning at least contributes to the implicit retention observed.

677

678 We measured the explicit contribution to learning via the use of an include/exclude design similar to
679 Werner et al. (2015), which probes the contribution at the end of learning. Other approaches such as
680 asking subjects to verbally report the aiming direction (Taylor et al., 2014) have the advantage of
681 probing the relative contributions of implicit and explicit processes throughout learning. However, it
682 has been suggested that this method may increase the explicit component by priming subjects that re-
683 aiming is beneficial (Leow et al., 2017; Taylor et al., 2014). Such priming may be particularly powerful
684 in paradigms like the current one as it has been shown that explicit awareness of the dimensions over
685 which to explore is required for motor learning based on binary feedback (Manley et al., 2014).
686 Alternatively, forcing subjects to respond at reduced reaction times can also suppresses the explicit
687 component of adapting to a rotation (Haith et al., 2015; Leow et al., 2017). However, Leow et al.
688 (2017) report that even at extremely short reaction times re-aiming to a single target, as used here, is
689 still possible. In future, approaches such as measuring eye movement (Rand and Rentsch, 2016) may
690 be beneficial to measure the explicit component during learning without priming subjects.

691

692 There is ongoing debate about the precise definition of the terms implicit and explicit when applied in
693 a motor learning context (Kleynen et al., 2014). As the authors note implicit and explicit learning may
694 not represent a dichotomy but instead ends of a continuum. The results of this experiment suggest that
695 indeed a binary distinction may not be possible as successful participants here demonstrate awareness
696 but mixed levels of verbalizable strategies, even when they are able to return to reaching at baseline
697 angles on request. Distinction of these possibilities is further complicated by relying on questionnaires
698 (Shanks and John, 1994). Moreover, responses are not always easy to classify into categories and
699 some subjects hold their views in low-confidence. Here we define the explicit component to learning
700 as the amount that participants could remove on request. Such a definition of explicit motor control
701 (Mazzoni and Wexler, 2009) could be more akin to awareness (Werner et al., 2015) or a form of
702 cognitive control (Cavanagh et al., 2009), rather than an explicit strategy which is often defined as a
703 subject's ability to verbalize the strategy they have employed.

704

705 In order to investigate the mechanism through which subjects learnt to counter the rotation we
706 employed the same analysis as Pekny et al., (2015). However, their study didn't involve learning as
707 such, as the rotation was immediately washed out. Despite this, our results are remarkably similar, in
708 that subjects in both studies made larger and more variable changes in actions following trials in
709 which they made an error. Sidarta et al. (2016) have also described a similar pattern of behavior when
710 subjects attempt to find a hidden target zone based on binary feedback, with greater reductions in
711 error following incorrect trials. Our results indicate that subjects who were unable to learn the full
712 rotation made smaller and less variable changes in response to errors and this was primarily driven by
713 their sensitivity to the outcome of the previous trial. Learning from errors has been suggested to be a
714 signature of explicit reinforcement learning, in contrast to learning from success in implicit learning
715 (Loonis et al., 2017). Therefore, the finding that the difference between successful and unsuccessful
716 subjects in the current experiments was in response to errors further supports the idea that it is the
717 sensitivity of the explicit system that is important for this task. However, from the data presented here
718 it is impossible to determine if the corrections following errors are explicit in nature or due to implicit
719 motor variability (He et al., 2016; Wu et al., 2014). In future, similar experiments investigating the

720 presence of neural signatures of explicit learning in tasks such as this may be able to shed light on
721 which process underlie trial-by-trial changes (Loonis et al., 2017). Interestingly, the pattern of reduced
722 sensitivity to errors found for unsuccessful subjects in the current experiment was similar to that
723 described for parkinsonian patients (Pekny et al., 2015). Genetic variability in various aspects of the
724 dopaminergic system has previously been linked to differential performance in reinforcement learning
725 (Frank et al., 2007, 2009), and the balance of model-free and model-based decision-making systems
726 (Doll et al., 2016). Future experiments assessing if the same genetic principles apply to motor learning
727 based on reward may be useful in not only explaining the variation in response but also cementing the
728 links between the principles of reinforcement learning and motor learning (Chen et al., 2017, 2018).
729 Interestingly, the magnitude of changes made in response to errors in a binary feedback based motor
730 learning task was correlated with connectivity changes between motor areas, prefrontal cortex and the
731 intraparietal sulcus (Sidarta et al., 2016). The prefrontal cortex and intraparietal sulcus have been
732 associated with the model-based decision making system (Gläscher et al., 2010), adding further
733 evidence for a pivotal role of explicit systems in reward-based motor learning. However, it should be
734 noted that effects of attention and motivation cannot be ruled out in the current paradigm. Therefore,
735 accompanying neurophysiological measures of these variables may be useful in elucidating their
736 possible contribution.

737

738 The efficacy of the dual task paradigm employed here in preventing learning is remarkable. Dual
739 tasks have previously been employed in conjunction with motor adaptation to visuomotor rotations
740 (Galea et al., 2010), force-fields (Keisler and Shadmehr, 2010; Taylor and Thoroughman, 2007,
741 2008), as well as during the learning of motor skills (Maxwell et al., 2001) and sequence learning
742 (Brown and Robertson, 2007). Galea et al. (2010) demonstrated that a secondary task can slow the
743 rate of adaptation to both a gradually and abruptly introduced visuomotor rotation. Keisler and
744 Shadmehr (2010) found that a declarative memory task could interfere with the ‘fast’ adaptation
745 system but that a demanding cognitive task without the memory component did not. Furthermore,
746 inhibition of the ‘fast’ process led to an increase in the ‘slow’, non-declarative process. Similarly in a
747 sequence learning task a dual task with a declarative element increased the procedural learning

748 suggesting that these two aspects of learning may be in competition (Brown and Robertson, 2007). It
749 could therefore be hypothesized that the use of a dual task in the current paradigm would shift
750 learning from explicit to the implicit system. However, the current data suggest that this did not occur
751 and for this paradigm the explicit system is necessary to compensate for large rotations, and cannot be
752 substituted for by an increase in the use of the implicit learning system. Alternatively, if the implicit
753 system is not engaged by the nature of this task then it would be impossible for it to compensate for
754 the disruption of the explicit system. Arguing against this possibility is the fact that implicit retention
755 was observed in this paradigm, suggesting that the implicit system is indeed engaged, at least to some
756 degree. Whereas previous experiments have employed secondary tasks that involve more verbal
757 systems (Galea et al., 2010; Keisler and Shadmehr, 2010; Taylor and Thoroughman, 2007), we
758 selected the dual task which would have the maximum likelihood of disrupting the explicit system
759 (Anguera et al., 2009; Georgopoulos and Massey, 1987). As the difficulty of the secondary task has
760 been linked with the amount of disruption (Taylor and Thoroughman, 2008), it is also possible that
761 the specific nature of the task may also be important and this is an interesting area for future study.
762 One other possibility is that constant impairment of performance due to the secondary task may
763 reduce intrinsic motivation of subjects (Liao and Masters, 2001).

764

765 The distinction between implicit and explicit reinforcement systems engaging in learning motor tasks
766 is not merely academic. At least part of the increased interest in the addition of reward to motor
767 adaptation and learning is due to the finding that it increases retention (Abe et al., 2011; Dayan et al.,
768 2014, 2014; Galea et al., 2015; Shmuelof et al., 2012; Therrien et al., 2016), along with the promise
769 this may have in a rehabilitation setting (Goodman et al., 2014; Quattrocchi et al., 2017). However, if
770 the benefits are primarily due to explicit or strategic processes, they may be poorly transferred to other
771 environments and be susceptible to disruption. In line with this, it has been demonstrated that motor
772 skills, such as golf putting or playing table tennis, are less disrupted by manipulations such as dividing
773 cognitive load, reducing reaction times or performing in stressful situations when learnt implicitly
774 (Liao and Masters, 2001; Maxwell et al., 2001). If the final goal of the addition of reward to motor
775 learning tasks is to increase retention for practical rehabilitation then it may be that methods that

776 increase the implicit contribution are required such as employing learning by analogy, reducing errors
777 during learning or the addition of dual tasks (Liao and Masters, 2001). However, the choice and
778 difficulty of the dual task should be made with caution as from the data presented here it may be too
779 disruptive and ultimately prevent learning.

780

781 **Acknowledgements**

782 The authors would like to thank Dr. Xiuli Chen for advice concerning the analysis of the data.

783

784 **Grants**

785 PH, OC, and JM were supported by the European Research Council grant MotMotLearn 637488.

786 **Disclosures**

787 The authors declare no competing financial interests.

788

789 **References**

790 Abe, M., Schambra, H., Wassermann, E.M., Luckenbaugh, D., Schweighofer, N., and Cohen, L.G.
791 (2011). Reward Improves Long-Term Retention of a Motor Memory through Induction of Offline
792 Memory Gains. *Curr. Biol.* 21, 557–562.

793 Anguera, J.A., Reuter-Lorenz, P.A., Willingham, D.T., and Seidler, R.D. (2009). Contributions of
794 Spatial Working Memory to Visuomotor Learning. *J. Cogn. Neurosci.* 22, 1917–1930.

795 Anguera, J.A., Bernard, J.A., Jaeggi, S.M., Buschkuhl, M., Benson, B.L., Jennett, S., Humfleet, J.,
796 Reuter-Lorenz, P.A., Jonides, J., and Seidler, R.D. (2012). The effects of working memory resource
797 depletion and training on sensorimotor adaptation. *Behav. Brain Res.* 228, 107–115.

798 Bond, K.M., and Taylor, J.A. (2015). Flexible explicit but rigid implicit learning in a visuomotor
799 adaptation task. *J. Neurophysiol.* 113, 3836–3849.

800 Bond, K.M., and Taylor, J.A. (2017). Structural Learning in a Visuomotor Adaptation Task Is
801 Explicitly Accessible. *ENeuro* 4.

802 Brown, R.M., and Robertson, E.M. (2007). Inducing motor skill improvements with a declarative
803 task. *Nat. Neurosci.* 10, nn1836.

804 Cavanagh, J.F., Cohen, M.X., and Allen, J.J.B. (2009). Prelude to and Resolution of an Error: EEG
805 Phase Synchrony Reveals Cognitive Control Dynamics during Action Monitoring. *J. Neurosci.* 29,
806 98–105.

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

809 Chen, X., Holland, P., and Galea, J.M. (2018). The effects of reward and punishment on motor skill
810 learning. *Curr. Opin. Behav. Sci.* 20, 83–88.

811 Christou, A.I., Miall, R.C., McNab, F., and Galea, J.M. (2016). Individual differences in explicit and
812 implicit visuomotor learning and working memory capacity. *Sci. Rep.* 6.

813 Dayan, E., Averbeck, B.B., Richmond, B.J., and Cohen, L.G. (2014). Stochastic reinforcement
814 benefits skill acquisition. *Learn. Mem.* 21, 140–142.

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

817 Doll, B.B., Bath, K.G., Daw, N.D., and Frank, M.J. (2016). Variability in Dopamine Genes
818 Dissociates Model-Based and Model-Free Reinforcement Learning. *J. Neurosci.* 36, 1211–1222.

819 Frank, M.J., Moustafa, A.A., Haughey, H.M., Curran, T., and Hutchison, K.E. (2007). Genetic triple
820 dissociation reveals multiple roles for dopamine in reinforcement learning. *Proc. Natl. Acad. Sci. U.*
821 *S. A.* 104, 16311–16316.

822 Frank, M.J., Doll, B.B., Oas-Terpstra, J., and Moreno, F. (2009). The neurogenetics of exploration
823 and exploitation: Prefrontal and striatal dopaminergic components. *Nat. Neurosci.* 12, 1062–1068.

824 Galea, J.M., Sami, S.A., Albert, N.B., and Miall, R.C. (2010). Secondary tasks impair adaptation to
825 step- and gradual-visual displacements. *Exp. Brain Res.* 202, 473–484.

826 Galea, J.M., Mallia, E., Rothwell, J., and Diedrichsen, J. (2015). The dissociable effects of
827 punishment and reward on motor learning. *Nat. Neurosci.* 18, 597–602.

828 Georgopoulos, A.P., and Massey, J.T. (1987). Cognitive spatial-motor processes. *Exp. Brain Res.* 65,
829 361–370.

830 Gläscher, J., Daw, N., Dayan, P., and O’Doherty, J.P. (2010). States versus Rewards: Dissociable
831 Neural Prediction Error Signals Underlying Model-Based and Model-Free Reinforcement Learning.
832 *Neuron* 66, 585–595.

833 Goodman, R.N., Rietschel, J.C., Roy, A., Jung, B.C., Diaz, J., Macko, R.F., and Forrester, L.W.
834 (2014). Increased reward in ankle robotics training enhances motor control and cortical efficiency in
835 stroke. *J. Rehabil. Res. Dev.* 51, 213–228.

836 Haith, A.M., and Krakauer, J.W. (2013). Model-Based and Model-Free Mechanisms of Human Motor
837 Learning. *Adv. Exp. Med. Biol.* 782, 1–21.

838 Haith, A.M., Huberdeau, D.M., and Krakauer, J.W. (2015). The Influence of Movement Preparation
839 Time on the Expression of Visuomotor Learning and Savings. *J. Neurosci.* 35, 5109–5117.

840 He, K., Liang, Y., Abdollahi, F., Bittmann, M.F., Kording, K., and Wei, K. (2016). The Statistical
841 Determinants of the Speed of Motor Learning. *PLOS Comput Biol* 12, e1005023.

842 Huang, J.-J., Yen, C.-T., Tsao, H.-W., Tsai, M.-L., and Huang, C. (2014). Neuronal Oscillations in
843 Golgi Cells and Purkinje Cells are Accompanied by Decreases in Shannon Information Entropy. *The*
844 *Cerebellum* 13, 97–108.

845 Huberdeau, D.M., Krakauer, J.W., and Haith, A.M. (2015). Dual-process decomposition in human
846 sensorimotor adaptation. *Curr. Opin. Neurobiol.* 33, 71–77.

847 Izawa, J., and Shadmehr, R. (2011). Learning from Sensory and Reward Prediction Errors during
848 Motor Adaptation. *PLOS Comput Biol* 7, e1002012.

849 Izawa, J., Criscimagna-Hemminger, S.E., and Shadmehr, R. (2012). Cerebellar Contributions to
850 Reach Adaptation and Learning Sensory Consequences of Action. *J. Neurosci.* 32, 4230–4239.

851 Keisler, A., and Shadmehr, R. (2010). A Shared Resource between Declarative Memory and Motor
852 Memory. *J. Neurosci.* 30, 14817–14823.

853 Kleynen, M., Braun, S.M., Bleijlevens, M.H., Lexis, M.A., Rasquin, S.M., Halfens, J., Wilson, M.R.,
854 Beurskens, A.J., and Masters, R.S.W. (2014). Using a Delphi Technique to Seek Consensus
855 Regarding Definitions, Descriptions and Classification of Terms Related to Implicit and Explicit
856 Forms of Motor Learning. *PLOS ONE* 9, e100227.

857 van der Kooij, K., and Overvliet, K.E. (2016). Rewarding imperfect motor performance reduces
858 adaptive changes. *Exp. Brain Res.* 234, 1441–1450.

859 Leow, L.-A., Gunn, R., Marinovic, W., and Carroll, T.J. (2017). Estimating the implicit component of
860 visuomotor rotation learning by constraining movement preparation time. *J. Neurophysiol.*
861 jn.00834.2016.

862 Liao, C.-M., and Masters, R.S.W. (2001). Analogy learning: A means to implicit motor learning. *J.*
863 *Sports Sci.* 19, 307–319.

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

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

868 Maxwell, J.P., Masters, R.S.W., Kerr, E., and Weedon, E. (2001). The implicit benefit of learning
869 without errors. *Q. J. Exp. Psychol. Sect. A* 54, 1049–1068.

870 Mazzoni, P., and Wexler, N.S. (2009). Parallel Explicit and Implicit Control of Reaching. PLOS ONE
871 4, e7557.

872 McDougle, S.D., Bond, K.M., and Taylor, J.A. (2015). Explicit and Implicit Processes Constitute the
873 Fast and Slow Processes of Sensorimotor Learning. J. Neurosci. 35, 9568–9579.

874 Morehead, J.R., Taylor, J.A., Parvin, D.E., and Ivry, R.B. (2017). Characteristics of Implicit
875 Sensorimotor Adaptation Revealed by Task-irrelevant Clamped Feedback. J. Cogn. Neurosci. 29,
876 1061–1074.

877 Nikooyan, A.A., and Ahmed, A.A. (2015). Reward feedback accelerates motor learning. J.
878 Neurophysiol. 113, 633–646.

879 Pekny, S.E., Izawa, J., and Shadmehr, R. (2015). Reward-Dependent Modulation of Movement
880 Variability. J. Neurosci. 35, 4015–4024.

881 Peters, M., and Battista, C. (2008). Applications of mental rotation figures of the Shepard and Metzler
882 type and description of a mental rotation stimulus library. Brain Cogn. 66, 260–264.

883 Quattrocchi, G., Greenwood, R., Rothwell, J.C., Galea, J.M., and Bestmann, S. (2017). Reward and
884 punishment enhance motor adaptation in stroke. BioRxiv 106377.

885 Rand, M.K., and Rentsch, S. (2016). Eye-Hand Coordination during Visuomotor Adaptation with
886 Different Rotation Angles: Effects of Terminal Visual Feedback. PLOS ONE 11, e0164602.

887 Saijo, N., and Gomi, H. (2010). Multiple Motor Learning Strategies in Visuomotor Rotation. PLOS
888 ONE 5, e9399.

889 Shanks, D.R., and John, M.F.S. (1994). Characteristics of dissociable human learning systems. Behav.
890 Brain Sci. 17, 367–395.

891 Shepard, R.N., and Metzler, J. (1971). Mental Rotation of Three-Dimensional Objects. Science 171,
892 701–703.

893 Shmuelof, L., Huang, V.S., Haith, A.M., Delnicki, R.J., Mazzoni, P., and Krakauer, J.W. (2012).
894 Overcoming Motor “Forgetting” Through Reinforcement Of Learned Actions. *J. Neurosci.* 32,
895 14617–14621.

896 Sidarta, A., Vahdat, S., Bernardi, N.F., and Ostry, D.J. (2016). Somatic and Reinforcement-Based
897 Plasticity in the Initial Stages of Human Motor Learning. *J. Neurosci.* 36, 11682–11692.

898 Smith, M.A., Ghazizadeh, A., and Shadmehr, R. (2006). Interacting Adaptive Processes with
899 Different Timescales Underlie Short-Term Motor Learning. *PLOS Biol.* 4, e179.

900 Taylor, J.A., and Thoroughman, K.A. (2007). Divided Attention Impairs Human Motor Adaptation
901 But Not Feedback Control. *J. Neurophysiol.* 98, 317–326.

902 Taylor, J.A., and Thoroughman, K.A. (2008). Motor Adaptation Scaled by the Difficulty of a
903 Secondary Cognitive Task. *PLoS ONE* 3.

904 Taylor, J.A., Krakauer, J.W., and Ivry, R.B. (2014). Explicit and Implicit Contributions to Learning in
905 a Sensorimotor Adaptation Task. *J. Neurosci.* 34, 3023–3032.

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

908 Tseng, Y., Diedrichsen, J., Krakauer, J.W., Shadmehr, R., and Bastian, A.J. (2007). Sensory
909 Prediction Errors Drive Cerebellum-Dependent Adaptation of Reaching. *J. Neurophysiol.* 98, 54–62.

910 Werner, S., Aken, B.C. van, Hulst, T., Frens, M.A., Geest, J.N. van der, Strüder, H.K., and Donchin,
911 O. (2015). Awareness of Sensorimotor Adaptation to Visual Rotations of Different Size. *PLOS ONE*
912 10, e0123321.

913 Wu, H.G., Miyamoto, Y.R., Castro, L.N.G., Ölveczky, B.P., and Smith, M.A. (2014). Temporal
914 structure of motor variability is dynamically regulated and predicts motor learning ability. *Nat.*
915 *Neurosci.* 17, 312–321.

916

917 **Figure Captions**

918 **Figure 1. Experimental design.** *A*, Subjects held the handle of robotic manipulandum with their right
919 hand, the position of the arm and handle was hidden from sight and feedback was provided on a
920 horizontal screen. *B*, Subjects made ‘shooting’ movements from a starting position (green circle)
921 towards a target (red circle), after the initial practice trials the position of the cursor (white circle)
922 was no longer visible at any point. *C*, Successful trials were indicated to the subject with the display
923 of a green tick after the cursor had passed through a region centered on the target, over the course of
924 the paradigm the position of the reward region gradually moved (solid green circle to dashed green
925 circle) whilst the visible target (red circle) remained in the central location. By the end of the learning
926 period a successful reach (dotted white line) was rotated by a maximum of either 15° or 25°. *D*, Time-
927 course of Experiment 2, at the same time as the target appeared on screen a ‘shape’ was also
928 displayed slightly above it, the subject was asked to memorize this shape. After the reach was
929 completed and the hand returned to the starting position subjects used their left hand to respond with
930 a button press as to whether they believed the new shape shown on screen was a rotated version of the
931 shape or an entirely different shape.

932

933 **Figure 2. Experiment 1: group performance.** *A*, Reach angle averaged over blocks of 5 trials, solid
934 colored lines represent the mean of each group and the shaded region represents SEM. The average
935 behavior of subjects in the 15Rot paradigm (Orange) fell consistently within the rewarded region
936 (grey shaded region) indicating successful learning. *B*, Average reach angle over blocks for all
937 subjects in the 25Rot paradigm (magenta) and also the same subjects split into two groups based on
938 success at the final angle (25RotSuccess – red, 25RotFail – blue). *C*, Distribution plots displaying the
939 reach angles for subjects in the three groups at various timepoints throughout the experiment with
940 individual data points overlaid on an estimate of the distribution. Horizontal black line in the
941 distribution represents the group median. *D*, Distribution plots of the computed variables of Implicit
942 (‘Remove-Baseline’) and Explicit (‘Maintain-Implicit’) retention. Significance stars above horizontal
943 black bars indicate differences between the groups (* $P < 0.05$, ** $P < 0.01$, *** $P < 0.001$).
944 Significance stars below the distributions represent a significant difference from zero. *E*, Reach angle
945 averaged over blocks of 5 trials for subjects in the 25RotControl group. There was no reduction in

946 reach angle during the time taken for the control questions between Maintain 1 and Maintain 2
947 blocks. However, when subjects were subsequently asked to remove their strategy, the period between
948 Maintain 2 and Remove blocks, a significant reduction in reach angle was observed.

949

950 **Figure 3. Experiment 1: trial-by-trial behavior.** Example of trial by trial reach angles from a subject
951 who was successful at the final angle (**A**) and one who was unsuccessful (**B**). In each case rewarded
952 trials are indicated with a circular marker and non-rewarded trials with a 'x'. The grey shaded
953 region indicates the reward region. **C**, Failure points for subjects in the 25RotFail group, thick lines
954 are the mean reach angle for each subject at each rotation angle, thin lines represent mean of each
955 block (average of 5 trials), colors go from hot to cold matching failure angles ranging from high to
956 low. Vertical lines represent the last angle at which mean reach fell within rewarded region for each
957 subject. The mean and standard deviation of all angles of failure is displayed as text.

958

959 **Figure 4. Experiment 1: performance after correct and incorrect trials.** Analysis of the effects of the
960 success of the previous trial and reward history on trial by trial changes in reach angle for the three
961 groups in Experiment 1 (15Rot – Orange, 25RotSuccess – Red, 25RotFail – Blue). Median (**A**) and
962 MAD (**B**) of change in reach angle separated by the success of the previous trial. Median (**C**) and
963 MAD (**D**) of the absolute change in reach angle separated by the success of the previous trial. **E**, The
964 absolute change in reach angle following all combinations of trial success over the previous three
965 trials. **F**, Sensitivity to the outcomes of each of the previous trials. Significance stars above horizontal
966 black bars indicate differences between the groups (* $P < 0.05$, ** $P < 0.01$).

967

968 **Figure 5. Experiment 2: group performance.** Change in reach angle over blocks (average of 5 trials)
969 during the dual task experiment. **A**, Group performance for the DualTask (Green) and Control
970 (Purple) task groups, the line indicates the mean and shaded region the SEM. The grey shaded region
971 represents the reward region. **B**, the split of the control task group into ControlSuccess (Dark Red)
972 and ControlFail (Blue). **C**, Distribution plots displaying the performance at different time points for
973 the dual task, and split control groups. The shaded region represents an estimation of the distribution

974 and is overlaid with data for each individual subject. **D**, Distribution plots of the difference in reach
975 angle during retention phases indicating the implicit and explicit components of retention.
976 Significance stars above horizontal black bars indicate differences between the groups (* $P < 0.05$,
977 ** $P < 0.01$).

978

979 **Figure 6. Experiment 2: trial-by-trial behavior.** Example of trial by trial reach angles from a subject
980 performing the dual task (**A**) rewarded trials are indicated with a circular marker and non-rewarded
981 trials with a 'x'. The grey shaded region represents the reward region. **B**, Failure points for subjects
982 in the DualTask group, thick lines are the mean reach angle for each subject at each rotation angle,
983 thin lines represent mean of each block, colors go from hot to cold matching failure angles ranging
984 from high to low. Vertical lines represent the last angle at which mean reach fell within rewarded
985 region for each subject. The mean and standard deviation of the angle of failure is reported as text in
986 the figure.

987

988 **Figure 7. Experiment 2: performance after correct and incorrect trials.** Analysis of the effects of the
989 success of the previous trial and reward history on trial by trial changes in reach angle for the two
990 groups performing the control task in Experiment 2. Distribution plots for median (**A**) and MAD (**B**)
991 of change in reach angle separated by the success of the previous trial. Median (**C**) and MAD (**D**) of
992 the absolute change in reach angle separated by the success of the previous trial. **E**, the absolute
993 change in reach angle following all combinations of trial success over the previous three trials. **F**,
994 sensitivity to the outcomes of each of the previous trials.

995

