1   *Reward-based invigoration of sequential reaching*

2

3   **Sebastian Sporn[1], Xiuli Chen[1] & Joseph M Galea[1]**

4

5   *[1]School of Psychology and Centre for Human Brain Health,*

6   *University of Birmingham, B15 2TT, UK*

7

8   <u>Conflict of Interest:</u> The authors declare no competing financial interests

11   <u>Correspondence</u>:

12   Sebastian Sporn

13   Email: SSS748@bham.ac.uk

14

15

16

17

18

19

20

21

22

23

24

25

26

27

28

29

30

31

Abstract

Seeking reward is a powerful tool for shaping human behaviour. While it has been demonstrated that reward invigorates performance of simple movements, its effect on more complex sequential actions is less clear. In addition, it is unknown why reward-based improvements for discrete actions are transient, i.e. performance gains are lost once reward is removed, but appear long lasting for sequential actions. We show across three experiments that reward invigorates sequential reaching performance. Driven by a reward-based increase in speed, movements also exhibited greater coarticulation, smoothness and a closer alignment to a minimum jerk trajectory. Critically, these performance gains were maintained across multiple days even after the removal of reward. We propose that coarticulation, the blending together of sub-movements into a single continuous action, provides a mechanism by which reward can invigorate sequential performance whilst also increasing efficiency. This change in efficiency appears essential for the retention of reward-based improvements in motor behaviour.

Introduction

From drinking a cup of coffee to cleaning your teeth, sequential actions form a fundamental component of our daily life activities. When first encountered, these actions are often executed as a series of distinct sub-movements with pronounced stop periods between them [1–8]. However, with learning these discrete sub-movements are gradually blended together to form a continuous action that is executed with increased speed, smoothness and energetic-efficiency [9–12]. This process, known as coarticulation, is an essential mechanism for understanding skilled sequential performance acquisition as it reflects the evolution of behaviour, both temporally and spatially, towards increased efficiency [7–12]. Coarticulation is observed across many motor behaviours such as speech production [13,14], sign language [15], piano playing [16], typing [17–19] and various other upper limb actions [20–25]. Within sequential reaching, the coarticulation of sub-movements leads to the gradual development of a new motor primitive that is globally planned and, once initiated, must run to completion [10]. Although this process can take weeks of practise, these newly formed motor primitives are highly generalizable and are not simply the result of increased movement speed [10].

It is important to emphasize that coarticulation represents a fundamentally different process from chunking which appears to feature similar characteristics. Chunking refers to the production of a series of discrete movements (i.e. button presses) that are temporally aligned [21,22,26–28]. This is represented at a behavioural level through shorter reaction times between actions [7,29], however each action is still performed discretely with a pronounced stop period between them [21,30]. In contrast, coarticulation reflects the merging of neighbouring movements into a continuous and kinematically distinct motor primitive [9–12].

Despite coarticulation being important to the development of skilled sequential behaviour, our ability to facilitate this process is limited [9–11,31]. In addition, the continuous actions that arise from coarticulation appear to breakdown in several motor disorders [32,33]. For instance, stroke patients suffering motor impairments produce actions that are decomposed into jerky sub-movements, with successful recovery being associated with a return to smooth continuous actions [34–37]. Therefore, identifying a mechanism which can facilitate coarticulation could have far reaching implications across health and disease.

Seeking reward is a powerful tool for shaping behaviour [38,39]. For example, the expectation of reward causes individuals to perform saccadic and reaching actions with greater speed and accuracy [40–47]. As a result, there has been significant interest in the

95   potential of using reward to enhance motor behaviour in both healthy individuals and clinical

96   populations such as stroke patients [48,49]. However, it has been shown that these reward-based

97   effects on movement are cognitively demanding, energetically-inefficient and transient i.e.

98   the effects are lost when reward is removed [40–42]. Yet, these results stem from simple tasks

99   that involve singular discrete actions e.g. movement towards a single static target. In contrast,

100   during more complex sequential or continuous tasks, the beneficial effects of reward appear

101   long lasting and persist even after the removal of reward [50–52]. The mechanism, however, by

102   which reward induces these long-lasting effects and the reason why sequential behaviour

103   seems critical is unknown. An interesting possibility is that reward-based invigoration is

104   associated with increased coarticulation, a mechanism which enables sequential movements

105   to be performed faster but with greater efficiency and therefore promotes long-term

106   retention of these performance benefits.

107

108   **Results**

109   Here we tested this hypothesis, namely that reward can facilitate coarticulation and

110   thereby promote energetically efficient performance gains that persist even in the absence

111   of reward. To this end, our main experiment (N=42) assessed the effect of reward on

112   sequential reaching performance and the evolution of coarticulation over the course of two

113   testing days. We then carried out two further experiments to assess the robustness of these

114   performance gains during an additional testing day without reward availability (N=5), and to

115   investigate whether reward or performance-based feedback drove these observed

116   improvements (N=60).

117   We developed a novel reaching task in which participants made 8 sequential reaching

118   movements to designated targets (1 trial) using a motion tracking device (Figure 1a,b).

119   Participant behaviour could range from executing 8 individual movements (i.e., stopping at

120   each via point; Figure 1c) to a series of 5 coarticulated movements that would reflect the

121   outcome of minimising jerk across the trial (Methods Equation 3; Figure 1d) [12]. Changes in the

122   number of reaching movements were associated with marked differences in the velocity

123   profile. Specifically, individual reaches were characterised by pronounced stops in-between

124   movements, with velocity dropping close to zero. As reaching movements merge through

125   coarticulation, these dips in velocity gradually disappear (Figure 1c). As suggested by the

126   minimum jerk model (Methods, Equation 3), we hypothesised that coarticulation would

127   mainly occur in the central set of three in-centre-out reaching movements (Figure 1c,d) [9–12,53].

a

b

c

d

e
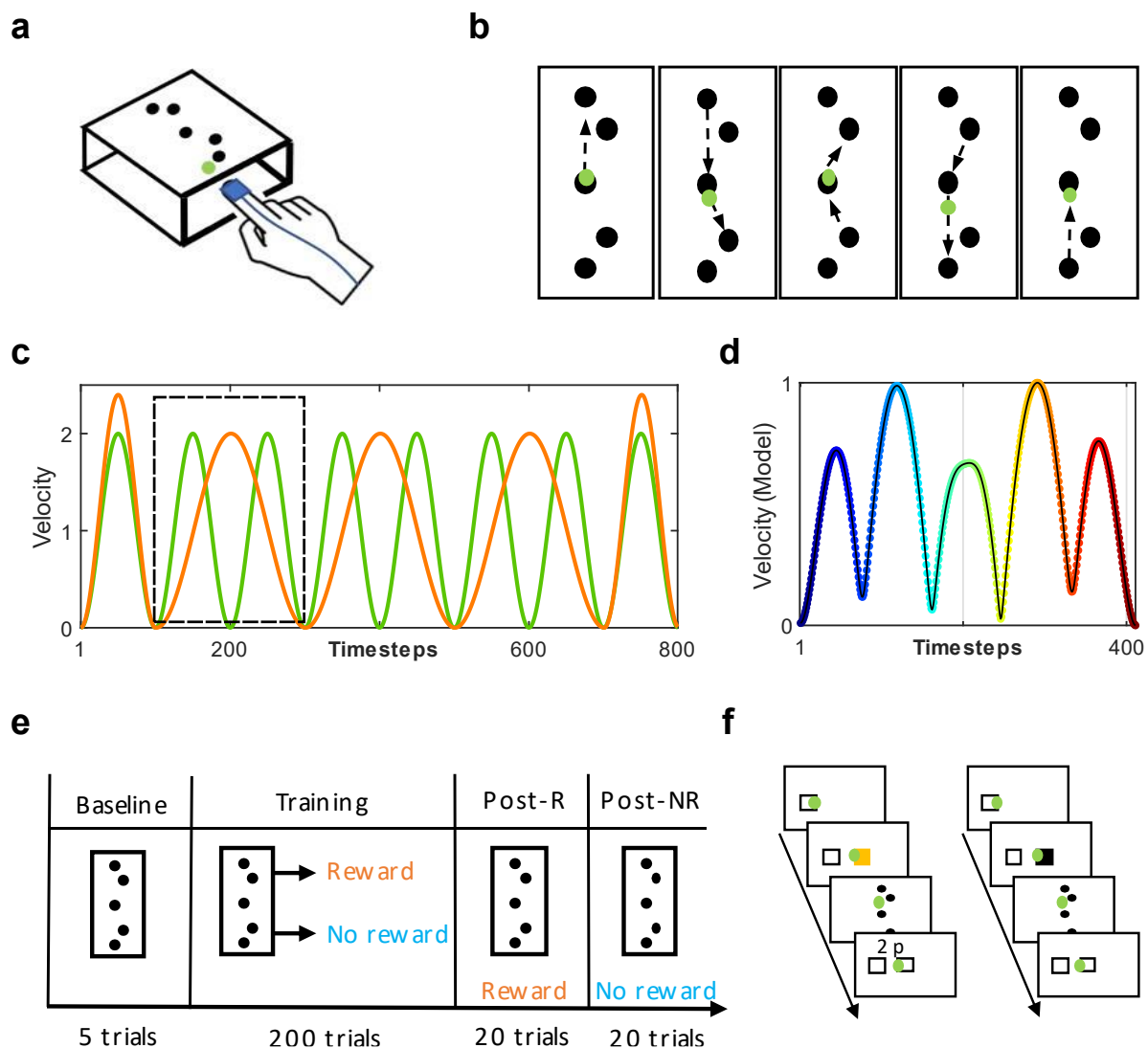
f

128

129

130   **Figure 1 | Experimental setup. a)** Participants wore a motion-tracking device on the index finger and
131   the unseen reaching movements were performed across a table whilst a green cursor matching the
132   position of index finger was viewable on a screen. **b)** 8 movement sequential reaching task. The
133   participants started from the centre target. **c)** Schematic velocity profiles highlighting the difference
134   between single versus coarticulated reaching movements for a trial. Based on the predictions of a
135   minimum jerk model, we assumed that the three sets of in-centre-out reaching movements could be
136   coarticulated which would reduce the number of velocity peaks from 8 (green) to 5 (orange). **d)**
137   Velocity profile for a trial predicted by a minimum jerk model. **e)** Study design. Prior to the start of the
138   experiment, participants were trained on the reaching sequence and were then asked to perform 10
139   baseline trials. Randomly allocated to a reward and no reward group, participants completed 200 trials
140   during training and an additional 20 trials in each post assessment; one with reward (post-R) and one
141   without (post-NR) (counterbalanced across participants). This design was repeated 24 hours later (Day
142   2). **f)** Rewarded trials were cued using a visual stimulus prior to the start of the trial. At the end of the
143   trial, participants received trial-based monetary feedback based on their last 20 trials (close-loop
144   design). In no reward trials, participants were instructed to be as fast and accurate as possible.

5

145   Participants were randomly allocated to a reward or no reward group. Prior to the

146   start of the experiment, participants were trained on the sequence without a time constraint

147   until reaching a learning criterion of 5 successful trials in a row. Therefore, any performance

148   gains could only be attributed to improvements in the execution and not memory of the

149   sequence. There was then a baseline period (10 trials), during which both groups were

150   encouraged to complete each trial 'as fast and as accurately as possible'. Afterwards, during

151   training, both groups completed 200 training trials. Participants in the reward group were

152   able to earn money depending on their movement time (MT) and received performance-

153   based feedback (the amount of money awarded in a given trial) after reaching the last target

154   in each trial. Reward trials were cued with a yellow start box (Figure 1f) and the monetary

155   reward was calculated using a closed-loop design comparing MT performance on a given trial

156   with performance on the last 20 trials. Participants in the no reward group were instructed to

157   move as fast and accurately as possible. Importantly, participants did not receive feedback on

158   their level of coarticulation. Missing a target resulted in an immediate abortion of the current

159   trial, which they then had to repeat, thereby forcing participants to accurately hit each target.

160   After training, both groups engaged in a rewarded (post-R) and no rewarded (post-NR) post-

161   assessment (20 trials each). The post-assessment intended to compare performance between

162   groups when under the same condition (Figure 1e). This design was repeated 24 hours later.

163

Reward invigorates sequential reaching movements

165   Movement time (MT) reflected the total movement duration from exiting the start

166   box until reaching the last target. Our results highlight that reward instantaneously

167   invigorates sequential reaching behaviour with these performance gains being maintained

168   even in the absence of reward (Figure 2a). Specifically, we found a significant decrease in MT

169   over the course of training on both days for both groups (mixed-effect ANOVA; timepoint

170   (early (1st 20 trials) vs late (last 20 trials)) x group; main effect for timepoint, day 1: F = 18.29,

171   p < 0.0001; day2: F = 8.51, p = 0.0058; Figure 2b,d). Importantly, despite no differences in MT

172   between groups during baseline on day 1 (Wilcoxon test; Z = -1.38, p = 0.17), we found that

173   the reward group produced significantly faster MTs than the no reward group across training

174   on both days (main effect for group, day 1: F = 18.96, p < 0.0001; day 2: F = 17.58, p < 0.0001;
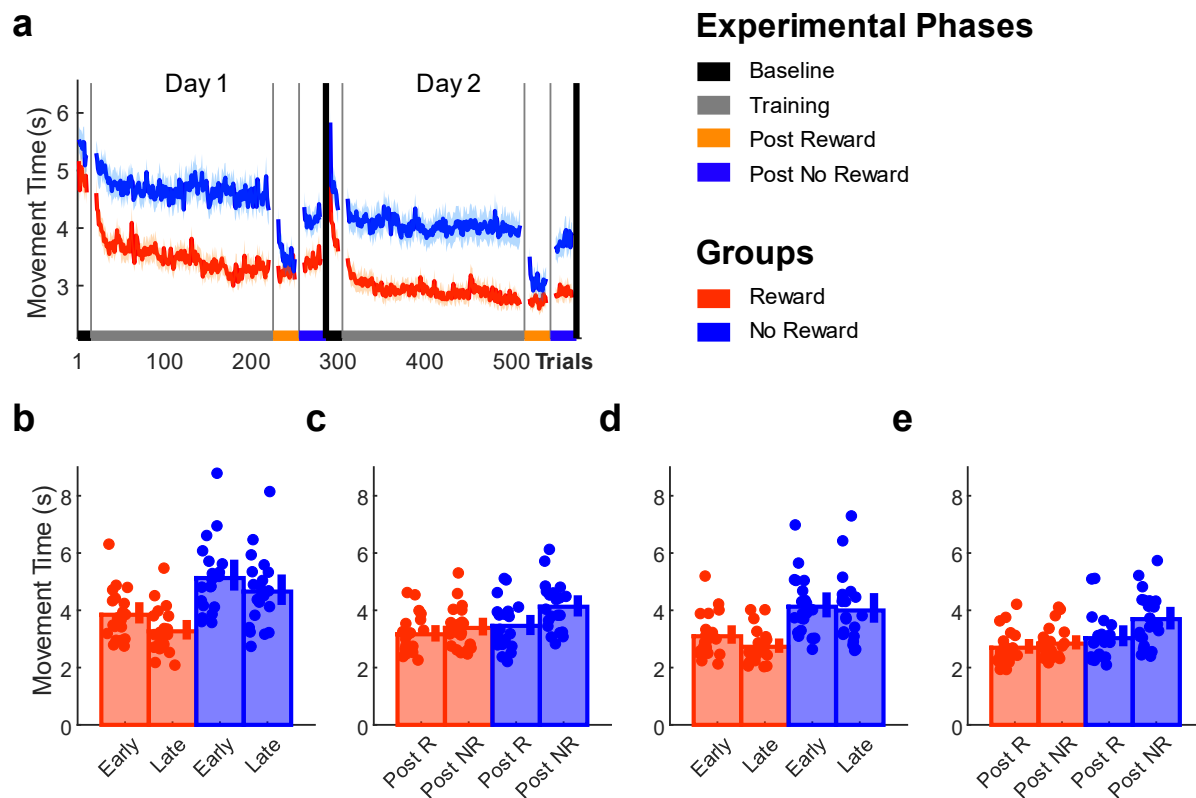
175   Figure 2b,d).

6

Figure 2 | Reward-based improvements in MT. a) Trial-by-trial changes in MT averaged over participants for both groups. **b-e)** Median MT values for each participant for both groups. **b)** Comparing MT performance during training on day 1 (early (first 20 trials) vs late (last 20 trials)). **c)** Post assessment performance (day 1; post-R vs post-NR). **d)** Training (day 2; early vs late). **e)** Post assessment performance (day 2; post-R vs post-NR)

Across post assessments, a significant interaction was found between phase (post-R (all 20 trials) vs post-NR (all 20 trials)) and group (mixed ANOVA; condition x group; interaction, day 1: F = 18.07, p < 0.0001; day 2: F = 19.99, p < 0.0001). Specifically, there was a significant difference in MT between groups during post-NR (Wilcoxon test; day 1: Z = -2.82, p = 0.0192; day 2: Z = -3.27, p = 0.0044; Figure 2c,e) but not post-R (Wilcoxon test; day 1: Z = -1.13, p = 1; day 2: Z = -1.38, p = 1). This indicates that the no reward group were able to instantaneously invigorate their performance during post-R. However, these performance gains were not maintained during post-NR, suggesting that they remained transient in nature. In contrast, the enhanced MT performance in the reward group did not change significantly between post assessments, implying that performance gains had become reward-independent (Wilcoxon test; day 1: Z = -1.11, p = 1; day 2: Z = -0.91, p = 1; Figure 2c,e). The counterbalancing of post assessment order did not affect these results (Supplementary Figure 1).

196   These decreases in MT could be driven by two processes: (1) an invigoration of each

197   reaching movement's peak velocity which has been observed to underlie the transient

198   performance gains in singular discrete reaching tasks [40–42] and (2) the coarticulation of

199   sequential reaching movements which decreases MTs via a reduction in dwell time around

200   the central target.

201

202   **Reward-based invigoration of peak velocities is instantaneous**

203   To assess changes in peak velocities (Methods, equation 1), we averaged peak velocity

204   across the 8 reaching movements (Figure 3a). Over the course of training, we found significant

205   increases in peak velocity on both days (mixed ANOVA; timepoint x group; main effect for

206   timepoint, day 1: $F = 9.72$, $p = 0.0034$; day 2: $F = 8.56$, $p = 0.0057$; Figure 3b,d). Despite there

207   being no differences in peak velocities during baseline on day 1 (Wilcoxon test; $Z = 0.70$, $p =$

208   0.4812), the reward group produced significantly higher peak velocities than the no reward

209   group across training (main effect for group, day 1: $F = 19.42$, $p < 0.0001$; day 2: $F = 19.28$, p

210   < 0.0001; Figure 3b,d). This supports existing findings that reward-based invigoration of MT

211   can be driven by increases in peak velocities [40–42]. Yet again, the no reward group exhibited a

212   pronounced 'on-off' effect across post assessments (post-R vs post-NR) for both days (mixed-

213   effect ANOVA; condition x group; interaction, day 1: $F = 8.02$, $p = 0.0072$; day 2: $F = 24.92$, p

214   < 0.0001). Specifically, we found significant differences between groups during post-NR

215   (Wilcoxon test; day 1: $Z = 2.84$, $p = 0.018$; day 2: $Z = 3.07$, $p = 0.0084$; Figure 3c,e) but not post-

216   R (Wilcoxon test; day 1: $Z = 0.86$, $p = 1$; day 2: $Z = 0.93$, $p = 1$; Figure 3c,e). Similarly to MT,

217   these results suggest that the no reward group were able to instantaneously increase peak

218   velocity during post-R, but these performance gains remained transient in nature and were

219   not maintained during post-NR. In contrast, peak velocities in the reward group remained

220   elevated across post assessments irrespective of reward availability (Wilcoxon test; day 1: $Z =$

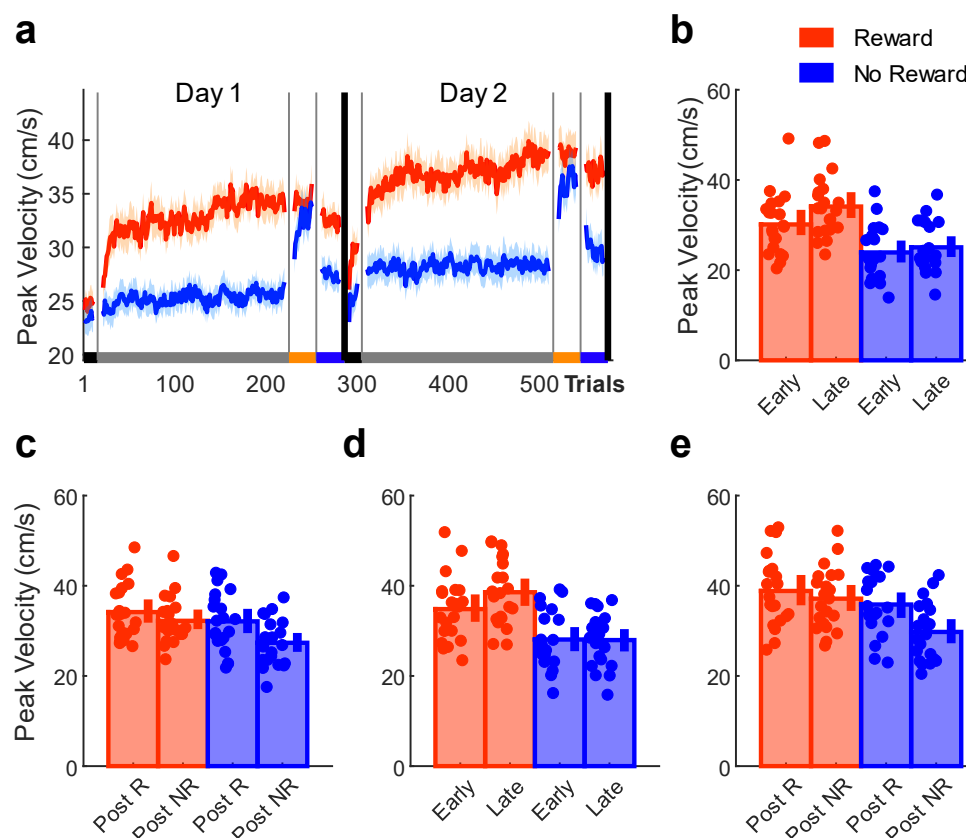221 0.86, $p = 1$; day 2: $Z = 0.68$, $p = 1$; Figure 3c,e).

222

Figure 3 | Reward-based improvements in peak velocity a) Trial-by-trial changes in peak velocity averaged over participants for both groups. **b-e)** Median peak velocity values for each participant for both groups. **b)** Comparing peak velocity performance during training on day 1 (early (first 20 trials) vs late (last 20 trials)). **c)** Post assessment performance (day 1; post-R vs post-NR). **d)** Training (day 2; early vs late). **e)** Post assessment performance (day 2; post-R vs post-NR)

Reward-based facilitation of coarticulation is training-dependent

Coarticulation describes the blending of individual motor elements into a combined smooth action. This is represented in the velocity profile by the stop period between two movements gradually disappearing and being replaced by a single velocity peak (Figure 1c). To measure coarticulation, we developed a coarticulation index (CI; Methods, Equation 2) that compared the mean peak velocities of two sequential reaches with the minimum velocity around the via point. The smaller the difference between these values, the greater coarticulation had occurred of these two movements as reflected by CI value closer to 1 (Figure 4a). As the central 3 segments of the movement could potentially be coarticulated the CI ranged from 0-3 for each trial (Figure 4b). Our results show that reward facilitates movement coarticulation and leads to stable changes in behaviour even in the absence of reward. Additionally changes in coarticulation, in contrast to the changes in peak velocities,

242 appears to be training-dependent. Across training, CI levels increased on both days (mixed

243 ANOVA; timepoint x group; main effect for timepoint, day 1: $F = 21.70$, $p < 0.0001$; day 2: $F =$

244 21.45, $p < 0.0001$; Figure 4b). Crucially, although there were no differences between groups

245 in CI levels during baseline (Wilcoxon test; $Z = 1.31$, $p = 0.1908$), we found significantly higher

246 CI levels for the reward group on both days (main effect for group, day 1: $F = 6.81$, $p = 0.0127$;

247 day 2: $F = 9.10$, $p = 0.0044$; Figure 4c,e). However, unlike MT and peak velocity, no significant

248 increases in CI levels were observed for the no reward group on either day during post-R. This

249 suggests that coarticulation cannot be invigorated instantaneously but represents a training-

250 dependent process that is facilitated by reward. Importantly, CI levels were maintained during

251 post-NR on both days for the reward group, highlighting that changes in coarticulation had

252 become reward-independent (mixed-effect ANOVA; condition x group; main effect for group,

253 day 1: $F = 4.91$, $p = 0.0324$; day 2: $F = 7.38$, $p = 0.0097$; Figure 4d,f). To understand whether CI

254 levels are related to the retention of MT performance, we correlated MT values with CI levels

255 during post-NR (Figure 4g,h). We found a significant correlation between CI levels and MT

256 performance for both day 1 (partial correlation controlling for group; post-NR: $\rho = -0.51$; $p <$

257 0.0001, Figure 4g) and day 2 (partial correlation controlling for group; post-NR: $\rho = -0.66$, $p <$

258 0.0001, Figure 4h). Although not causal, this indicates that faster MTs during post-NR were

259 associated with higher levels of coarticulation and that the reward group showed better

260 performance for both (Figure 4g,h).

261   To summarise, these results demonstrate that improvements in MT are driven by two

262 processes which are both reward-sensitive but follow different time courses. The invigoration

263 of peak velocities is instantaneous, whereas coarticulation is training-dependent. Invigoration

264 in the no reward group was mainly driven by increases in peak velocities which were transient

265 in nature, thereby displaying a pronounced 'on-off' effect. In contrast, participants in the

266 reward group capitalised on both strategies: the invigoration of peak velocities and

267 additionally increases in coarticulation which led to persistent and ultimately reward-

268 independent improvements in MT performance. Importantly, despite an increase in

269 execution errors on day 1 (reward group), we found that execution errors were not associated

270 with performance levels during late training (day 2; Supplementary Figure 2). In addition, no

271 significant differences between groups were found on day 2 (mixed-effect ANOVA; condition

272 x group; main effect for group, day 2: $F = 1.88$, $p = 0.1785$), suggesting that performance gains

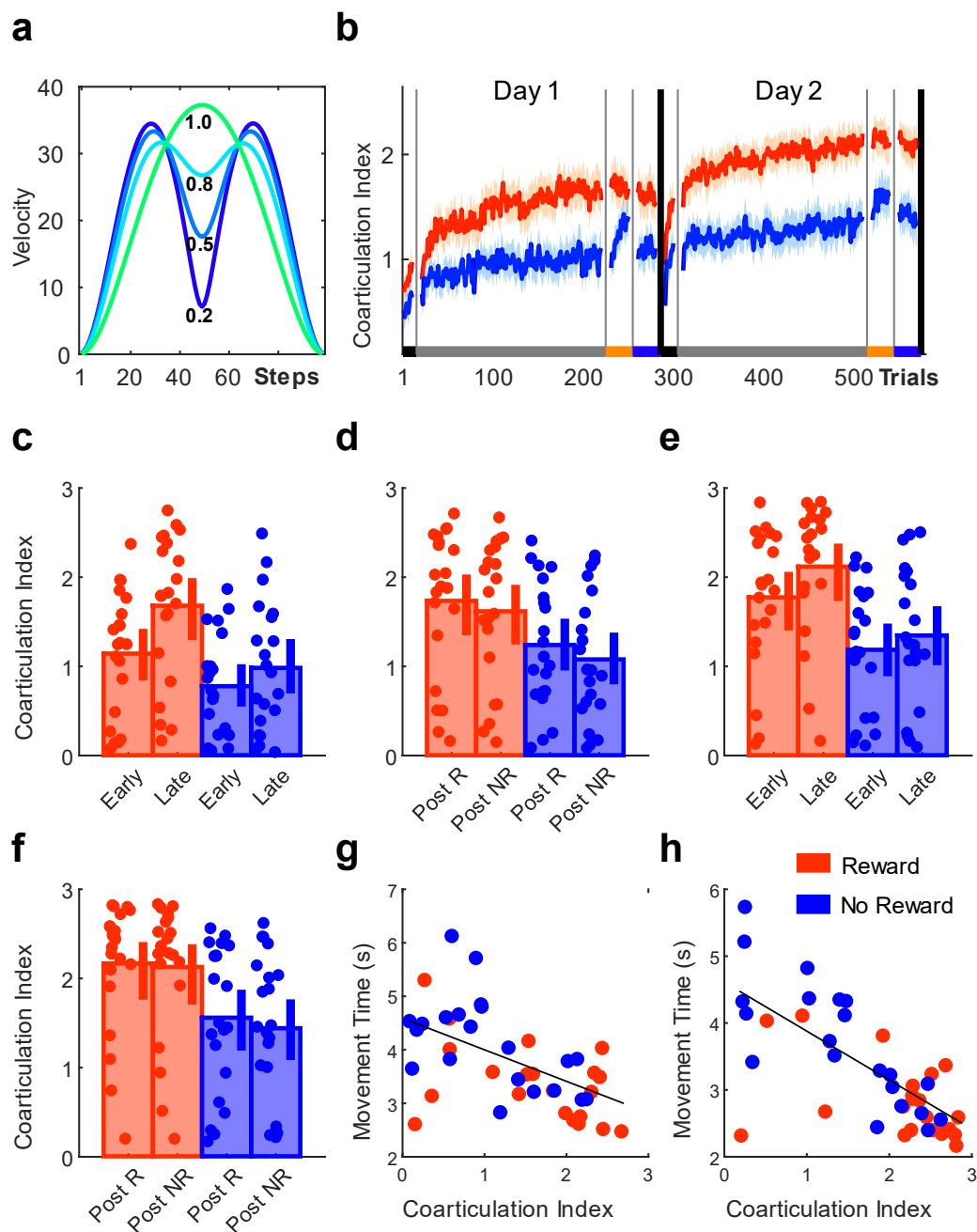273 by the second testing day reflected true improvements in skill.

10

274

275

Figure 4 | Reward-based improvements in CI levels. a) Illustration of CI levels. **b)** Trial-by-trial changes in CI levels averaged over participants for both groups. **c-f)** Median CI levels for each participant for both groups. **c)** Comparing CI levels during training on day 1 (early (first 20 trials) vs late (last 20 trials)). **d)** Post assessment performance (day 1; post-R vs post-NR). **e)** Training (day 2; early vs late). **f)** Post assessment performance (day 2; post-R vs post-NR). **g-h)** Scatterplots displaying the relationship between MT and CI levels during post-NR on **g)** day 1 and **h)** day 2 with a linear line fitted across groups.

283

284

11

285 **Reward-based improvements in smoothness are associated with reward-independent**

286 **maintenance of performance gains**

287   It has been suggested that coarticulation reduces motor costs via the reduction of jerk

288 and thereby reflects more efficient (smoother) execution [53,54], which in turn could explain why

289 participants in the reward group maintained their improved performance levels even in the

290 absence of reward. Using spectral arc length as a smoothness metric[55], we show that the

291 performance in the reward group became progressively smoother which could reflect

292 increased energetic efficiency (Figure 5a) [54].
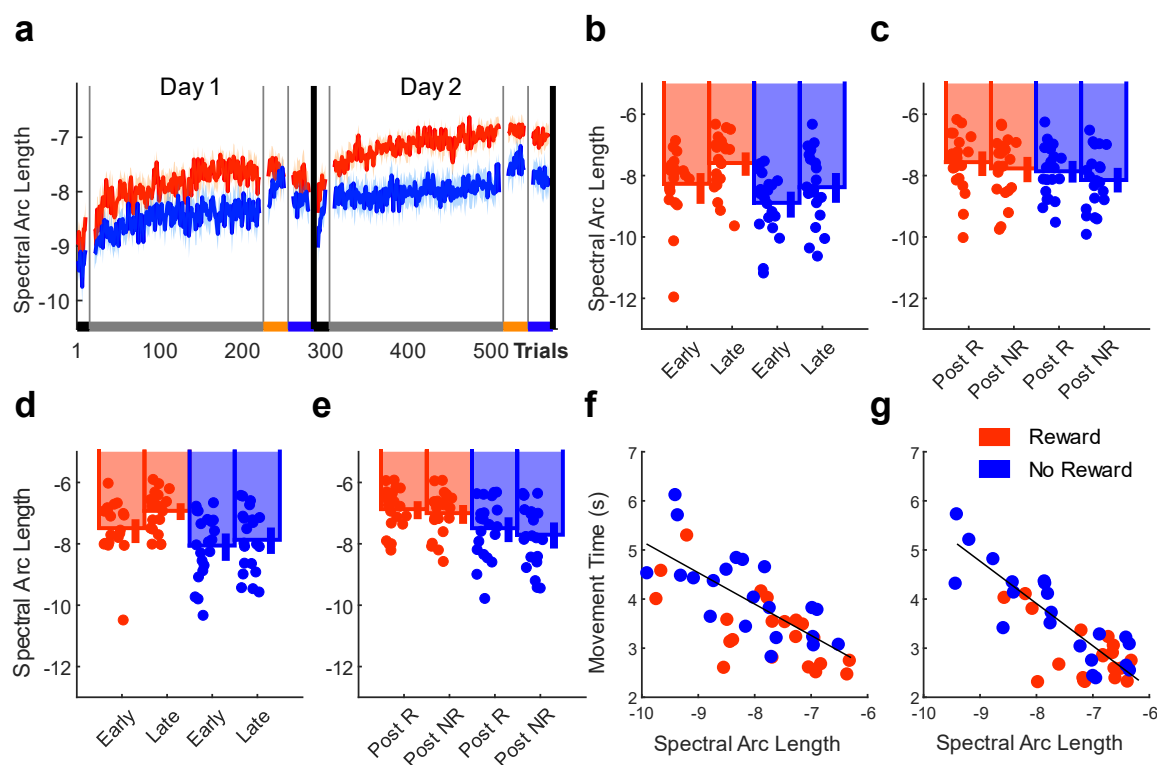
293

294
295

296 **Figure 5 | Reward-based improvements in smoothness. a)** Trial-by-trial changes in spectral arc length
297 (smoothness) averaged over participants for both groups. A value closer to zero indicates greater
298 smoothness[55]. **b-e)** Median smoothness values for each participant for both groups. **b)** Comparing
299 smoothness during training on day 1 (early (first 20 trials) vs late (last 20 trials)). **c)** Post assessment
300 performance (day 1; post-R vs post-NR). **d)** Training (day 2; early vs late). **e)** Post assessment
301 performance (day 2; post-R vs post-NR). **f-g)** Scatterplots displaying the relationship between MT and
302 smoothness during post-NR on **f)** day 1 and **g)** day 2 with a linear line fitted across groups.

303

304   Movement smoothness significantly improved across training on both days (mixed

305 ANOVA; timepoint x group; main effect for timepoint, day 1: $F = 27.56$, $p < 0.0001$; day 2: $F =$

306 18.19, $p < 0.0001$, Figure 5a). Despite there being no differences between groups during

307 baseline (Wilcoxon test; $Z = 1.79$, $p = 0.0741$), the reward group showed a greater

12

308　improvement in smoothness throughout training (main effect for group, day 1: F = 5.31, p =

309　0.027; day 2: F = 7.78, p = 0.0080; Figure 5b,d). During post assessments, we found a

310　significant group effect for day 2 (mixed-effect ANOVA; post assessment x group; group, day

311　1: F = 1.44, p = 0.2369; day 2: F = 6.41, p = 0.0154; Figure 5c,e). This suggests that

312　improvements in smoothness were greater in the reward group and became reward-

313　independent. Additionally, we found that increased smoothness was strongly associated with

314　faster MTs when reward was not available. We correlated smoothness values with MTs during

315　post-NR on both days (Figure 5f,g). We found a significant correlation between smoothness

316　and MT performance for both day 1 (partial correlation controlling for group; Post-NR: ρ = -

317　0.69 ; p < 0.0001, Figure 5f) and day 2 (partial correlation controlling for group; post-NR: ρ = -

318　0.79, p < 0.0001, Figure 5g). These results align with the aforementioned results showing that

319　CI levels were associated with faster MTs during post-NR, and point to improvements in

320　movement efficiency as a potential mechanism enabling reward-independent maintenance

321　of performance gains during post-NR.

322

323　Progressive alignment to the predictions of a minimum jerk model is facilitated by reward

324　　　　We then assessed whether performance aligned with the predictions of a minimum

325　jerk model [12] by calculating the mean squared error between the model's and the actual

326　velocity profile on a trial-by-trial basis (Methods, Equation 3; Figure 6a). In comparison to the

327　no reward group, performance in the reward group became significantly more aligned to the

328　predictions of the minimum-jerk model. Mean squared error progressively decreased across

329　days although this was not significant for day 2 (mixed ANOVA; timepoint x group; main effect

330　for timepoint, day 1: F = 16.19, p < 0.0001; day 2: F = 2.23, p = 0.1429, Figure 6b). Despite no

331　differences between groups during baseline (Wilcoxon test; Z = -1.16, p = 0.2472), the reward

332　group's performance showed significantly greater similarity to the model's predictions (main

333　effect for group day 1: F = 5.61, p = 0.0228; day 2: F = 7.56, p = 0.0089, Figure 6c,e). Across

334　post assessments, we found a significant interaction for day 2 (mixed-effect ANOVA; condition

335　x group; main effect for interaction, for day 1: F = 3.92 p = 0.0548; day 2: F = 7.58, p = 0.0088,

336　Figure 6b,f), however all post-hoc comparisons were not significant. Nevertheless, there was

337　a significant group effect for day 2 (main effect for group, for day 1: F = 3.26 p = 0.0787; day

338　2: F = 4.98, p = 0.0313, Figure 6b,f), suggesting that the degree of similarity to the minimum-

339　jerk model was maintained across post assessments irrespective of reward availability.
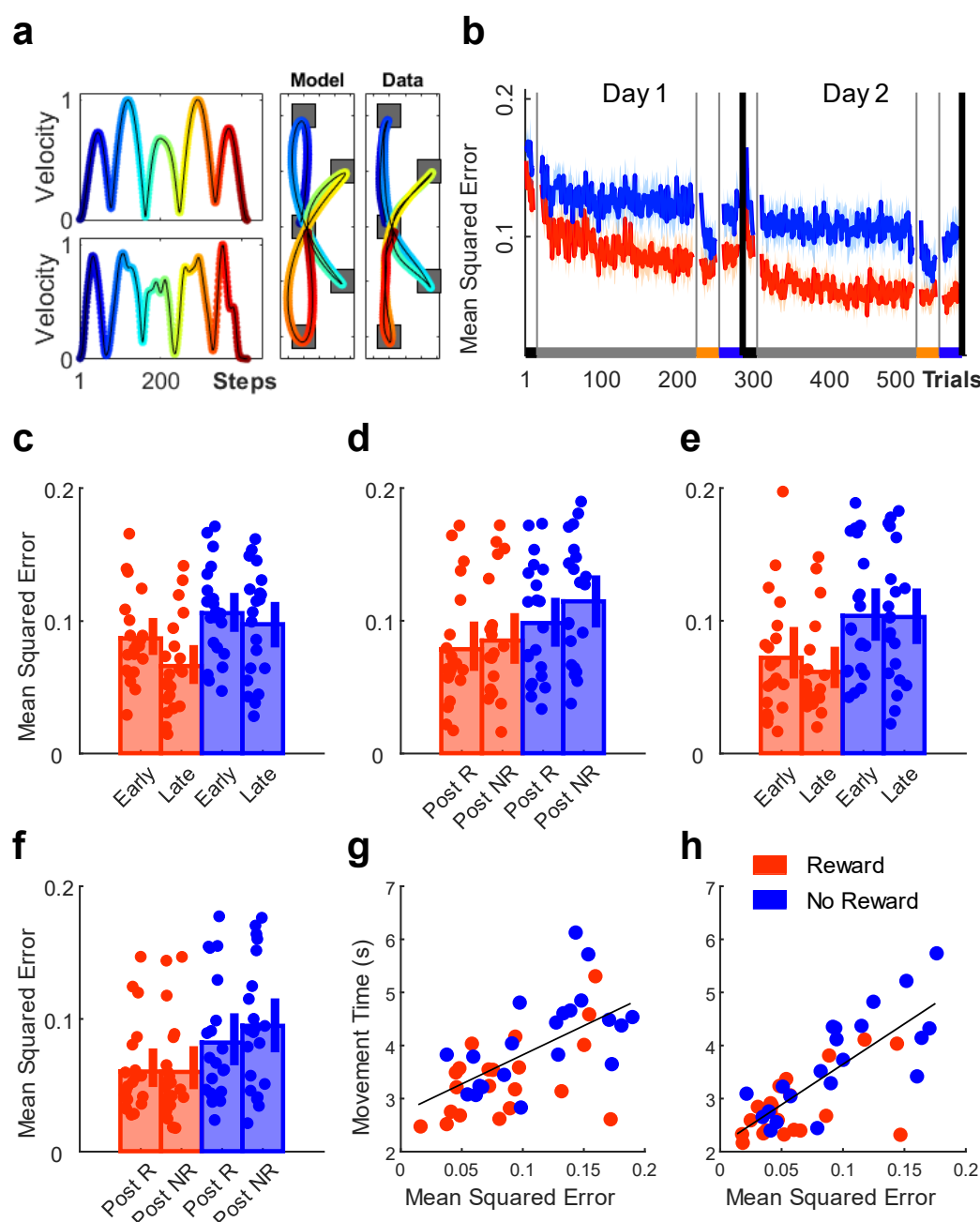
13

Figure 6 | Progressive alignment to the predictions of a minimum jerk model is facilitated by reward a) Comparisons between data and the predictions of a minimum jerk model for both trajectory (right panel) and velocity profiles (left panel) for a single participant. **b)** Trial-by-trial changes in mean square error averaged over participants for both groups. **c-f)** Median mean square error for each participant in both groups. **c)** Day 1 early vs late training **d)** Day 1 post assessment (post-R vs post-NR). **e)** Day 2 early vs late training **f)** Day 2 post assessment. **g-h)** Scatterplots displaying the relationship between MT and mean squared error during post NR on **g)** day 1 and **h)** day 2 with a linear line fitted across groups.

Within this context, we aimed to assess whether more optimal performance with regards to the model prediction was related to the maintenance of performance. Using partial linear correlations, we found a correlation between model similarity and MT performance

14

353   during post-NR for both day 1 (partial correlation controlling for group; Post-NR: ρ = 0.55 ; p

354   < 0.0001, Figure 6g) and day 2 (partial correlation controlling for group; post-NR: ρ = 0.70, p

355   < 0.0001, Figure 6h). These results complement our findings on smoothness and CI. Overall,

356   our results show that over the course of the experiment performance in the reward group

357   became smoother and more optimal with regards to the predictions of a minimum jerk model.

358   Potentially due to being more energetically efficient [56], participants in the reward group

359   maintained MT performance gains even in the absence of reward.

360

361   **Spatial reorganisation identifies the final stages of coarticulation and is enhanced by reward**

362   Coarticulation can also be expressed spatially as the radial distance between the peak

363   velocity on the sub-movements and the minimum velocity around the via point (Figure 7a).

364   This distance becomes smaller with high levels of coarticulation [9,10] (Figure 7a; Supplementary

365   Figure 3), suggesting that spatial reorganisation reflects the final stages of two movements

366   merging together. We found that the reward group expressed spatial reorganisation with

367   significant decreases in the radial distance between peaks and the via point. No difference

368   between groups in radial distance was observed during baseline (Wilcoxon test; Z = -1.18, p =

369   0.2371). However, there was significant main effects for timepoint (mixed ANOVA; timepoint

370   x group; main effect for timepoint, day 1: F = 7.32, p = 0.0100; day 2: F = 18.14, p < 0.0001;

371   Figure 7b) as well as group for the second testing day (main effect for group, day 1: F = 1.95,

372   p = 0.1706; day 2: F = 7.52, p = 0.0091; Figure 7c,e). This indicates that radial distance

373   progressively decreased across both days, with the reward group showing greater changes on

374   day 2. Across post assessments, the no reward group showed no reward-driven changes in

375   spatial reorganisation, whereas the reward group maintained their performance although the

376   differences between groups was only significant on day 2 (mixed-effect ANOVA; post

377   assessment x group; main effect for group, day 1: F = 3.08, p = 0.0868; day 2: F = 5.76, p =

378   0.0211; Figure 7d,f).

379   To understand the relationship between CI and spatial reorganisation, we plotted

380   them against each other and detected a pronounced drift in radial distance values (%) with

381   increasing CI levels resulting in a curvilinear shape (Figure 7g). After fitting a two-segment

382   piecewise linear function to the data, we found an inflection point at ~1.66 (CI) and a strong

383   correlation between CI levels and radial distance values for the second segment (partial

384 correlation controlling for group; segment 1: ρ = -0.16; p < 0.0001; segment 2: ρ = 0.89; p <
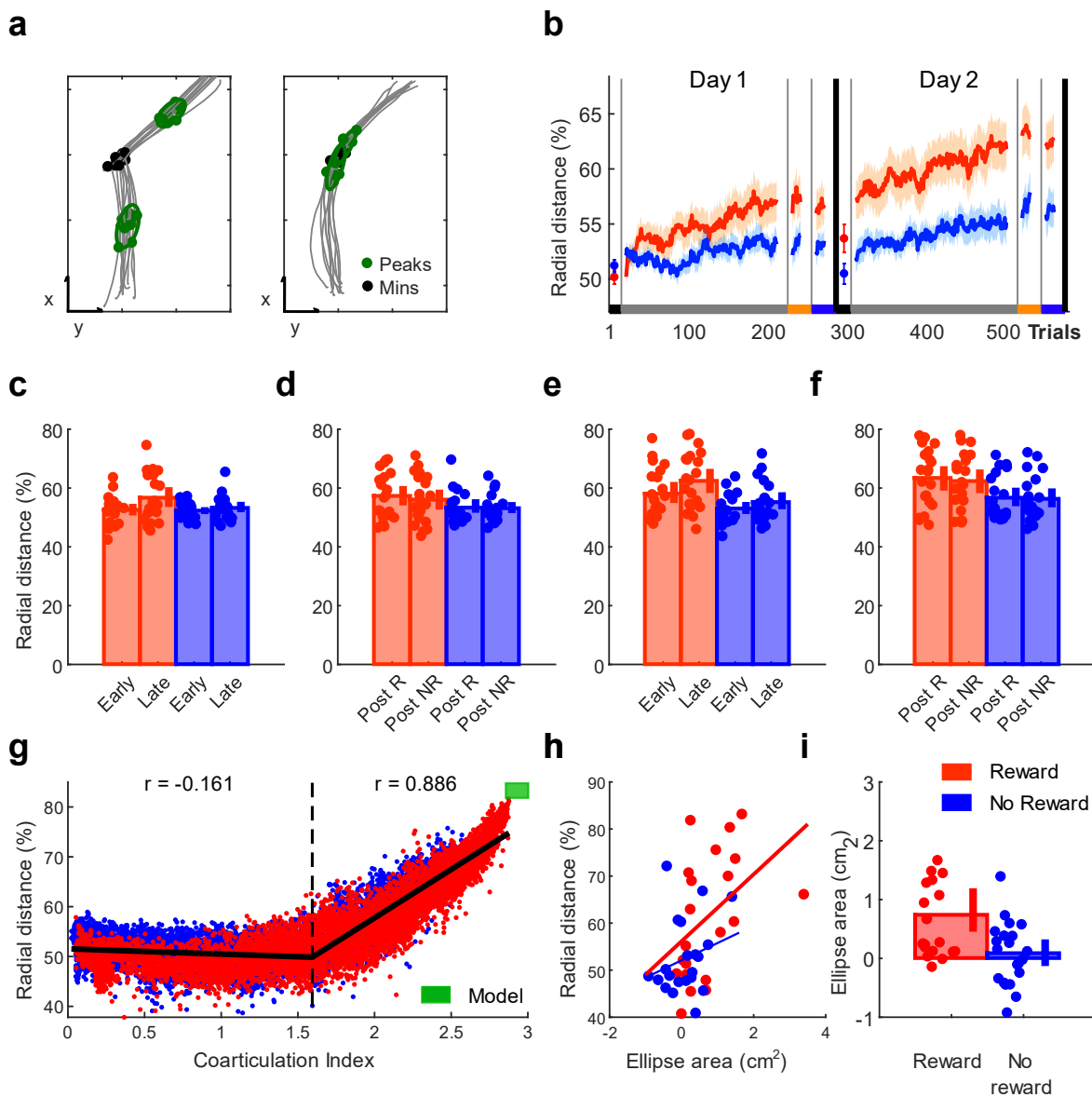
385 0.0001).

Figure 7| Reward-based improvements in spatial reorganisation. a) Example data for the spatial location of peak velocities when performing two individual (left panel) and one coarticulated (right panel) reaching movement. **b)** Trial-by-trial changes in radial distance averaged over participants for both groups. **c-f)** Median radial distance for each participant for both groups. **c)** Comparing radial distance during training on day 1 (early (first 20 trials) vs late (last 20 trials)). **d)** Post assessment performance (day 1; post-R vs post-NR). **e)** Training (day 2; early vs late). **f)** Post assessment performance (day 2; post-R vs post-NR). **g)** Scatterplot illustrating the relationship between mean CI levels and spatial reorganisation (radial distance %). It includes a two-segment piecewise linear function fitted to the data, while the green area represents the model prediction. **h)** Scatter plot displaying the relationship between variability (ellipse area cm^2) during early training and spatial coarticulation (radial distance %) during late training. **i)** Bar plot comparing the reward and no reward group early training variability including jittered scatter of participant-based median (ellipse area cm^2).

386

16

387   This suggests that in order to fully coarticulate two consecutive movements, spatial

388   reorganisation is essential [9–11] with this process appearing significantly more pronounced in

389   the reward group. Overall, these findings indicate that improvements in movement efficiency

390   through a quantitative change in how the task is performed may enable the retention of

391   reward-based performance gains.

392   Within the context of reinforcement learning, it has been suggested that reward leads

393   to improvements in performance via increases in exploration during the early stages of

394   training [57,58]. To assess whether an early increase in spatial variability was associated with

395   improved performance, we measured the area of the confidence ellipses used to determine

396   the radial distance to the via point. We then correlated the ellipse area (cm^2), which was

397   normalised to baseline, over the first 30 trials on day 1 with radial distance (%) over the last

398   30 trials on day 2 (Figure 7h). We found a significant correlation (partial correlation controlling

399   for group; $\rho = 0.40$; $p = 0.0104$) suggesting that early increases in spatial variability were

400   associated with increases in spatial coarticulation towards the end of the experiment.

401   Comparing ellipse area between groups during early training (first 30 trials), we found that

402   the reward group exhibited higher levels of variability (Wilcoxon test, $Z = 2.52$, $p = 0.0119$,

403   Figure 7i). These results indicate that early reward-driven increases in spatial variability may

404   benefit future performance.

405

Performance gains are maintained across an additional testing day without reward

407   We next aimed to assess the robustness of these performance gains in an experiment

408   including an additional testing day without reward availability (elongated washout condition).

409   In experiment two, participants underwent the same regime as the reward group in

410   experiment 1 on the first two days. On the third day, participants were asked to complete 200

411   no reward trials. Our results show that even after 24 hours, and over the course of 200

412   additional unrewarded trials, participants maintained similar MT performance levels. We

413   used a repeated measures ANOVA with timepoint (early vs late across all testing days) as the

414   within factor to assess changes across testing days (repeated measures ANOVA, main effect

415   for timepoint, $F = 28.65$, $p < 0.0001$; Figure 8a; Supplementary Figure 4). These results indicate

416   that performance improved over the course of the experiment. However, no changes in MT

417   performance could be observed between late training on day 2 and early training on day 3

418   (Wilcoxon test, $Z = -1.21$, $p = 0.3016$) and between early and late training on day 3 (Wilcoxon

419 test, Z = -1.48 , p = 0.4444). Similarly, coarticulation levels appeared stable across the

420 additional testing day without reward (repeated measures ANOVA, main effect for timepoint,

421 F = 19.19 p < 0.0001; Figure 8b; Supplementary Figure 4), with no changes in performance

422 between late day 2 and early day 3 (Wilcoxon test, Z = -0.94, p = 1). Similarly to MT, no changes

423 were observed across day 3 (Wilcoxon test; early vs late; day 3, Z = -0.67, p = 1). Yet again, we

424 found that peak velocities were persevered transitioning to and across day 3 (Wilcoxon test,

425 late training day 2 x early training day 3, Z = -1.21, p = 1; early training day 3 x late training

426 day 3, Z = -1.21, p = 1; Figure 8c; Supplementary Figure 4). Taken together, these findings

427 highlight that performance gains were preserved even after an additional testing day without

428 reward availability.

429   When assessing changes in smoothness using spectral arc length, we found similar

430 results to experiment 1. Smoothness improved over the course of the experiment (repeated

431 measures ANOVA, main effect for timepoint, F = 57.14 p < 0.0001; Figure 8d; Supplementary

432 Figure 4), while no changes could be observed between late training on day 2 and early

433 training on day 3 (Wilcoxon test, Z = -0.67, p = 1) and between early and late training on day

434 3 (Wilcoxon test, Z = -1.21 , p = 1). In line with these results, we found that performance

435 aligned progressively with the predictions of the minimum jerk model (repeated measures

436 ANOVA, main effect for timepoint, F = 23.38 p < 0.0001; Figure 8e; Supplementary Figure 4),

437 while no significant changes in similarity could be observed transitioning to and across day 3

438 (Wilcoxon test, late training day 2 x early training day 3, Z = -2.02, p = 1; early training day 3 x

439 late training day 3, Z = -1.21, p = 1). In addition, we found that participants progressively

440 reorganised their spatial movement output, with a significant decrease in radial distance

441 found across the experiment (repeated measures ANOVA, main effect for timepoint, F = 57.14

442 p < 0.0001; Figure 8f; Supplementary Figure 4). However, no changes in radial distance could

443 be observed between late training on day 2 and early training on day 3 (Wilcoxon test, Z = -

444 2.02, p = 1) and between early and late training on day 3 (Wilcoxon test, Z = -1.75 , p = 1).

445 These findings support our results from experiment 1 and indicate that improvements in

446 movement efficiency through a quantitative change in how the task is performed may enable

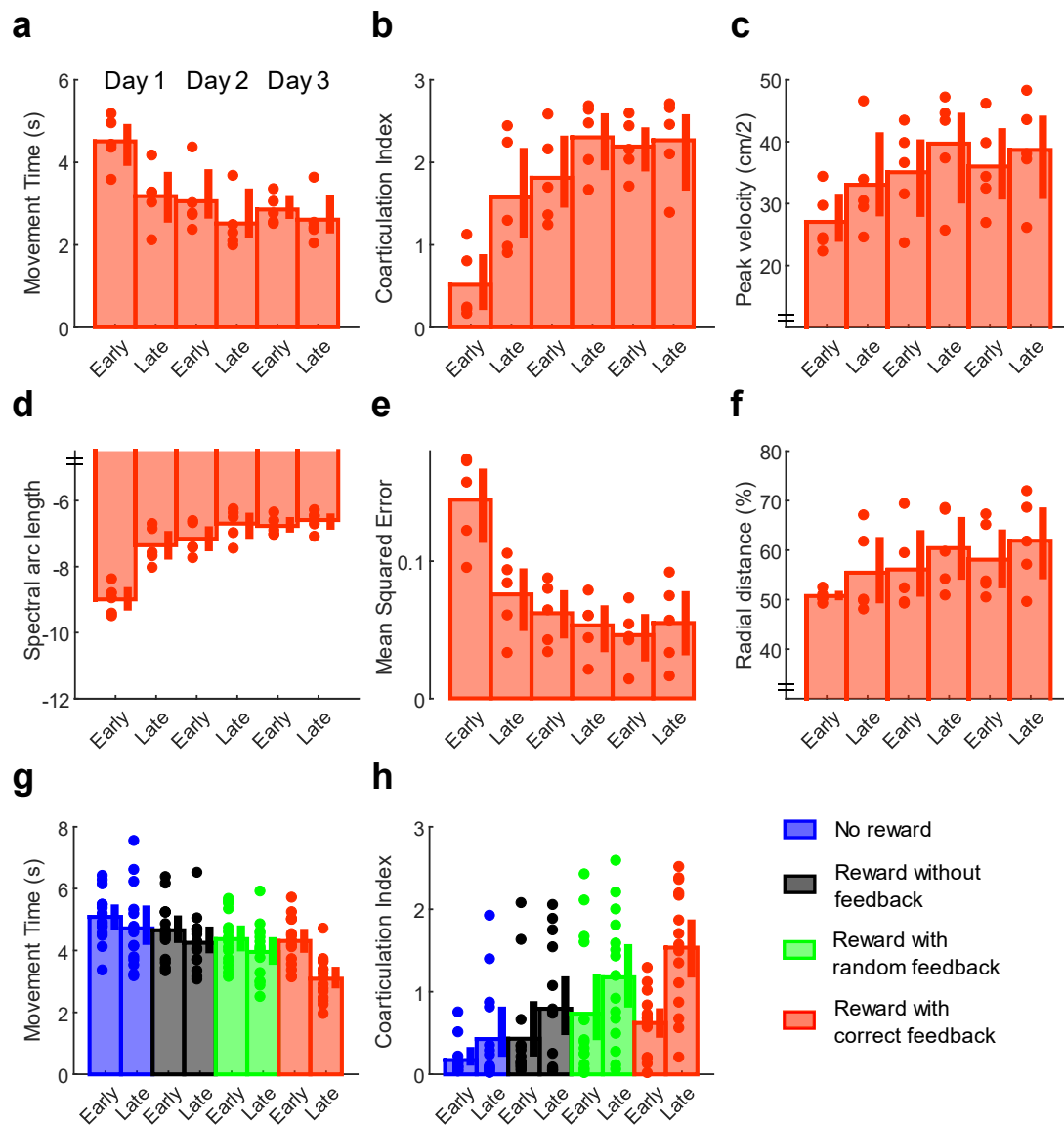447 long-term retention of reward-based performance gains.

448

Figure 8| Long-term maintenance of performance without reward. a-f: Experiment 2. Median data across all 3 days (early (first 20 trials) vs late (last 20 trials) training). **a)** MT **b)** CI level **c)** Peak velocity **d)** Spectral arc length (smoothness) **e)** Mean squared error between data and minimum jerk model prediction **f)** Radial distance (spatial reorganisation) g-**h: Experiment 3. g)** Median MT across groups. **h)** Median CI level.

Reward based on performance is most effective in producing behavioural change

Our third experiment was intended to assess whether the observed changes on day 1 were specific to training with reward-based feedback of performance. Participants were allocated to one of the four groups: (1) no reward, (2) reward without performance-based feedback, (3) reward with random feedback and (4) reward with accurate feedback. Only groups 2-4 received monetary reward; reward for groups 1 and 4 were equivalent to that

457   used in experiment 1. Participants in group 3 received feedback about the reward delivered

458   on each trial which was randomly drawn from the feedback experienced during experiment

459   1. It therefore matched group 4 in terms of reward probability but this did not correspond to

460   actual performance. In contrast, participants in group 2 did not receive any performance-

461   based feedback after completing a given trial. Upcoming reward trials, however, were still

462   cued. Since groups 1 and 4 underwent the same regime as the no reward and reward group

463   in experiment 1 respectively, we were able to test whether our results replicated. Group 4

464   produced the largest improvements in performance, being the only group who showed

465   significant improvements in MT (mixed ANOVA; timepoint x group; interaction, F = 3.76, p =

466   0.0158; Wilcoxon test; early vs late group 4: Z = 3.40, p = 0.0028; Figure 8g; Supplementary

467   Figure 5) and coarticulation (mixed ANOVA; timepoint x group; interaction, F = 3.18, p =

468   0.0308; Wilcoxon test; early vs late group 4: Z = -3.24, p = 0.0048; Figure 8h; Supplementary

469   Figure 5). This indicates that the expectation of reward is not sufficient to induce large

470   behavioural improvements with a link to performance being essential to optimise gains.

471   These results showed that performance-based reward invigorates sequential

472   reaching. Driven by a reward-based increase in speed, movements also exhibited greater

473   coarticulation, smoothness and a closer alignment to the predictions of a minimum jerk

474   model. Importantly, these performance gains were maintained across multiple days even

475   after the subsequent withdrawal of reward. This highlights the importance of coarticulation

476   to skilful sequential reaching performance, and the potential of this mechanism to produce

477   long-lasting reward-driven improvements in behaviour.

478

479   **Discussion**

480   Motor skill learning is integral to everyday life and describes improvements in

481   performance above baseline levels [59]. Improvements in skill have often been assessed at the

482   levels of speed, accuracy and more recently efficiency [60]. In our task, improvements in speed

483   (MT) could be achieved via two strategies: (1) increases in peak velocities of the individual

484   reaching movements and (2) reduction in dwell times around the via points (coarticulation).

485   We were able to show that reward invigorated peak velocities, supporting previous findings

486   on the reward-based invigoration of simple discrete reaching movements [41,42]. Moreover, for

487   the first time, we demonstrate that reward accelerates the slow learning process associated

488   with coarticulation [9–11,31]. With respect to movement efficiency, these two strategies differ

489   regarding the metabolic costs they incur. Recently, it has been demonstrated that in discrete

490   reaching movements, improvements in peak velocity can be achieved at no cost to accuracy

491   through increased arm stiffness [41]. Although an attractively simple strategy, it comes with a

492   marked escalation in metabolic costs [56]. In fact, it has been shown that in discrete reaching

493   tasks such invigoration is reward-dependent and transient in nature [40–42]. We suggest that

494   this could be due to the energetic demands of the invigorated movement requiring the

495   continued presence of reward to negate this added cost. In contrast, coarticulation enables

496   the performance of sequential movements to become similar to a minimum jerk trajectory

497   through the merging of neighbouring movements into singular smooth elements [9–12,53].

498   Increases in smoothness have been shown to reduce metabolic costs, thereby enhancing

499   overall movement efficiency [54]. Here we showed that performance in the reward group was

500   smoother and exhibited greater similarity to a minimum jerk trajectory. The subsequent

501   hypothesised improvement in energetic efficiency might explain why the enhanced

502   performance of the reward group became reward-independent, whereas performance in the

503   no reward group, who showed little coarticulation, followed the previously observed

504   transient 'on-off' effect.

505   Despite previous studies demonstrating that reward can enhance retention across a

506   wide variety of sequential and continuous motor tasks, the underlying mechanism for this

507   effect has been described at a very abstract level [50–52]. Specifically, it is unclear how reward

508   strengthens a motor memory so that improved performance is maintained even when the

509   incentive is no longer provided. We believe that for reward to induce such long-term

510   improvements in motor skill, it must not only lead to enhanced performance but also

511   improvements in efficiency. Reduction in the cost of this enhanced performance (faster or

512   more accurate movements [40]), enables it to be performed long-term without further

513   incentive/reward. Therefore, reward may not enhance the memory of the action but instead

514   lead to the task being performed in a fundamentally different, and more efficient, manner. In

515   support of this, we found that participants in the reward group produced a quantitative

516   change (spatial reorganisation of the velocity profile between neighbouring movements) in

517   how the task was executed. This highlights that careful analysis of how reward influences the

518   performance of complex actions is essential for dissociating its effects on execution and

519   retention.

520   Although coarticulation might appear to feature similar movement characteristics as

521   chunking, it is important to emphasise that they represent different processes. Chunking

522   often refers to a series of discrete movements (e.g. button presses in sequence learning

523   paradigms) which are temporally aligned (chunked) over the course of training [7,26–29]. The

524   chunks are represented at a behavioural level through shorter reaction times between

525   actions, and often allow for a faster execution overall [7,8]. Yet the elements within a chunk are

526   still performed discretely with a clear stop period between them [22,29]. In addition, at a neural

527   level, these elemental movements are planned individually through competitive queuing,

528   highlighting again that some level of independence is maintained [61]. In contrast, coarticulation

529   reflects the merging of neighbouring movements into a single motor primitive that no longer

530   features a pronounced stop period and must run to completion once initiated [9–11,31]. While

531   kinematically appearing to represent a single movement, it remains an open question how

532   these newly formed primitives are represented at a neural level in humans.

533   Recent work in animals suggests that sequence learning is a dopamine-dependent

534   process [62,63]. By tracking dopamine concentration changes in the nucleus accumbens core of

535   rats, it has been shown that dopamine levels dynamically change with task proficiency [62].

536   Further evidence originates from research in Parkinson's disease (PD) in which PD patients

537   OFF medication exhibit specific impairments in movement chunking [32,33]. However, whether

538   dopamine underpins coarticulation similarly to chunking is still unexplored (a notable

539   exception being [33]). Specifically, whether dopamine directly acts on coarticulation or whether

540   any mediating relationship is due to the effects of dopamine on movement invigoration,

541   requires further examination [33].

542   Here, we show that coarticulation is associated with smoother and, with regards to

543   minimising jerk, more efficient execution. Interestingly, reaching movements performed by

544   stroke patients exhibit reduced smoothness [34–37], with increases in jerk being due to a

545   decomposition of a movement into a series of sub-movements [34–37]. However, over the course

546   of the recovery process, performance becomes smoother as these sub-movements are

547   progressively blended [34–37]. Considering this theoretical proximity to the concept of

548   coarticulation, we speculate that stroke recovery and coarticulation may follow similar

549   principles. Consequently, coarticulation facilitated by reward could be a powerful tool in

550   stroke rehabilitation to promote smooth and efficient sequential actions which form an

551   essential component of everyday life activities.

552   In conclusion, this work highlights that coarticulation provides a mechanism by which

553   reward can invigorate sequential performance whilst also improving efficiency. This

554   improvement in efficiency through a quantitative change in how the action is performed

555   appears essential for the retention of reward-based improvements in motor behaviour.

556

Methods

558   **Participants:** 107 participants (17 males; age range 18 - 35) were recruited to participate in three

559   experiments, which had been approved by the local research ethics committee of the University of

560   Birmingham. All participants were novices to the task paradigm and were free of motor, visual and

561   cognitive impairment. Most participants were self-reportedly right-handed (N = 7 left-handed

562   participants) and gave written informed consent prior to the start of the experiment. For their

563   participation, participants were remunerated with either course credits or money (£7.5/hour) and

564   were able to earn additional money during the task depending on their performance. Depending on

565   the experiment, participants were pseudo-randomly allocated to one of the available groups.

566   **Experimental Apparatus:** All experiments were performed using a Polhemus 3SPACE Fastrak tracking

567   device (Colchester, Vermont U.S.A; with a sampling rate of 110Hz). Participants were seated in front

568   of the experimental apparatus which included a table, a horizontally placed mirror 25cm above the

569   table and a screen (Figure 1a). The low-latency Apple Cinema placed 25cm above the mirror had a

570   refresh rate of 60Hz and displayed the workspace and participants' hand position (represented by a

571   green cursor – diameter 1cm). On the table, participants were asked to perform 2-D reaching

572   movements. Looking into the mirror, they were able to see the representation of their hand position

573   reflected from the screen above. This setup effectively blocked their hand from sight. The experiment

574   was run using MATLAB (The Mathworks, Natwick, MA), with Psychophysics Toolbox 3.

575   **Task design:** Participants were asked to hit a series of targets displayed on the screen (Figure 1c). Four

576   circular (1cm diameter) targets were arranged around a centre target ('via target'). Starting in the via

577   target, participants had to perform eight continuous reaching movements to complete a trial. Target

578   1 and 4 were displaced by 10cm on the *y*-axis, whereas Target 2 and 3 were 5cm away from the via

579   target with an angle of 126 degrees between them (Figure 1c). Our task design was based on previous

580   work [9–11,31] in which the authors were able to observe coarticulation using similar angles and reaching

581   distance configurations. To start each trial, participants had to pass their cursor though the

582   preparation box (2x2cm) on the left side of the workspace, which triggered the appearance of the

23

583   start box (2x2cm) in the centre of the screen. After moving the cursor into the start box, participants

584   had to wait for 1.5s for the targets to appear. This ensured that participants were stationary before

585   reaching for the first target. Target appearance served as the go-signal and the start box turned into

586   the via target (circle). Upon reaching the last target (via target), all targets disappeared, and

587   participants had to wait for 1.5s before being allowed to exit the start box to reach for the preparation

588   box to initiate a new trial. Participants had to repeat a trial if they missed a target or performed the

589   reaching order incorrectly. Similarly, exiting the start box too early either at the beginning or at the

590   end of each trial resulted in a missed trial.

591   **Reward Structure and Feedback.** Participants in experiment 1 and experiment 2 experienced either

592   reward or no reward trials depending on the current experimental phase: (1) Reward trials were cued

593   using a visual stimulus prior to the start of the trial. Once participants moved into the preparation box,

594   the start box appeared in yellow (visual stimulus) rather than in black (Figure 1f). Participants were

595   informed that faster MTs will earn them more money, with a maximum amount of 5p available in each

596   trial. While participants moved from the start box to the preparation box to initiate a new trial, the

597   amount earned in the previous trial was displayed on the top of the screen (i.e. 'You have earned 2p

598   out of 5p'). We used a closed-loop design to calculate the amount of reward earned in each trial. To

599   calculate this, we included the MT values of the last 20 trials and organised them from fastest to

600   slowest to determine the rank of the current trial within the given array. A rank in the top three (<=

601   90%) returned a value of 5p, ranks >= 80% and <90% were valued at 4p; ranks >=60% and <80% were

602   awarded 3p; ranks >=40% and < 60% earned 2p while 1p was awarded for ranks >=20% and < 40%. A

603   rank in the bottom three (<20%) returned a value of 0p. When participants started a new experimental

604   block, performance in the first trial was compared to the last 20 trials of the previously completed

605   block. (2) No reward trials were not cued, and no reward was available for participants. However,

606   participants were instructed to 'move as fast and accurately as possible'. In experiment 3, participants

607   were randomly allocated to one of the four groups: (1) no reward, (2) reward without performance-

608   based feedback, (3) reward with random feedback and (4) reward with accurate feedback. Participants

609   in the no reward (1) and reward with accurate feedback (4) groups underwent the same regime as the

610   no reward and reward groups in experiment 1 respectively. To investigate whether reward and/or

611   feedback drove performance changes, we changed the reward and feedback structure for the two

612   remaining groups. Participants in the reward without performance-based feedback group (2) were

613   able to earn money depending on their performance in each trial. Reward trials were cued with a

614   yellow start box prior to the start of the trial. However, participants did not receive any feedback on

615   their performance after completing a given trial. They were asked to initiate a new trial and 'be as fast

616   and accurate as possible to earn more money'. In contrast to this, participants in the reward with

617   random feedback group (3) received feedback after completing a reward trial during training, which

618   were also cued with a yellow start box. However, feedback in this group was not performance-based,

619   but was drawn randomly from feedback given to participants in experiment 1. To this end, we strung

620   together all reward values given to participants in the first experiment and randomly chose a value for

621   feedback in a given trial in experiment 3. Participants, therefore, received feedback which was similar

622   in reward probability without corresponding to actual performance.

Experimental Procedure

Experiment 1

625   In this experiment, we investigated whether reward can invigorate performance on a sequential

626   reaching task. The experiment included an initial learning phase prior to the start of the experiment

627   as well as a baseline, training and two post assessments. The same design without the learning phase

628   was repeated 24 hours later (Figure 1b). Participants were pseudo-randomly allocated to either the

629   reward or no reward group (N = 21 each) and were informed that at some point during the experiment

630   they would be able to earn additional money depending on their performance.

631   *Learning*: We included a learning phase prior to the start of the experiment for participants to be able

632   to memorise the reaching sequence. This allowed us to attribute any performance gains to

633   improvements in execution rather than memory. Once participants waited 1.5s inside the start box,

634   the targets appeared which were numbered clockwise from 1 to 4 starting with the central top target.

635   Participants were also able to see a number sequence at the top left of the screen displaying the order

636   of target reaches (1 – 3 – 2 - 4). Participants were instructed to hit the targets according to the number

637   sequence while also hitting the via target in between target reaches. They had to repeat a trial if they

638   missed a target or performed the reaching order incorrectly. Similarly, exiting the start box too early

639   either at the beginning or at the end of each trial resulted in a missed trial. After a cued trial,

640   participants were asked to complete a trial from memory without the number sequence or numbers

641   inside the targets. If participants failed a no cue trial more than twice, cues appeared in the following

642   trial as a reminder. After a maximum of 10 cue and 10 no cue trials participants completed this block.

643   *Baseline*: Participants in both groups completed 10 baseline trials, which were used to assess whether

644   there were any pre-training differences between groups. Both groups were instructed to 'move as fast

645   and accurately as possible', while no performance-based feedback was given at the end of each trial.

646   *Training:* Participants in the reward group were informed that during this part they would be able to

647   earn money depending on how fast they complete each trial (200 reward trials). In contrast,

648   participants in the no reward group engaged in 200 no reward trials and were again instructed to

649   move as fast and as accurately as possible.

650 *Post assessments*: On both testing days, participants from both groups were asked to complete two

651 post assessments (20 trials each); one with reward trials (post-R) and one with no reward trials (post-

652 NR). The order was counter-balanced across participants.

653 **Experiment 2**

654 In this experiment, we aimed to test how robust reward-driven performance gains were over an

655 additional testing day without reward availability. Participants (N = 5) underwent the same regime as

656 the reward group in experiment 1 on the first two testing days. On the third testing day after baseline,

657 participants were asked to complete 200 no reward trials.

658 **Experiment 3**

659 Here, we investigated whether the observed performance gains depend on reward expectation,

660 performance-based feedback or a combination of both. To this end, we allocated participants (N = 60)

661 to one of the four groups: (1) no reward, (2) reward without performance-based feedback, (3) reward

662 with random feedback and (4) reward with accurate feedback (see Reward Structure and Feedback

663 for more information). Participants underwent the same procedure as participants in experiment 1

664 over the course of the first testing day. After a learning phase and a baseline part, participants engaged

665 in 200 training trials which differed with regards to their reward and feedback structure. Similarly, to

666 experiment 1 and 2, participants then completed the post assessments, which were counter-balanced

667 across participants.

668 **Data Analysis**

669 Analysis code is available on the Open Science Framework website, alongside the experimental

670 datasets at https://osf.io/62wcz/. The analyses were performed in Matlab (Mathworks, Natick, MA).

671 **Movement Time (MT):** MT was measured as the time between exiting the start box and reaching the

672 last target. This excludes reaction time, which describes the time between target appearance and

673 when the participants' start position exceeded 2cm. Trials with MTs beyond 9.0s were excluded from

674 analysis, which amounted to 0.37% of all trials.

675 **Peak Velocity:** Through the derivative of positional data (*x, y*), we obtained velocity profiles for each

676 trial which were smoothed using a gaussian smoothing kernel (σ = 2). The velocity profile was then

677 divided into segments representing movements to each individual target (8 segments) by identifying

678 when the positional data was within 2cm of a target. We measured the peak velocity (v_{peak}) of each

679 segment by finding the maximum velocity:

$$v_{peak} \triangleq max_{\epsilon\,[t_1\,t_2]}\,v(t) \qquad\qquad \text{Equation 1}$$

681 Where *v(t)* is the velocity of segment *t*, and *t_1 and t_2* represent the start and end of segment *t*

682 respectively.

683   **Coarticulation Index (CI):** Coarticulation describes the blending together of individual motor elements

684   into a singular smooth action. This is represented in the velocity profile by the stop period between

685   the two movements gradually disappearing and being replaced by a single velocity peak (Figure 1c) [9–

686   11]. Based on the predictions of a model which minimised jerk (Equation 3; Figure 1d), we hypothesised

687   that coarticulation would mainly occur between out-centre centre-out reaching movements (Figure

688   1e). We therefore excluded the first and last target reach from this analysis. To measure coarticulation,

689   we compared the mean peak velocities of two sequential reaches with the minimum velocity around

690   the via point. The smaller the difference between these values, the greater coarticulation had

691   occurred between the two movements (Figure 2c) [25]. The velocity profile was cut into 3 segments

692   depending on the peak velocities of the out-centre and centre-out reaching movements. The

693   minimum velocity of these segments was calculated and compared to the average of the peak

694   velocities:

695

$$Coarticulation\ Index \triangleq 1 - \frac{\left(\frac{v_{peak1} - v_{peak2}}{2}\right) - v_{min}}{\left(\frac{v_{peak1} - v_{peak2}}{2}\right)} \qquad \text{Equation 2}$$

696   with v_{peak1} and v_{peak2} representing the velocity peak value of the out-centre and centre-in reaching

697   movement of a given segment, respectively, and v_{min} representing the minimum value between these

698   two points. We normalised the obtained difference, ranging from 0 to 1, with 1 indicating a fully

699   coarticulated movement. Given that in this task three pairs of movements were able to be

700   coarticulated, the maximum CI value was 3 in each trial. Note that this measure is MT insensitive since

701   it focuses on the difference in values not on their absolute value. This means slow and fast movements

702   could have a similar CI value.

703   **Error:** We operationalised error as the amount of repetitions necessary to complete a given trial. Trials

704   had to be repeated if participants missed a target or if they exited the start box before the targets

705   appeared or disappeared.

706   **Minimum-jerk model:** A traditional minimum-jerk model for motor control is guided by optimisation

707   theory, where a 'cost' is minimised over the trajectory [12,53]. In the case of the minimum-jerk model,

708   the cost is defined as the squared jerk (3rd derivative of position with respect to time):

709

710

$$jerk \triangleq \frac{1}{2} \int_{t_1}^{t_2} \left(\left|\frac{d^3 x}{dt^3}\right|^2 + \left|\frac{d^3 y}{dt^3}\right|^2 \right) dt \qquad \text{Equation 3}$$

711

712   Here x and y represent the position of the index finger over time (t), while t_1 and t_2 define the start

713   and end of a trial in seconds (t). The Matlab code provided by Todorov and Jordan (1998)[64] was used

714   to compute the minimum jerk trajectory (trajectory that minimised Equation 3), and the

715   accompanying velocity profile, given a set of via points, start/end position and movement time [12]. We

716   then calculated the mean square error (immse function in Matlab) between the predicted and actual

717   velocity profile, which were both normalised and interpolated (N = 500), to estimate the fit on a trial-

718   by-trial basis. Due to the two-dimensional structure of trajectories, we chose the velocity profiles of

719   rather than the trajectories for this comparison.

720   **Spectral Arc Length:** To further assess movement smoothness, we measured spectral arc length.

721   Although we decided to use the traditional jerk metric in our modelling analysis, to allow for

722   comparisons with prior literature, spectral arc length has been shown to be less sensitive to

723   differences in MT and more sensitive to changes in smoothness [55,65]. The spectral arc length is derived

724   from the arc length of the power spectrum of a Fourier transformation of the velocity profile. We used

725   an open-source Matlab toolbox to calculate this value for each trajectory [66].

726   For both spectral arc length and the minimum-jerk model, we only included non-corrected trials. Trials

727   that were classified as corrected included at least one corrective movement to hit a previously missed

728   target. These additional movements added peaks to the velocity profile which complicated model

729   comparison and increased jerkiness disproportionally. Therefore, 1820 trials were excluded for both

730   analyses (8.68% of all trials).

731   **Spatial Reorganisation:** In addition to CI, coarticulation can also be expressed spatially as the radial

732   distance between the peak velocity (v_{peak}) on the sub-movements and the minimum velocity (v_{min})

733   around the via point (Figure 7a). This distance becomes smaller with increased coarticulation [9,10] and

734   reflects the merging of two sub-movements into one (Supplementary Figure 3). To measure these

735   changes in radial distance between peaks and via points, we used a sliding window approach of 10

736   trials at a time. For each target reach (excluding the first and the last) we fitted a confidence ellipse [67]

737   with a 95% confidence criterion around the scatter of the spatial position (*x, y*) of each peak velocity

738   of the included trials (Figure 7a). The confidence ellipses were obtained using principal component

739   analysis to determine the minimum and maximum dispersion of the included data points in the *x-y*

740   plane. To measure the distance between the scatter and its corresponding via point, we determined

741   the ellipse's centroid (point of intersection of ellipse's axes) and calculated the radial distance to the

742   via point. The obtained distance values were normalised and ranged from 0-100%, with 100%

743   representing 0 cm distance between the centroid and the via point. Considering that individual

744   reaching movements display a bell-shaped velocity profile, with the v_{peak} situated approximately in the

745   centre of the movement, radial distance values between 45-55% can be expected if each movement

746   is executed individually (Supplementary Figure 3). To understand the optimal radial distance to the via

747   point, we measured the radial distance for a trajectory which minimised jerk. This suggested that

748   values of 82-85% represent the optimal range for coarticulated movements in this task (Figure 7g).

749   **Variability:** To assess changes in variability we measured the area of the peak velocity ellipses using

750   the same approach as for spatial reorganisation. The area of the ellipse represents the total variance

751   of the included data points in the *x-y* plane and was calculated by multiplying the axes of the ellipse

752   with pi (π) [67,68]. Data was normalised to the baseline for each group.

753   **Statistical Analysis**

754   Wilcoxon tests were used to analyse differences in performance during baseline. Mixed model

755   ANOVAs were used to assess statistical significance of our results in experiment 1. We carried out

756   separate analyses for training with timepoint (early training (first 20 trials), late training (last 20 trials))

757   and group (reward, no reward), and post assessment with condition (post-R, post-NR (all 20 trials in

758   each)) and group (reward, no reward) as factors for both days. We used one-sample Kolmogorov-

759   Smirnov tests to test our data for normality and found that all measures were non-parametric. Median

760   values were therefore used as input in all mixed model ANOVAs. Wilcoxon tests were employed when

761   a significant interaction and/or main effects were reported and corrections for multiple comparisons

762   were performed using Bonferroni correction. Linear partial correlations (fitlm function in Matlab)

763   were used to measure the degree of association between the chosen variables, while accounting for

764   the factor group. Piecewise linear spline functions were fitted through the scatter of spatial distance

765   values and CI levels using least square optimisation by means of shape language modelling (SLM) [69].

766   We used three knots as input for the linear model.

767   A repeated-measure ANOVA was used to test for significance of our results in experiment 2. We

768   compared MT and CI performance separately with timepoint (early training, late training over all 3

769   testing days) as the within factor. Due to our data being non-parametric after using one-sample

770   Kolmogorov-Smirnov tests, we included median values as input for all repeated-measure ANOVAs.

771   Wilcoxon test was used as post-hoc test and multiple comparisons were corrected using Bonferroni

772   corrections.

773   We used mixed model ANOVAs to statistically analyse our results from experiment 3. Separate

774   analyses were carried out for training and post assessment for both MT and CI. Timepoint (early

775   training, late training) and group (no reward, reward without feedback, reward with random feedback,

776   reward with correct feedback) were factors to assess performance in training. Post assessment

777   performance was analysed with a mixed model ANOVA that had condition (Post-R, Post-NR) and group

778   (no reward, reward without feedback, reward with random feedback, reward with correct feedback)

779   as factors. We used median values as input, because one-sample Kolmogorov-Smirnov tests confirmed

780   a lack of normality. Wilcoxon tests were employed when a significant interaction and/or main effects

781   were reported and corrections for multiple comparisons were performed using Bonferroni correction.

782

References

1.   Fowler, C. A. Coarticulation and theories of extrinsic timing. 21.

2.   Willingham, D. B. A Neuropsychological Theory of Motor Skill Learning. 27.

3.   Doeringer, J. A. & Hogan, N. Serial processing in human movement production. *Neural Networks* **11**, 1345–1356 (1998).

4.   Jin, X., Tecuapetla, F. & Costa, R. M. Basal ganglia subcircuits distinctively encode the parsing and concatenation of action sequences. *Nat Neurosci* **17**, 423–430 (2014).

5.   Shah, A., Barto, A. & Fagg, A. A Dual Process Account of Coarticulation in Motor Skill Acquisition. *Journal of Motor Behavior* **45**, 531 (2013).

6.   Hansen, E., Grimme, B., Reimann, H. & Schöner, G. Anticipatory coarticulation in non-speeded arm movements can be motor-equivalent, carry-over coarticulation always is. *Exp Brain Res* **236**, 1293–1307 (2018).

7.   Diedrichsen, J. & Kornysheva, K. Motor skill learning between selection and execution. *Trends Cogn. Sci. (Regul. Ed.)* **19**, 227–233 (2015).

8.   Ramkumar, P. *et al.* Chunking as the result of an efficiency computation trade-off. *Nat Commun* **7**, 12176 (2016).

9.   Sosnik, R., Hauptmann, B., Karni, A. & Flash, T. When practice leads to co-articulation: the evolution of geometrically defined movement primitives. *Exp Brain Res* **156**, 422–438 (2004).

10.   Sosnik, R., Chaim, E. & Flash, T. Stopping is not an option: the evolution of unstoppable motion elements (primitives). *Journal of Neurophysiology* **114**, 846 (2015).

11.   Sosnik, R., Flash, T., Hauptmann, B. & Karni, A. The acquisition and implementation of the smoothness maximization motion strategy is dependent on spatial accuracy demands. *Exp Brain Res* **176**, 311–331 (2007).

12.   Todorov, E. & Jordan, M. I. Smoothness Maximization Along a Predefined Path Accurately Predicts the Speed Profiles of Complex Arm Movements. *Journal of Neurophysiology* **80**, 696–714 (1998).

13.   Kuhnert, B. & Nolan, F. The origin of coarticulation. *FIPKM)* **35**, (1997).

14.   Kent, R. D. & Minifie, F. D. Coarticulation in recent speech production models. *Journal of Phonetics* **5**, 115–133 (1977).

15.   Jerde, T. E., Soechting, J. F. & Flanders, M. Coarticulation in Fluent Fingerspelling. *The Journal of Neuroscience* **23**, 2383 (2003).

16.   Winges, S. A. & Furuya, S. Distinct digit kinematics by professional and amateur pianists. *Neuroscience* **0**, 643 (2015).

17.   Soechting, J. F. & Flanders, M. Organization of sequential typing movements. *Journal of Neurophysiology* (1992) doi:10.1152/jn.1992.67.5.1275.

18.   Soechting, J. F. & Flanders, M. Flexibility and Repeatability of Finger Movements During Typing: Analysis of Multiple Degrees of Freedom. *J Comput Neurosci* **4**, 29–46 (1997).

19.   Rumelhart, D. E. & Norman, D. A. Simulating a Skilled Typist: A Study of Skilled Cognitive-Motor Performance. *Cognitive Science* **6**, 1–36 (1982).

20.   Breteler, M. D. K., Hondzinski, J. M. & Flanders, M. Drawing Sequences of Segments in 3D: Kinetic Influences on Arm Configuration. *Journal of Neurophysiology* (2003) doi:10.1152/jn.01062.2002.
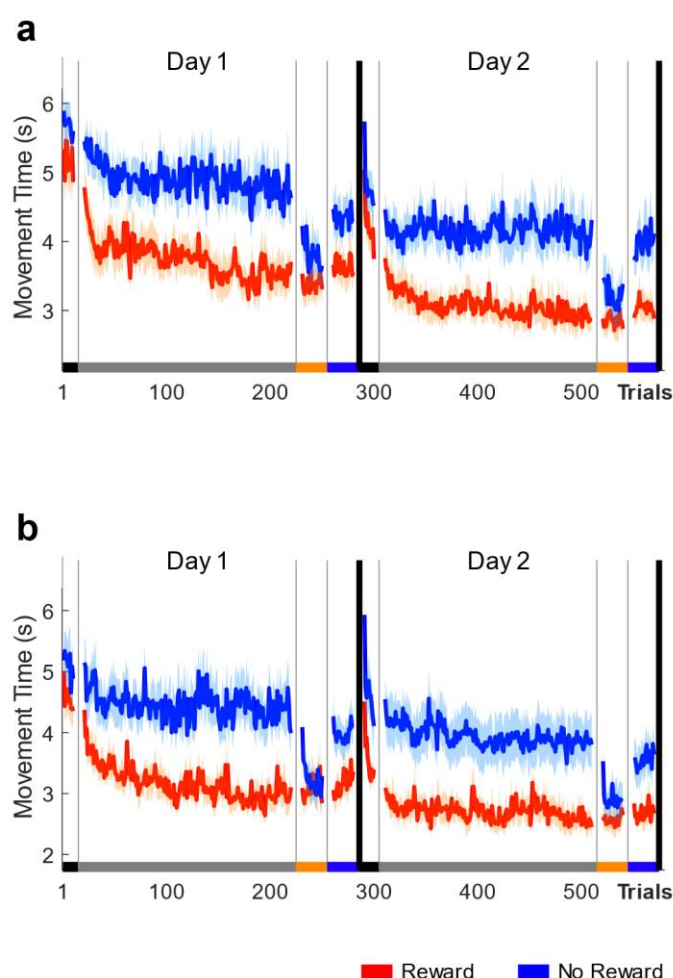
822   21.   Verwey, W. B. & Dronkert, Y. Practicing a Structured Continuous Key-Pressing Task: Motor Chunking
823   or Rhythm Consolidation? *J Mot Behav* **28**, 71–79 (1996).

824   22.   Verwey, W. B., Groen, E. C. & Wright, D. L. The stuff that motor chunks are made of: Spatial instead of
825   motor representations? *Exp Brain Res* **234**, 353–366 (2016).

826   23.   Wymbs, N. F., Bassett, D. S., Mucha, P. J., Porter, M. A. & Grafton, S. T. Differential recruitment of the
827   sensorimotor putamen and frontoparietal cortex during motor chunking in humans. *Neuron* **74**, 936 (2012).

828   24.   Thompson, J. J., McColeman, C. M., Blair, M. R. & Henrey, A. J. Classic motor chunking theory fails to
829   account for behavioural diversity and speed in a complex naturalistic task. *PLOS ONE* **14**, e0218251 (2019).

830   25.   Friedman, J. & Korman, M. Observation of an expert model induces a skilled movement coordination
831   pattern in a single session of intermittent practice. *Sci Rep* **9**, 1–15 (2019).

832   26.   Boyd, L. A. *et al.* Motor sequence chunking is impaired by basal ganglia stroke. *Neurobiology of
833   learning and memory* **92**, 35–44 (2009).

834   27.   Huntley, J., Bor, D., Hampshire, A., Owen, A. & Howard, R. Working memory task performance and
835   chunking in early Alzheimer's disease. *Br J Psychiatry* **198**, 398–403 (2011).

836   28.   Ariani, G. & Diedrichsen, J. Sequence learning is driven by improvements in motor planning. *J.
837   Neurophysiol.* **121**, 2088–2100 (2019).

838   29.   Verwey, W. B. Evidence for a multistage model of practice in a sequential movement task. *J EXP
839   PSYCHOL HUMAN* **25**, 1693–1708 (1999).

840   30.   Barnhoorn, J. S., Van Asseldonk, E. H. F. & Verwey, W. B. Differences in chunking behavior between
841   young and older adults diminish with extended practice. *Psychological Research* **83**, 275–285 (2019).

842   31.   Sosnik, R., Flash, T., Sterkin, A., Hauptmann, B. & Karni, A. The activity in the contralateral primary
843   motor cortex, dorsal premotor and supplementary motor area is modulated by performance gains. *Frontiers in
844   Human Neuroscience* **8**, (2014).

845   32.   Bidet-Ildei, C., Pollak, P., Kandel, S., Fraix, V. & Orliaguet, J.-P. Handwriting in patients with Parkinson
846   disease: effect of L-dopa and stimulation of the sub-thalamic nucleus on motor anticipation. *Hum Mov Sci* **30**,
847   783–791 (2011).

848   33.   Mazzoni, P., Hristova, A. & Krakauer, J. W. Why Don't We Move Faster? Parkinson's Disease,
849   Movement Vigor, and Implicit Motivation. *J. Neurosci.* **27**, 7105–7116 (2007).

850   34.   Rohrer, B. *et al.* Movement Smoothness Changes during Stroke Recovery. *The Journal of Neuroscience*
851   **22**, 8297 (2002).

852   35.   Rohrer, B. *et al.* Submovements Grow Larger, Fewer, and More Blended during Stroke Recovery.
853   *Motor Control* **8**, 472–483 (2004).

854   36.   Dipietro, L., Krebs, H. I., Fasoli, S. E., Volpe, B. T. & Hogan, N. Submovement changes characterize
855   generalization of motor recovery after stroke. *Cortex* **45**, 318–324 (2009).

856   37.   Gulde, P., Hughes, C. M. L. & Hermsdörfer, J. Effects of Stroke on Ipsilesional End-Effector Kinematics
857   in a Multi-Step Activity of Daily Living. *Front. Hum. Neurosci.* **11**, (2017).

858   38.   Dayan, P. & Balleine, B. W. Reward, motivation, and reinforcement learning. *Neuron* **36**, 285–298
859   (2002).

860   39.   Sutton, R. S. & Barto, A. G. Reinforcement Learning: An Introduction. 352.

861   40.   Manohar, S. G. *et al.* Reward Pays the Cost of Noise Reduction in Motor and Cognitive Control.
862   *Current Biology* **25**, 1707 (2015).

863   41.   Codol, O., Holland, P. J., Manohar, S. G. & Galea, J. M. Reward-based improvements in motor control
864   are driven by multiple error-reducing mechanisms. *J. Neurosci.* (2020) doi:10.1523/JNEUROSCI.2646-19.2020.

865   42.   Summerside, E. M., Shadmehr, R. & Ahmed, A. A. Control of Movement: Vigor of reaching
866   movements: reward discounts the cost of effort. *Journal of Neurophysiology* **119**, 2347 (2018).

867   43.   Codol, O., Galea, J. M., Jalali, R. & Holland, P. J. Reward-driven enhancements in motor control are
868   robust to TMS manipulation. *Exp Brain Res* 1–13 (2020) doi:10.1007/s00221-020-05802-1.

869   44.   Kawagoe, R., Takikawa, Y. & Hikosaka, O. Reward-Predicting Activity of Dopamine and Caudate
870   Neurons—A Possible Mechanism of Motivational Control of Saccadic Eye Movement. *Journal of*
871   *Neurophysiology* **91**, 1013–1024 (2004).

872   45.   Pekny, S. E., Izawa, J. & Shadmehr, R. Reward-Dependent Modulation of Movement Variability. *The*
873   *Journal of Neuroscience* **35**, 4015 (2015).

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

876   47.   Carroll, T. J., McNamee, D., Ingram, J. N. & Wolpert, D. M. Rapid Visuomotor Responses Reflect Value-
877   Based Decisions. *The Journal of Neuroscience* **39**, 3906 (2019).

878   48.   Goodman, R. N. *et al.* Increased reward in ankle robotics training enhances motor control and cortical
879   efficiency in stroke. *J Rehabil Res Dev* **51**, 213–227 (2014).

880   49.   Quattrocchi, G., Greenwood, R., Rothwell, J. C., Galea, J. M. & Bestmann, S. Reward and punishment
881   enhance motor adaptation in stroke. *J. Neurol. Neurosurg. Psychiatry* **88**, 730–736 (2017).

882   50.   Abe, M. *et al.* Reward improves long-term retention of a motor memory through induction of offline
883   memory gains. *Curr. Biol.* **21**, 557–562 (2011).

884   51.   Wächter, T., Lungu, O. V., Liu, T., Willingham, D. T. & Ashe, J. Differential effect of reward and
885   punishment on procedural learning. *J. Neurosci.* **29**, 436–443 (2009).

886   52.   Steel, A., Silson, E. H., Stagg, C. J. & Baker, C. I. The impact of reward and punishment on skill learning
887   depends on task demands. *Sci Rep* **6**, (2016).

888   53.   Flash, T. & Hogan, N. The coordination of arm movements: an experimentally confirmed
889   mathematical model. *J. Neurosci.* **5**, 1688–1703 (1985).

890   54.   Huang, H. J., Kram, R. & Ahmed, A. A. Reduction of Metabolic Cost during Motor Learning of Arm
891   Reaching Dynamics. *Journal of Neuroscience* **32**, 2182–2190 (2012).

892   55.   Balasubramanian, S., Melendez-Calderon, A., Roby-Brami, A. & Burdet, E. On the analysis of
893   movement smoothness. *Journal of NeuroEngineering and Rehabilitation* **12**, 112 (2015).

894   56.   Gribble, P. L., Mullin, L. I., Cothros, N. & Mattar, A. Role of cocontraction in arm movement accuracy.
895   *J. Neurophysiol.* **89**, 2396–2405 (2003).

896   57.   Dhawale, A. K., Smith, M. A. & Ölveczky, B. P. The Role of Variability in Motor Learning. *Annu Rev*
897   *Neurosci* **40**, 479–498 (2017).

898   58.   Sternad, D. It's not (only) the mean that matters: variability, noise and exploration in skill learning.
899   *Current Opinion in Behavioral Sciences* **20**, 183–195 (2018).

900   59.   Shmuelof, L., Krakauer, J. W. & Mazzoni, P. How is a motor skill learned? Change and invariance at the
901   levels of task success and trajectory control. *J Neurophysiol* **108**, 578–594 (2012).

902   60.   Schmidt, R. A. & Lee, T. D. *Motor control and learning: A behavioral emphasis, 5th ed*. ix, 581 (Human
903   Kinetics, 2011).

904   61.   Kornysheva, K. *et al.* Neural Competitive Queuing of Ordinal Structure Underlies Skilled Sequential
905   Action. *Neuron* **101**, 1166-1180.e3 (2019).

906   62.   Santos, F. J., Oliveira, R. F., Jin, X. & Costa, R. M. Corticostriatal dynamics encode the refinement of
907   specific behavioral variability during skill learning. *Elife* **4**, e09423 (2015).

908   63.   Collins, A. L. *et al.* Dynamic mesolimbic dopamine signaling during action sequence learning and
909   expectation violation. *Sci Rep* **6**, 20231 (2016).

910   64.   https://www.mathworks.com/matlabcentral/mlc-
911   downloads/downloads/submissions/58403/versions/7/previews/min_jerk_todorov.m/index.html.

912   65.   Gulde, P. & Hermsdörfer, J. Smoothness Metrics in Complex Movement Tasks. *Front. Neurol.* **9**,
913   (2018).

914   66.   siva82kb/smoothness. *GitHub* https://github.com/siva82kb/smoothness.

915   67.   Sadnicka, A. *et al.* High motor variability in DYT1 dystonia is associated with impaired visuomotor
916   adaptation. *Scientific Reports* **8**, 1–11 (2018).

917   68.   Revol, P. *et al.* Pointing errors in immediate and delayed conditions in unilateral optic ataxia. *Spat Vis*
918   **16**, 347–364 (2003).

919   69.   D'Errico, J. SLM - Shape Language Modeling.
920   https://de.mathworks.com/matlabcentral/fileexchange/24443-slm-shape-language-modeling (2020).

921

922

923

924

925

926

927

928

929

930

931

932

933

934

935

936 **Supplement**

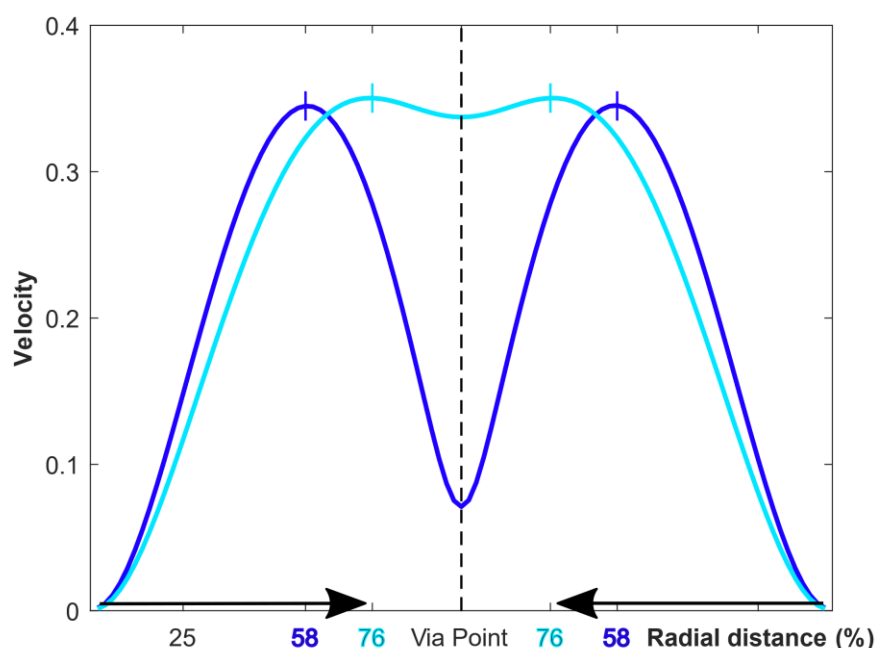937 **Supplementary Figure 1**

938

Reward ▮ No Reward ▮

939 **Supplementary Figure 1 | Counterbalancing did not affect MT results.** Data was split depending on the
940 counterbalancing. Participants in **a)** completed post-R prior to post-NR, whereas participants encountered the
941 reversed order during their post assessment (**b**).

942

943

944

945

946

947

948

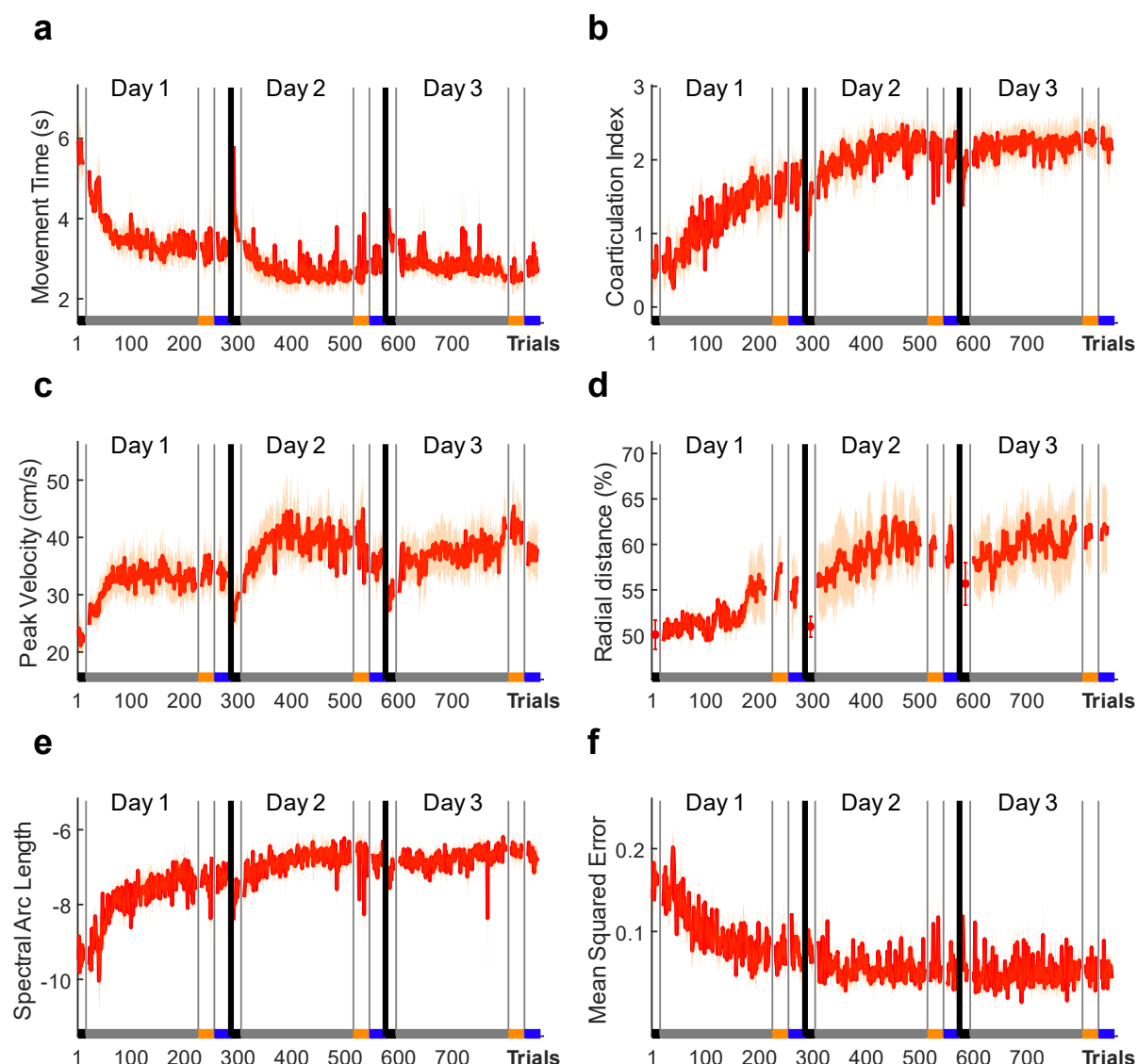949   **Supplementary Figure 2**

950
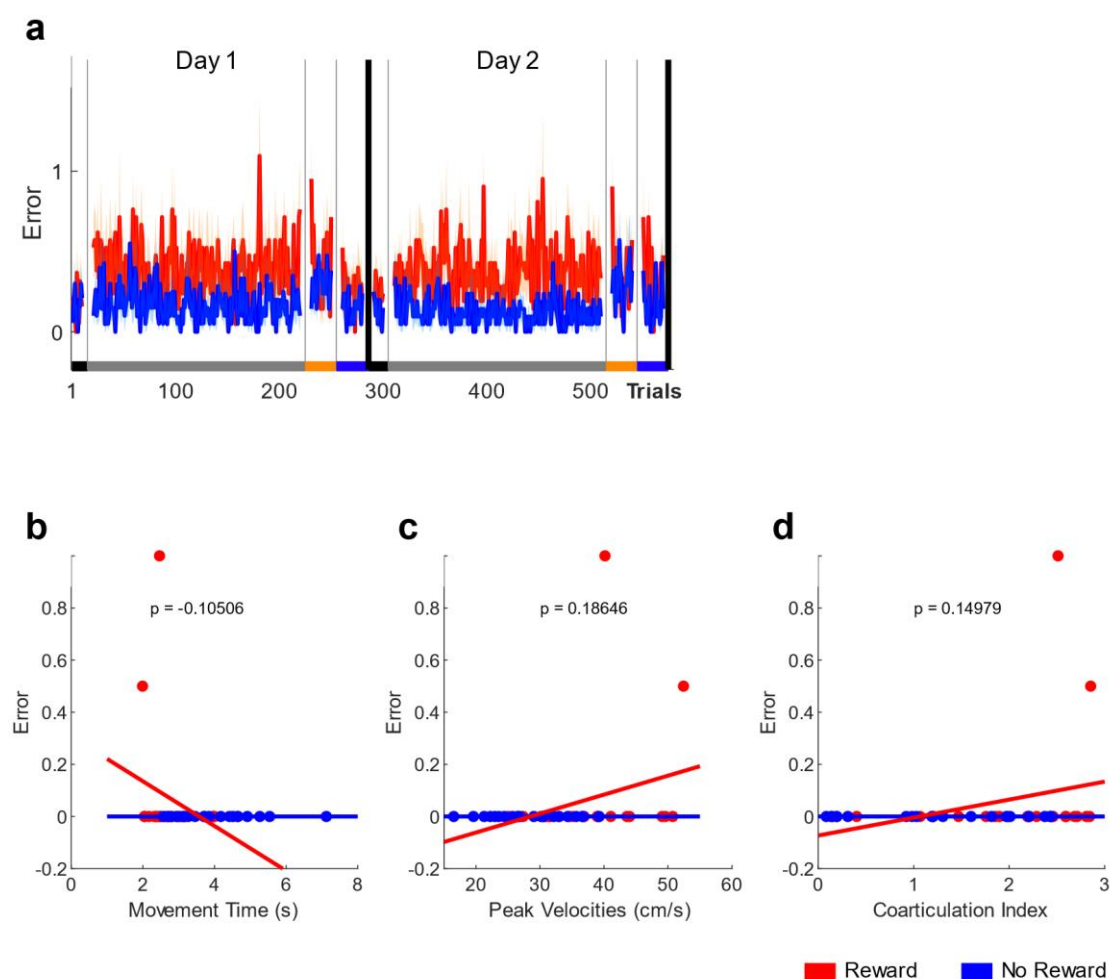
951   **Supplementary Figure 2 | Number of errors did not affect performance during later stages of training. a)**
952   Trial-by-trial data showing the number of errors across participants for both groups). **b-d)** Scatterplots
953   comparing median error data during early training with median **(b)** MT, **(c)** peak velocity and **(d)** CI data during
954   late training.

955

956

957

958

959

960

961

962
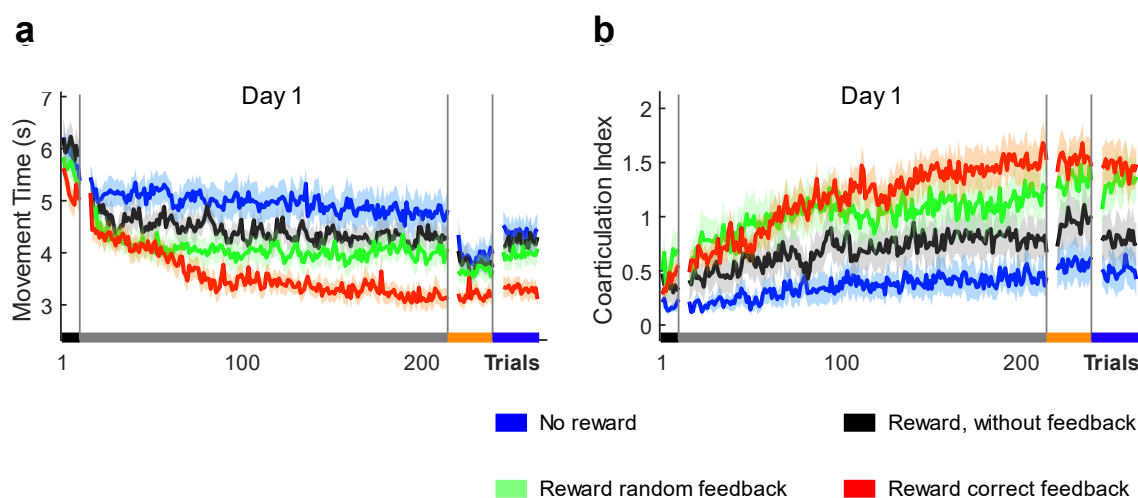
963 **Supplementary Figure 3**

964

965

966 **Supplementary Figure 3 | Illustration of radial distance (%) values depending on different levels of**
967 **coarticulation.** Considering two target reaches (same length) the spatial position (x,y) changes depending on the
968 level of coarticulation thereby exhibiting less radial distance to the via point. We used the predictions of the
969 minimum jerk model used in this paper to plot the optimal solutions for the here included sequential target
970 reaches.

971

972

973

974

975

976

977

978

979

980

981

36

Supplementary Figure 4

a

b

c

d

e

f

Supplementary Figure 4 |Performance gains are stable even over a prolonged washout period. Trial-by-trial data for **a)** MT, **b)** CI levels **c)** peak velocity **d)** radial distance, **e)** spectral arc length, **f)** mean squared error CI for the second experiment which was scheduled on three consecutive days to assess behavioural change over a prolonged washout period. Participants received reward during training on the first two days, however no reward was available on the third day.

Supplementary Figure 5

Supplementary Figure 5 | Reward based on correct performance feedback is most efficient at driving behavioural change. Trial-by-trial data for **a)** MT and **b)** CI levels for the third experiment which included 4 groups that differed with regards to reward availability and feedback type during training to investigate what drove behavioural changes observed in experiment 1.