A critical re-evaluation of fMRI signatures of motor sequence learning

Eva Berlot[1], Nicola J. Popp[1], & Jörn Diedrichsen[1-3]

[1] The Brain and Mind Institute, University of Western Ontario, Canada
[2] Department of Statistical and Actuarial Sciences, University of Western Ontario, Canada
[3] Department of Computer Science, University of Western Ontario, Canada

Correspondence

Jörn Diedrichsen, The Brain and Mind Institute, University of Western Ontario, London, Canada. jdiedric@uwo.ca

Abstract

29

30
31 Despite numerous studies, there is little agreement about what brain changes accompany motor

32 sequence learning, partly because of a general publication bias that favors novel results. We

33 therefore decided to systematically reinvestigate proposed functional magnetic resonance

34 imaging correlates of motor learning in a preregistered longitudinal study with four scanning

35 sessions over 5 weeks of training. Activation decreased more for trained than untrained

36 sequences in premotor and parietal areas, without any evidence of learning-related activation

37 increases. Premotor and parietal regions also exhibited changes in the fine-grained, sequence-

38 specific activation patterns early in learning, which stabilized later. No changes were observed in

39 the primary motor cortex (M1). Overall, our study provides evidence that human motor sequence

40 learning occurs outside of M1. Furthermore, it shows that we cannot expect to find activity

41 increases as an indicator for learning, making subtle changes in activity patterns across weeks

42 the most promising fMRI correlate of training-induced plasticity.

Introduction

43

44

45   Humans have the remarkable ability to learn complex sequences of movements. While

46   behavioural improvements in sequence learning tasks are easily observable, the underlying

47   neural processes remain elusive. Understanding the neural underpinnings of motor sequence

48   learning could provide clues about more general mechanisms of plasticity in the brain. This

49   motivation has led numerous functional magnetic resonance imaging (fMRI) studies to investigate

50   the brain changes related to motor sequence learning. However, there is little agreement about

51   how and where in the brain learning-related changes are observable. Previous studies include

52   reports of signal increases across various brain regions (Floyer-Lea & Matthews, 2005; Grafton,

53   Hazeltine, & Ivry, 1995; Hazeltine, Grafton, & Ivry, 1997; Karni et al., 1995; Lehéricy et al., 2005;

54   Penhune & Doyon, 2002), as well as signal decreases (Jenkins, Brooks, Nixon, Frackowiak, &

55   Passingham, 1994; Peters, Lee, Hedrick, Neil, & Komiyama, 2017; Toni, Krams, Turner, &

56   Passingham, 1998; Ungerleider, Doyon, & Karni, 2002; Wiestler & Diedrichsen, 2013), nonlinear

57   changes in activation (Ma et al., 2010; Xiong et al., 2009), spatial shifts in activity (Lehéricy et al.,

58   2006; Steele & Penhune, 2010), changes in multivariate patterns (Wiestler & Diedrichsen, 2013;

59   Wymbs & Grafton, 2015), and changes in inter-regional functional connectivity (Bassett, Yang,

60   Wymbs, & Grafton, 2015; Bassett et al., 2010; Doyon et al., 2002; Mattar et al., 2016). Additionally,

61   some experiments have matched the speed of performance (Karni et al., 1995; Penhune &

62   Doyon, 2002; Steele & Penhune, 2010; Lehéricy et al., 2005; Seidler et al., 2002, 2005), while

63   others have not (Bassett et al., 2015; Lutz, Koeneke, Wüstenberg, & Jäncke, 2004; Wiestler &

64   Diedrichsen, 2013; Wymbs & Grafton, 2015). Given that fMRI analysis has many degrees of

65   freedom, these inconsistencies may not be too surprising. However, the implicit pressure in the

66   publication system to report findings may also have contributed to a lack of coherency. To address

67   this issue, we designed a comprehensive longitudinal study of motor sequence learning that

68   allowed us to systematically reinvestigate previous findings. In order to increase transparency,

69   we pre-registered the design, as well as all tested hypotheses on the Open Science Framework

70   (Berlot, Popp, & Diedrichsen, 2017; https://osf.io/etnqc), and make the full dataset available to

71   the research community.

72   The main aim of our study was to systematically evaluate different ideas of how learning-

73   related changes are reflected in the fMRI signal. In the context of motor sequence learning, the

74   most commonly examined brain region is the primary motor cortex (M1). Previous reports of

75   increased M1 activation after long-term learning have been interpreted as additional recruitment

76   of neuronal resources for trained behavior, taken to suggest the skill is represented in M1 (Floyer-

3

77 Lea & Matthews, 2005; Karni et al., 1995, 1998; Lehéricy et al., 2005; Penhune & Doyon, 2002;

78 for a review see Dayan & Cohen, 2011; **Fig. 1a**). Since then, several pieces of evidence have

79 suggested that sequence-specific memory may not reside in M1 (Beukema, Diedrichsen, &

80 Verstynen, 2019; Wiestler & Diedrichsen, 2013; Yokoi & Diedrichsen, 2019). However, some of

81 these reports studied skill acquisition over a course of a few days, while human skill typically

82 evolves over weeks (and months) of practice. Therefore, including several weeks of practice,

83 might be more suitable to test whether, and at what time point, M1 develops skill-specific

84 representations.

85 Outside of M1, learning-related activation changes have been reported in premotor and

86 parietal areas (Grafton, Hazeltine, & Ivry, 2002; Hardwick, Rottschy, Miall, & Eickhoff, 2013;

87 Honda et al., 1998; Penhune & Doyon, 2002; Tamás Kincses et al., 2008; Vahdat et al., 2015),

88 with activation increases commonly interpreted as increased involvement of these areas in the

89 skilled behavior. Yet, recent studies have mostly found that, as the motor skill develops, activation

90 in these areas predominantly decreases (Penhune & Steele, 2012; Wiestler & Diedrichsen, 2013;

91 Wu et al., 2004). Such reductions are harder to interpret as they could reflect a reduced areal

92 involvement in skilled performance or, alternatively, more energy efficient implementation of the

93 same function **(Fig. 1b)** (Picard, Matsuzaka, & Strick, 2013; Poldrack et al., 2005). To complicate

94 things further, regional activity increases and decreases could occur simultaneously in the same

95 area (**Fig. 1c**; Steele & Penhune, 2010). In such a scenario, the net activation in the region would

96 not change, yet, the trained sequences would engage slightly different subpopulations of the

97 region than untrained sequences.

98 A variant of this idea is that each specific sequence becomes associated with dedicated

99 neuronal subpopulation (and hence fMRI activity pattern). Such a representation would form the

100 neural correlate of sequence-specific learning – the part of the skill that does not generalize to

101 novel, untrained motor sequences (Karni et al., 1995). Sequence-specific activation patterns

102 should change early in learning **(Fig. 1d)**, when behavior improves most rapidly, and stabilize

103 later, once the skill has consolidated and an optimal pattern is established (Peters et al., 2017).

104 One possible way in which sequence-specific patterns could reorganize is by becoming more

105 distinct from one another **(Fig. 1e**; Wiestler & Diedrichsen, 2013). Having a distinctive code for

106 each sequence might be of particular importance to the system in a trained state, allowing it to

107 produce different dynamical sequences, while avoiding confusion or "tangling" of the different

108 neural trajectories (Russo et al., 2018).

109 To systematically examine the cortical changes associated with motor sequence learning,

110 we carried out a longitudinal study over 5 weeks of training with 4 sessions of high-field (7 Tesla)

111   fMRI scans. Behavioural performance in the first three scanning sessions was imposed to the

112   same speed of performance. This allowed us to inspect whether examined fMRI metrics reflect

113   brain reorganization, independent of behavioral change. However, controlling for speed incurs

114   the danger of not tapping into neural resources that are necessary for skilled performance (Orban

115   et al., 2010; Poldrack, 2000). We therefore compared the fMRI session with paced performance

116   at the end of behavioural training with one acquired with full speed performance **(Fig. 2)**. This

117   manipulation allowed us to systematically assess the role of speed on the fMRI metrics of learning,

118   thereby addressing an important methodological problem faced by virtually every study on motor

119   learning.

120
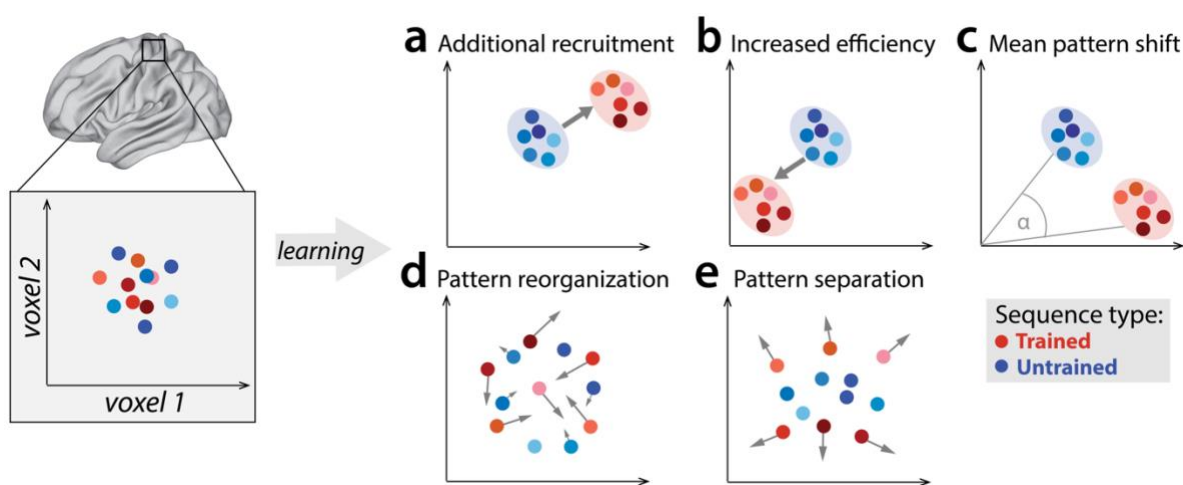
121
122   **Figure 1. Potential fMRI signatures of learning in a specific brain area.** Each panel shows hypothetical
123   activation for the six trained sequences (red) and the six untrained sequences (blue) in the space of two
124   hypothetical voxels. **a)** Activation could increase during learning across voxels, indicating additional
125   recruitment of resources involved in skilled behavior. **b)** Activation could decrease across voxels, implying
126   that the region performs its function more efficiently. **c)** Some voxels (x-axis) could increase activation with
127   training, while others (y-axis) could decrease. This would lead to a shift of the overall activity pattern in the
128   region without an overall net change in activation. **d)** Activation patterns specific to each trained sequence
129   could undergo more change than untrained sequences, reflective of plastic reorganization of the sequence
130   representation. Arrow length in the figure indicates the amount of reorganization. **e)** One specific form of
131   such reorganization would be increasing dissimilarities (pattern separation) between activity patterns for
132   individual trained sequences.

Results

Speed of sequence execution increases with learning

We trained 26 participants to perform six 9-digit sequences with their right hand on a keyboard device **(Fig. 2a)**. During training, they received visual feedback (green for correct and red for incorrect presses) and were rewarded for both accuracy and speed (see Materials and Methods). Over the course of 5 weeks, participants practiced ~4000 trials **(Fig. 2b)**. This led to substantial performance improvement, with the average movement time (MT) to complete a sequence decreasing from an initial 3.2 seconds to 1.2 seconds at the end of the training **(Fig. 2c)**. The training regime was complemented with behavioral assessments on four occasions designed to specifically assess participants' performance on trained sequences relative to untrained sequences (**Fig. 2d**, yellow underlay). Prior to training (test day 1), the speed of sequence execution did not differ between trained and untrained sequences. For all subsequent sessions, MTs were significantly faster for trained than untrained sequences (p<.001), implying sequence-specific learning. Additionally, performance of trained sequences improved between all subsequent sessions, even after week 3 (week 3-5: $t_{(25)}$=5.49, p=1.1e-5). Thus, participants' performance of trained sequences improved across the five weeks.

To assess fMRI changes with learning, participants underwent four fMRI scans (1st scan: before the main training; 2nd scan: week 2; 3rd & 4th scan: week 5), performing both trained and untrained sequences (**Fig. 2d** – grey underlay). During the first three sessions, participants were paced with a metronome so that all sequences, trained and untrained, were performed at the same speed as in the first scan. Performance in the fourth session was at maximum speed, resulting in significantly lower MTs for trained compared to untrained sequences **(Fig. 2d)**. To assess different neural signatures of observed behavioral learning, we first examined how the overall evoked activation changed over weeks of training for the same speed of movement.
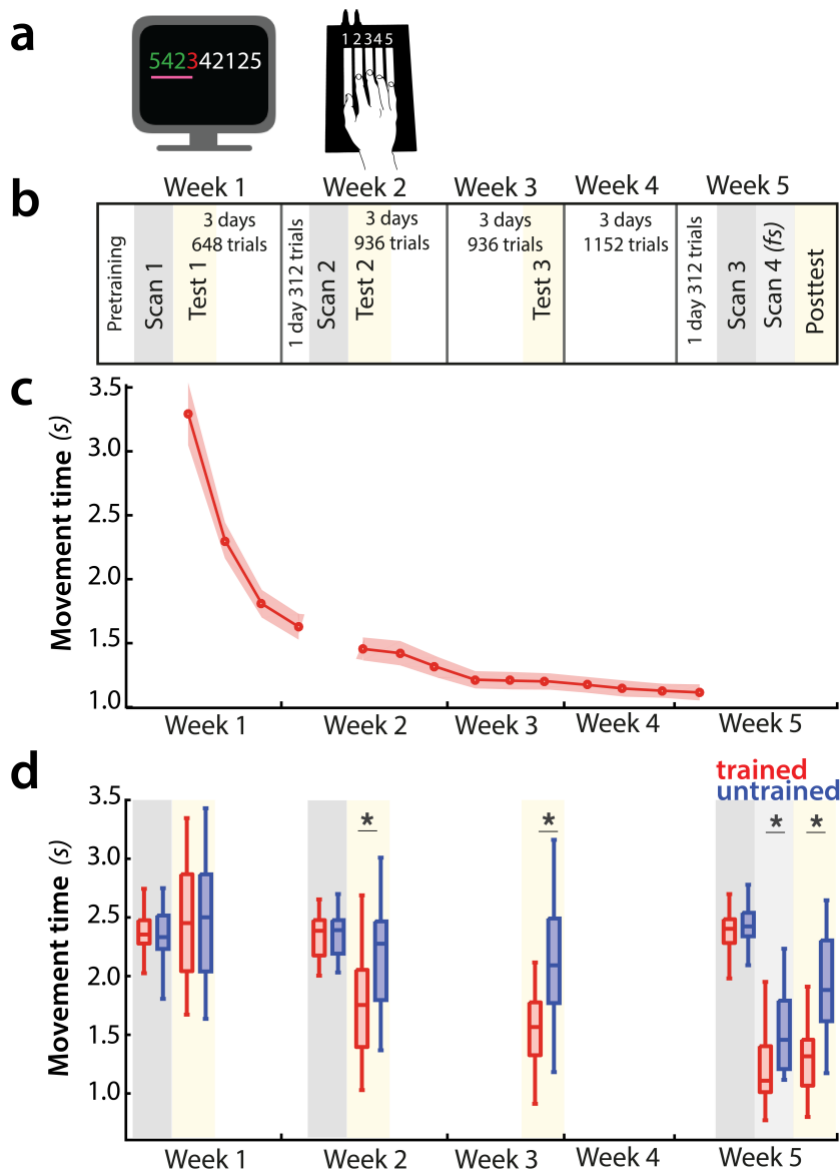
Figure 2. Experimental design and paradigm. a) Apparatus and task. Participants were trained to perform six 9-item sequences on a keyboard device. For each finger press, the corresponding digit on the screen turned green (correct) or red (incorrect). During fMRI scans 1-3, an expanding pink line under the numbers indicated the pace at which participants had to press the keys. See supplementary figure **S2** for trial structure during scanning sessions. **b)** Training protocol lasted for 5 weeks, and included four behavioral test sessions (yellow underlay) and four scans (grey underlay). Scans 1-3 were performed at a paced speed, while scan 4 performance was full speed *(fs)*. **c)** Average group performance executing trained sequences across the training sessions, measured in seconds. The average movement time (MT) decreased with learning. Shaded area denotes between-subject standard error. **d)** Performance during scanning sessions and behavioral tests, measured in seconds. Performance of trained sequences improved across all subsequent behavioral test sessions. Performance improved also for untrained sequences from week 2 onwards, suggesting some transfer in learning, but performance was still faster for trained sequences, indicating sequence-specific learning. Error bars indicate between-subject standard error. Stars denote significance levels lower than *p*<.001.

7

173 ***Overall activation does not change in M1***

174 First, we re-investigated the classical finding that activity, measured as the percent BOLD signal

175 change relative to rest, increased in M1 for matched performance after long-term training (Karni

176 et al., 1995; **Fig. 1a**). Our task elicited activation in a range of cortical areas (**Fig. 3a** for session

177 1 – i.e., prior to learning). A region of interest (ROI) analysis of the hand area of M1, contralateral

178 to the performing hand, however, showed no significant change across weeks (**Fig. 3b**,

179 $F_{(2,50)}$=1.82, p=.17). Neither did we find any difference between trained and untrained sequences

180 ($F_{(1,25)}$=0.19, p=.66), or a significant interaction between the two ($F_{(2,50)}$=2.01, p=0.14).

181 The absence of overall activity changes, however, should not be taken as evidence for an

182 absence of plasticity in the region. It is possible that some subregions of M1 increased in activation

183 for learned sequences, while other decreased, as suggested by Steele and Penhune (2010). Such

184 mixed changes would result in a shift of the overall pattern, which would lead to an increase in

185 the angle between the mean activity pattern for trained and untrained sequences **(Fig. 1c)**.

186 Because we calculated the angle between activity patterns for each participant separately, this

187 criterion does not assume that the observed shift is spatially consistent across individuals – any

188 idiosyncratic shift could be detected. Therefore it serves as a sensitive statistical criterion to detect

189 shifts in spatial location of activation, which were previously reported only descriptively (Steele &

190 Penhune, 2010).

191 However, in M1, the averaged cosine angle **(Fig. 3c)** remained unchanged across the

192 weeks ($F_{(2,50)}$=1.71, p=.19), indicating that the average activity pattern remained comparable

193 across trained and untrained sequences. In sum, we found no evidence for activation increases

194 (Karni et al., 1995), decreases, or relative shifts in activation patterns (Steele & Penhune, 2010)

195 in M1.

8

196
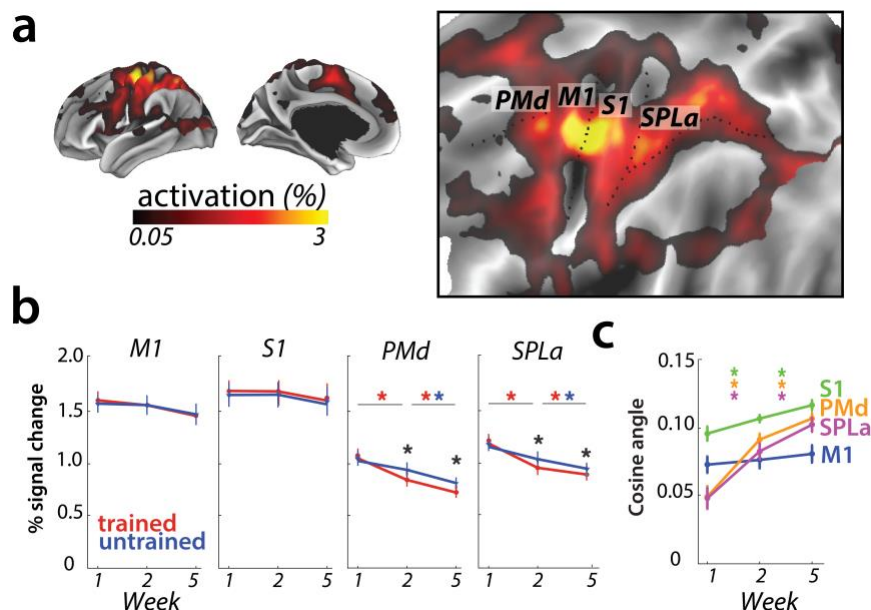
197 **Figure 3. Overall activation and changes with learning in defined regions of interest. a)** Average
198 activation during production of any sequence in scanning session 1 (prior to learning) in the hemisphere
199 contralateral to the performing hand. Activation was contrasted against resting baseline. On the right,
200 activation map is presented on a flattened surface, corresponding to surface maps in other figures. **b)**
201 Changes in activation across predefined areas – primary motor cortex (M1), primary somatosensory cortex
202 (S1), premotor dorsal area (PMd) and superior parietal lobule – anterior (SPLa). No significant changes in
203 activation were observed in M1 or S1 across weeks or between trained and untrained sequences (*
204 indicates $p<.01$). Error bars indicate between-subject standard error. **c)** The cosine angle dissimilarity
205 between average trained and untrained sequence across scanning weeks. The cosine angle increased
206 significantly across weeks in PMd, SPLa and S1, but not M1 (* indicates $p<.05$). Error bars indicate
207 between-subject standard error.

208

209 ***Learning-related activation changes in premotor and parietal areas***

210 To investigate activation changes in areas outside of M1, we calculated changes in activity

211 between the weeks in a map-wise approach (**Fig. 4a**). Over the three measurement time points,

212 we found no reliable activation increases in any cortical area that was activated by the task in

213 week 1. Instead, we observed widespread learning-related reductions in activity in premotor and

214 parietal areas **(Fig. 4a)**, in line with our pre-registered prediction. These activation reductions

215 were observed across both subsequent sessions (i.e. weeks 1-2, weeks 2-5) for trained and

216 untrained sequences, with bigger reductions for trained sequences. In weeks 2 and 5, trained

217 sequences elicited overall lower activity than untrained sequences (**Fig. 4b**; see supplementary

218 figure **S4** for statistical maps). These learning-related reductions in activity were also statistically

219 significant in our predefined ROIs in premotor (dorsal premotor cortex – PMd) and parietal cortices

220 (anterior superior parietal lobule – SPLa) **(Fig. 3b)**: In a 3 (week) x 2 (sequence type) ANOVA on

221 observed activation both main effects and interaction were highly significant in PMd (week:

9

222      $F_{(2,50)}=17.47$, $p=1.77e\text{-}6$; sequence type: $F_{(1,25)}=11.86$, $p=2.03e\text{-}3$; interaction: $F_{(2,50)}=13.22$,

223      $p=2.46e\text{-}5$) as well as in SPLa (week: $F_{(2,50)}=19.14$, $p=6.73e\text{-}7$; sequence type: $F_{(1,25)}=19.36$,

224      $p=1.77e\text{-}4$; interaction: $F_{(2,50)}=21.59$, $p=1.74e\text{-}7$). In contrast, no main effect of week was observed

225   in S1 ($F_{(2,50)}=0.44$, $p=.85$). There was a significant main effect of sequence type ($F_{(1,25)}=6.32$,

226      $p=.019$), but none of the post-hoc t-tests revealed a significant difference. The week x sequence

227   type interaction was not significant in S1 ($F_{(2,50)}=0.17$, $p=.84$). Thus, we observed widespread

228   activation decreases with learning across secondary and association cortical areas.

229   In a few smaller areas, activation increased with learning (red patches in **Fig. 4a-b**). This

230   was observed uniformly in areas with activity at or below baseline – thus these changes reflect

231   decreased suppression of activity rather than increases. It is likely that these activity increases

232   are not task relevant, but instead reflect the increasing automaticity and lower need for central

233   attentional resources with learning (see Discussion).
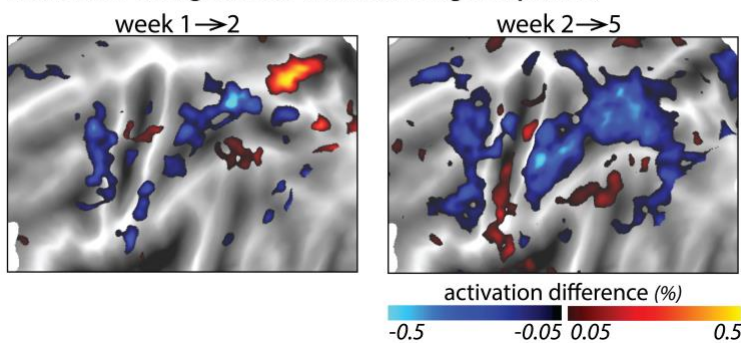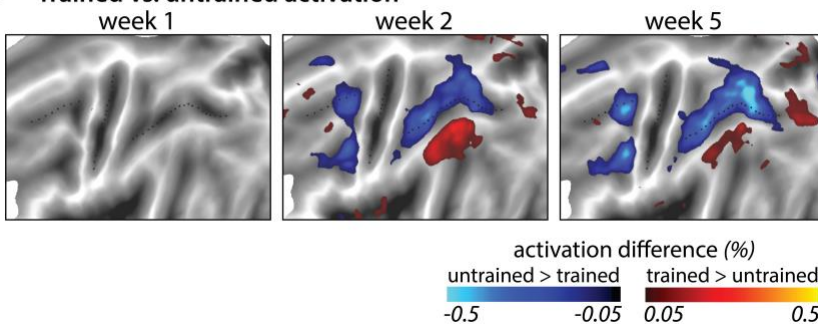
234
235 **Figure 4. Changes in average activation across the cortical surface. a)** Average change in activation
236 across subsequent sessions. Activation was measured as difference in percent signal change relative to
237 the resting baseline. Activation decreased (blue shades) in motor-related regions across sessions during
238 sequence execution. **b)** Contrast of activation for trained vs. untrained sequences per scanning session. In
239 weeks 2 and 5, trained sequences elicited lower activation in motor-related regions than untrained
240 sequences (blue shades; see supplementary figure **S4** for *t*-maps and statistical quantification of activation
241 clusters). Areas with observed increases in activation for trained sequences (red shades) lie in the default
242 mode network that showed on average lower activity during task than rest.
243

244   We also examined whether there were, in addition to the overall activity decreases, shifts

245   in the average activity patterns in the predefined regions of interest **(Fig. 1c)**. As for M1, we

246   calculated the cosine angle dissimilarity (see Materials and Methods) between the average

247   activity patterns for trained and untrained sequences, separately for each scanning session.

248   Figure **5a** shows cosine angle dissimilarities between trained and untrained sequences in PMd,

249   displayed using multidimensional scaling (MDS). Patterns for trained sequences moved away

250   from the starting point over weeks, and became more different from untrained patterns. Both in

251   parietal and premotor areas there was clear evidence for a shift – cosine angular dissimilarity

252   between the average trained and untrained sequence activation increased significantly across

253   weeks (PMd: $F_{(2,50)}=23.63$, $p=5.98e_{-8}$; SPLa: $F_{(2,50)}=23.19$, $p=7.49e_{-8}$) **(Fig. 3c)**. S1 also showed a

254   significant increase in cosine dissimilarity between trained and untrained patterns with learning

255   ($F_{(2,50)}=8.68$, $p=5.79e_{-4}$). These changes, however, were much less pronounced than those

256   observed in premotor and parietal areas.

257   To investigate whether the observed changes in the overall activity patterns in premotor

258   and parietal areas were spatially consistent across individuals, we normalized (z-scored)

259   activation maps in each region and assessed the relative contribution of subregions to overall

260   activation in weeks 1 and 5 **(Fig. 5b)**. Comparing the pattern of activation revealed that before

261   training (week 1, blue) sequences elicit relatively more activation in rostral parts of the premotor

262   and supplementary motor areas, and that activity was more caudal after training (week 5, red;

263   **Fig. 5c** displays the cross-section of relative activation changes). Some differences were also

264   observed in the posterior parietal cortex, with activation shifting from more posterior to anterior

265   subregions after learning **(Fig. 5c)**. Altogether, these results show that with learning, the execution

266   of sequences relies on slightly different subareas within premotor and parietal regions.
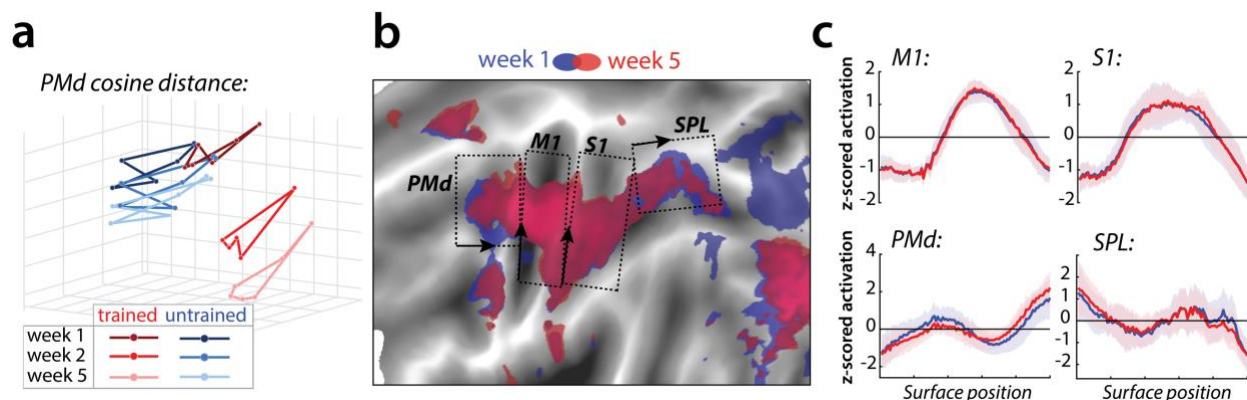
267
268 **Figure 5. Relative change in evoked activation. a)** Multidimensional scaling plot of cosine angle
269 dissimilarities for trained and untrained sequences in premotor dorsal area (PMd) across weeks 1-5. Each
270 dot represents a single sequence, and dots are connected for each session and sequence type separately.
271 Trained sequences on average become more distant from untrained sequences with learning. Untrained

272 sequences on average also progress across weeks, but less than trained sequences. **b)** Normalized
273 activation plots for trained sequences in week 1 (blue) and 5 (red). The arrows and brackets indicate the
274 direction and range of activation cross-sections presented in c). Areas: dorsal premotor cortex (PMd),
275 primary motor cortex (M1), primary somatosensory cortex (S1), superior parietal lobule (SPL). **c)** Cross-
276 section of elicited activation for trained sequences in defined areas, in weeks 1 (blue) and 5 (red).
277

278 ### *Sequence-specific activity patterns reorganize early in learning*

279 Our analyses so far have been concerned with changes in the overall pattern of trained vs.

280 untrained sequences, and showed widespread reductions in activation and some more subtle

281 changes in relative location. The sequence-specific performance advantage, however, indicates

282 that the brain must represent specific sequences – i.e. there should be activity patterns that are

283 unique to each individual sequence. Sequence-specific learning should then be reflected in

284 changes of these sequence-specific activity patterns with learning **(Fig. 1d)**. Consistent with

285 previous results (Wiestler & Diedrichsen, 2013; Yokoi & Diedrichsen, 2019), we detected

286 sequence-specific activity patterns, i.e. activity patterns that differentiate between the tested

287 motor sequences, in various cortical regions, even in session 1 **(Fig. 6a)**. This allowed us to

288 assess their reorganization across sessions.

289   Our pre-registered hypothesis (https://osf.io/etnqc) was that earlier in learning sequence-

290 specific activity patterns would change more for trained than untrained sequences, and would

291 stabilize later in learning. In contrast to the other ideas tested in this paper, this was a novel

292 hypothesis and not based on previous reports. Specifically, we predicted that the correlation of

293 each sequence-specific pattern between weeks 1 and 2 should be lower for trained as compared

294 to untrained sequences. The problem with performing a simple correlation analysis on the

295 patterns, however, is that the estimated correlation will be biased by noise – i.e., more within-

296 session variability for one set of sequences will result in a lower correlation (Diedrichsen, Yokoi,

297 & Arbuckle, 2017). To address this problem, we used the pattern component modelling (PCM)

298 framework which explicitly models and estimates the signal and noise for each session explicitly.

299 Using this approach, we estimated the likelihood of each participants' data under a series of

300 models, each assuming a true correlation in the range between *0* (uncorrelated patterns) and *1*

301 (perfect positive correlation; see Materials and Methods for details). Figure **6b** shows the log-

302 likelihood for each specific correlation model relative to the mean across all models. In SPLa, the

303 most likely correlation of the activity patterns for the trained sequences between weeks 1 and 2

304 was $r = 0.37$. For week 2-5, the likelihood peaked at $r = 0.6$. In contrast, the likelihood functions for

305 untrained sequences indicated that the most likely model was between $r = 0.6\text{-}0.7$ for both week

306 1-2 and 2-5. The advantage of this analysis is that we can be sure that the observed low

307 correlation in week 1-2 for trained sequence was not due to increased noise. In fact, if the noise

308   in one or both sessions was too high, then the model would be unable to distinguish between any

309   of the correlation models – i.e. the likelihood curve would be a flat line.

310   To statistically assess the difference in correlations across trained and untrained

311   sequences, we compared the likelihood of the data of trained sequences between two models:

312   the best-fitting model for the trained sequences (r =0.37 in SPLa) and the correlation model best

313   fitting the data of untrained sequences (r =0.6) (black dots and projections onto y-axis in **Fig. 6b**).

314   To avoid double-dipping, the 'best-fitting' model was chosen on 25 participants (n-1) and the

315   likelihood assessed on the left-out subject (see Materials and Methods). The difference in model

316   evidence was significant for correlation between weeks 1-2 in SPLa ($t_{(25)}$=2.88, p=8.0e-3). In

317   contrast, no difference in correlation was observed later in learning, between weeks 2 and 5

318   ($t_{(25)}$=1.21, p=0.24). A similar pattern of results was observed in PMd**,** with correlation of trained

319   sequences significantly lower than that of untrained sequences between weeks 1 and 2 ($t_{(25)}$=2.93,

320   p=7.2e-3), but not between weeks 2 and 5 ($t_{(25)}$=0.88, p=.39). No such change in correlation across

321   weeks 1-2 was observed in M1 ($t_{(25)}$=0.43, p=.67) or S1 ($t_{(25)}$=1.72, p=0.097). Overall, we found

322   significant evidence that sequence-specific trained patterns in SPLa and PMd reorganize more in

323   weeks 1-2 as compared to the untrained sequences, and stabilize later on with learning, in line

324   with our new pre-registered prediction.

325   To determine more generally where in the neocortex sequence-specific plasticity could be

326   detected, we fit PCM correlation models to regularly tessellated regions spanning the cortical

327   surface. Figure **6c** displays the correlation with the highest evidence for activity patterns across

328   weeks 1-2 and 2-5; separately for trained and untrained sequences. In general, the highest

329   correlations were found in core sensory-motor areas. Across weeks 1-2 for trained sequences,

330   correlations were significantly lower in a number of dorsal premotor, inferior frontal, and parietal

331   regions **(Fig. 6c)**. Across the cortex, correlation for trained patterns increased for weeks 2-5,

332   resulting in similar values which did not differ significantly between trained and untrained

333   sequences for most tessels (see supplementary figure **S6**). Together, these results confirmed that

334   sequence-specific activation patterns in secondary association areas show less stability early in

335   learning, but stabilize later on.

336   Can we obtain further insight into *how* the sequence-specific patterns change in these

337   areas? One specific preregistered prediction was that there would be an increase in

338   distinctiveness (dissimilarity) between fMRI patterns underlying each trained sequence (Wiestler

339   & Diedrichsen, 2013; **Fig. 1e**). To test this hypothesis, we calculated crossnobis dissimilarities

340   (Walther et al., 2016) between sequence-specific activations, separately for trained and untrained

341   sequences. In contrast to our prediction, no significant change in dissimilarity across weeks was

342 observed in any of the predefined regions **(Fig. 6d)**. This suggests that the reorganization

343 observed for trained sequences early in learning did not increase the average distinctiveness of

344 the sequence-specific patterns.
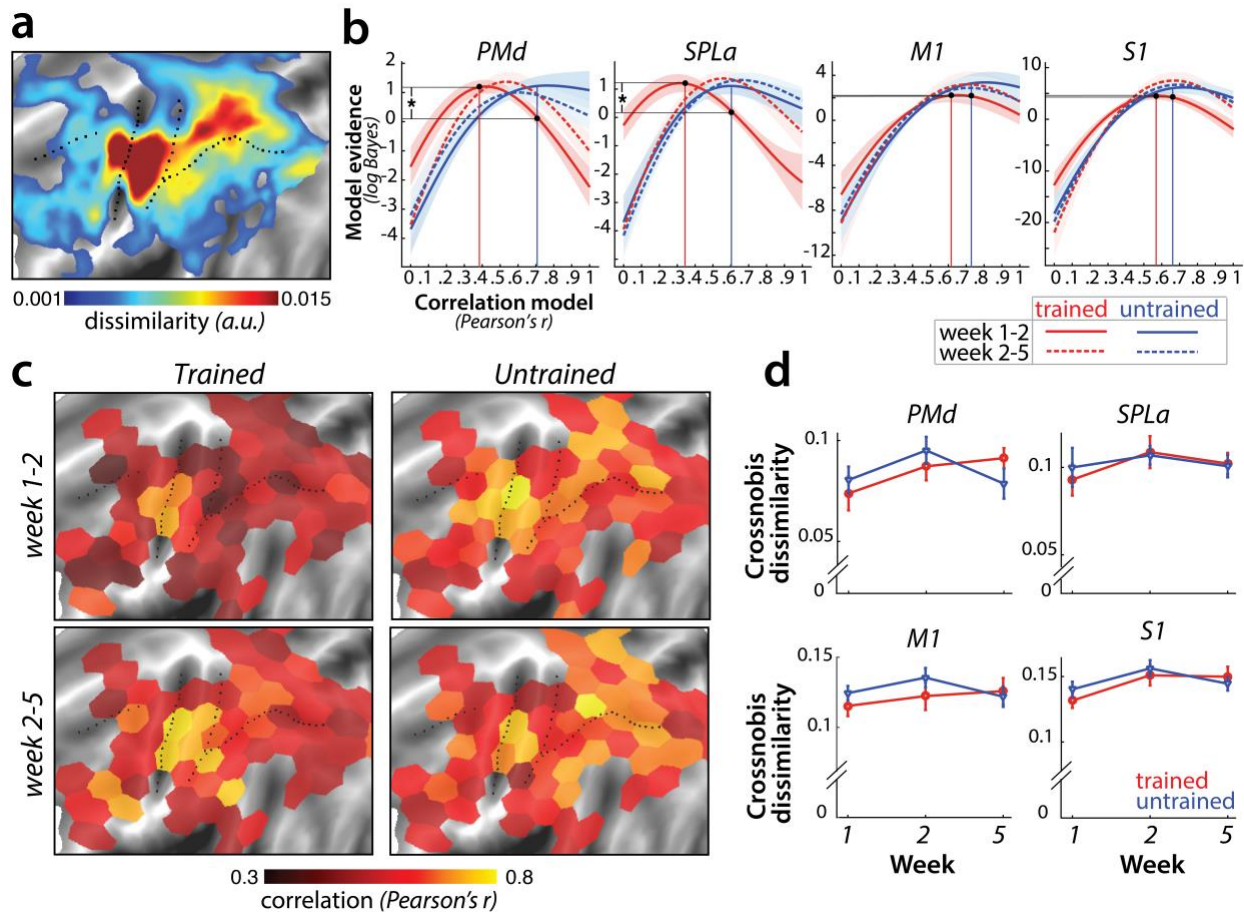
345
346 **Figure 6. Sequence-specific activity patterns reorganize across sessions. a)** Cortical surface map of
347 crossnobis dissimilarities between activity patterns for different sequences in session 1. These regions
348 encode which sequence is executed by the participant. **b)** Evidence explained by models of correlation
349 values between *r* =0 and *r* =1 for sequence-specific patterns across weeks 1-2 (solid) and 2-5 (dashed),
350 separately for trained (red) and untrained (blue) sequences. Evidence was assessed with a type-II log-
351 likelihood, relative to the average log-likelihood across models. Shaded areas indicate standard error
352 across participants. Difference between log-likelihoods can be interpreted as log-Bayes factor, with a
353 difference of 1 indicating positive evidence. Horizontal lines are drawn for the winning correlation model for
354 trained (red) and untrained (blue) patterns across weeks 1-2. Black dots are projections of the two winning
355 models onto the correlation function of trained sequences across weeks 1-2. The horizontal lines from the
356 two black dots indicate the likelihood of the trained data under the two models, which was tested in a
357 crossvalidated t-test. **c)** Map displaying the correlation of the winning model for trained and untrained
358 sequences across weeks 1-2 and 2-5. The correlation of the winning correlation model is shown in all
359 tessels where the difference between evidence for winning model vs. worst-fitting model exceeds log-Bayes
360 factor of 1 (averaged across participants). See **S6** for the difference in best model correlation between
361 trained and untrained sequences, and an indication of tessels where the difference is significant, as based
362 on the crossvalidated t-test. **d)** Crossnobis dissimilarities between trained and untrained sequence pairs

14

363   across weeks. No significant effect of week, sequence type or their interaction was observed in any of the
364   regions. Error bars indicate standard error across participants.
365

366   ***Trained sequences elicit distinct patterns during full speed performance***

367   In the last part of the experiment, we asked whether some of the negative findings (e.g. no

368   changes in M1, no increase in dissimilarities for trained sequences) might have been due to the

369   fact that participants were paced at a relatively slow speed. Matching the speed across sessions

370   allows for the comparisons of changes in neural activity for exactly the same behavioral output

371   (Karni et al., 1995; Lehéricy et al., 2005). However, it could be that controlling for speed impairs

372   our ability to study brain representations of motor skill; simply because after learning, the system

373   is not challenged enough to activate the neuronal representations supporting skilled performance.

374   Consequently, several studies have not (Bassett et al., 2010; Wymbs & Grafton, 2015), or not

375   strictly (Wiestler & Diedrichsen, 2013), matched performance across sessions or levels of training.

376   To examine the effect of performance speed, we added a fourth scanning session *(fs)*, just a day

377   after from the third session in week 5, in which participants were instructed to perform the

378   sequence as fast as possible.

379   Performance during the 4_{th} scan was 1010 ms faster than in the first session ($t_{(25)}$=15.7,

380   p=1.82e-14) and also 338 ms ($t_{(25)}$=9.92, p=4.58e-10) faster for trained than for untrained

381   sequences. Averaged over trained and untrained sequences, we found that the faster

382   performance in this session led to an increase in activity across premotor and parietal areas (**Fig.**

383   **7a,b**). Although trained sequences were executed faster than untrained sequences, activation

384   was still lower for trained compared to untrained sequences, similar to what we observed for

385   paced performance **(Fig. 7c;** see **S7a** for statistical maps). In M1 and S1, we found no difference

386   in activation between trained and untrained sequences (**Fig. 7a**; M1: $t_{(25)}$=1.78, p=.09; S1:

387   $t_{(25)}$=1.69, p=.10). Overall, the pattern of results for evoked activation did not change qualitatively

388   when participants performed at full speed.

389   Next, we examined whether the brain representations of individual sequences are similarly

390   engaged at slow and fast speeds. The correlation between sequence-specific patterns was

391   relatively high (r=0.62) across our regions of interest. We found no differences between the

392   different regions ($F_{(3,75)}$=1.47, p=.23), or sequence types (trained vs. untrained: $F_{(1,25)}$=0.25,

393   p=.62). Thus, the sequence-specific representations activated during performance at high skill

394   level (full speed) are at least partly activated even when performance slowed down.

395   Having established that the mean activation results are replicated across paced and full-

396   speed performance, and that similar sequence-specific representations are activated in both

397   cases, we tested whether activation patterns for different trained sequences are more distinct

15

398 during full speed performance, as reported in Wiestler & Diedrichsen (2013). Overall, crossnobis

399 dissimilarities increased at full speed for trained sequences in PMd and SPLa **(Fig. 7e)**. No such

400 changes were found in M1 or S1. Moreover, trained sequences showed larger dissimilarities than

401 untrained at full-speed performance across premotor and parietal cortices (**Fig. 7f**), which was

402 not the case for the last paced session. In our predefined ROIs, this difference was significant for

403 PMd (**Fig. 7d**), but also parietal areas showed significantly higher dissimilarities between trained

404 sequences at full speed (supplementary figure **S7b**). This suggests that while activity patterns at

405 full speed are correlated to those during paced performance, they are more distinguishable for

406 trained sequences.

407 Could this effect be driven by behavioral performance, with trained sequences performed

408 more differently at full speed (i.e. different speeds across trained sequences), while untrained

409 sequences were performed at a more equal speed? To test for this, we calculated crossnobis

410 dissimilarities between movement times associated with different trained and untrained

411 sequences. The dissimilarities based on speed of performance did not differ significantly across

412 trained and untrained sequences ($t_{(25)}$=0.57, p=.57). Therefore, increased dissimilarity of trained

413 compared to untrained patterns in premotor and parietal areas could not be explained by a

414 difference in execution speed. Instead, this effect likely reflects changes in activity patterns

415 underlying full speed skilled performance.

416

417

16
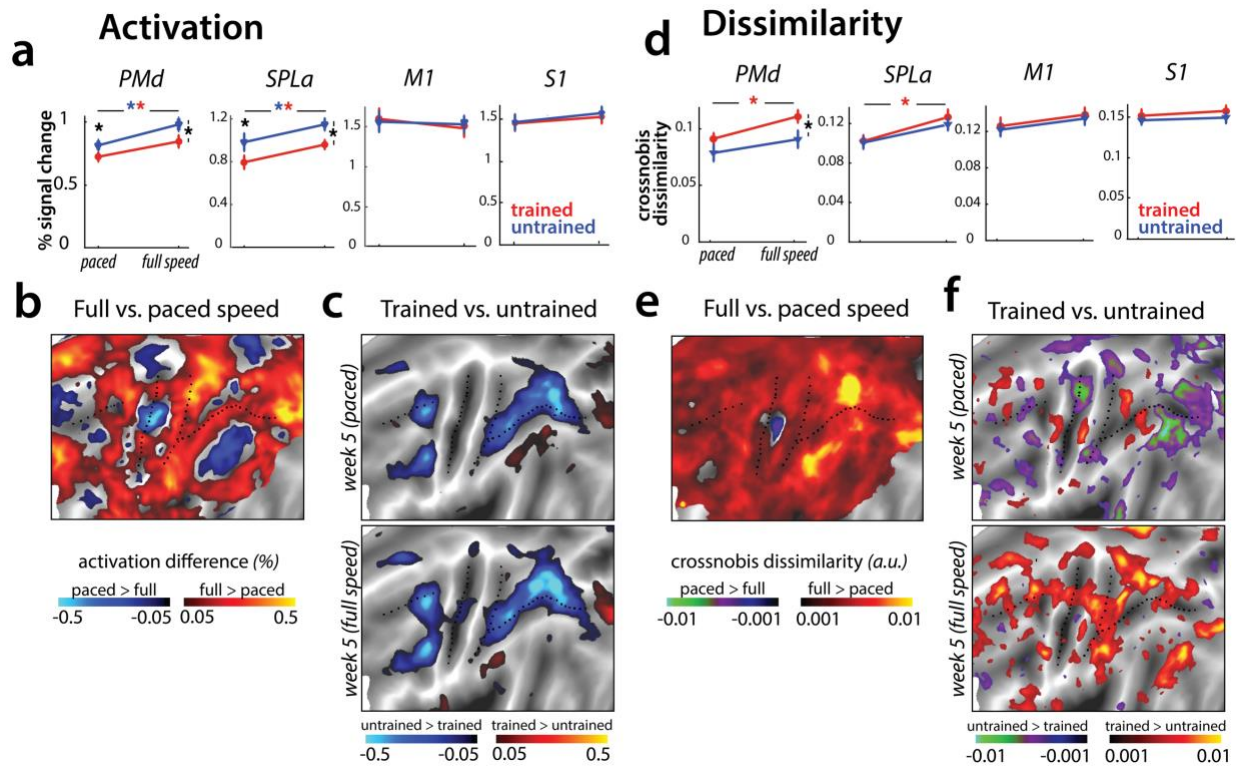
418 **Figure 7. Speed-related changes in activation and dissimilarities. a)** Overall activation in week 5 in
419 paced and full speed sessions for trained (red) and untrained (blue) sequences. Activation was measured
420 as percent signal change over resting baseline (* indicates *p*<.05). Error bars indicate standard error across
421 participants. **b)** Increase in activation for full speed compared to paced speed in percent signal change,
422 averaged across trained and untrained sequences. Red colors indicate an increase in activity during full
423 speed performance compared to paced performance. Blue colors indicate higher activation during paced
424 compared to full speed performance. **c)** Difference in activation elicited for trained relative to untrained
425 sequences, during the paced and full speed sessions (see supplementary figure **S7a** for statistical maps).
426 Trained>untrained is shown in red, untrained>trained in blue. **d)** Average crossnobis dissimilarity between
427 sequence-specific patterns in paced and full speed sessions for trained and untrained sequences.
428 Dissimilarities are significantly larger for trained (red), as compared to untrained (blue) patterns, in PMd for
429 full-speed session (* indicates *p*<.05). Error bars indicate standard error across participants. **e)** Difference
430 between crossnobis dissimilarities across full speed and paced sessions, averaged across trained and
431 untrained sequences. Higher dissimilarities for full speed than paced session are shown in red, whereas
432 blue/green hues indicate higher dissimilarities during paced than full speed session. **f)** Difference in
433 dissimilarities for trained relative to untrained sequences, during the paced and full speed sessions.
434 Trained>untrained is shown in red, untrained>trained in blue/green. Trained sequences elicited higher
435 dissimilarities than untrained in full speed, but not paced session (see **S7b** for statistical *t*-maps).

436

437 ### *Striatal activity patterns for trained sequences manifest at full speed performance*

438 We observed learning-related changes in cortical association areas, but not in the primary motor

439 cortex. Of course, learning could also be driven by neuronal changes in subcortical brain regions

440 (Ashby, Turner, & Horvitz, 2010; Graybiel, 2016; Graybiel & Grafton, 2015; Hikosaka et al., 1999;

441 Yin et al., 2009). The striatum in particular has been proposed as a structure where motor skills

442 are stored (Kawai et al., 2015; Lehéricy et al., 2006). Inspecting changes in overall activity across

443 sessions, we observed no difference in activity between trained and untrained sequences in either

444 putamen or caudate nucleus **(Fig. 8a)**.

445 　　　　Previous experiments have reported that with learning, activation moves from more

446 'cognitive' areas of the striatum (i.e. caudate nucleus) to more 'motor' areas (i.e. putamen)

447 (Coynel et al., 2010; Lehéricy et al., 2005; Reithler, van Mier, & Goebel, 2010). Our data fails to

448 replicate this result: Both the visual inspection **(Fig. 8b)**, and statistical quantification of the mean

449 pattern difference for trained and untrained sequences across sessions revealed no such

450 learning-specific shift of mean striatal activation pattern with learning.

451 　　　　Lastly, we examined if the striatum represents individual sequences. During the paced

452 sessions, activity patterns for different sequences were not distinguishable in either caudate

453 nucleus or putamen **(Fig. 8c)**. However, during full speed performance trained sequences elicited

454 distinct activity patterns in both regions (i.e. crossnobis dissimilarity>0: caudate nucleus:

455 $t_{(25)}$=2.27, *p*=0.032; putamen: $t_{(25)}$=2.44, *p*=.022; **Fig. 8c**). This effect was specific to the trained

456 sequences, with untrained sequences still exhibiting undistinguishable patterns of activity at full

457 speed. Thus, we found some evidence that trained motor sequences are represented in the form

458 of distinct activity patterns in the striatum during full speed skilled performance.
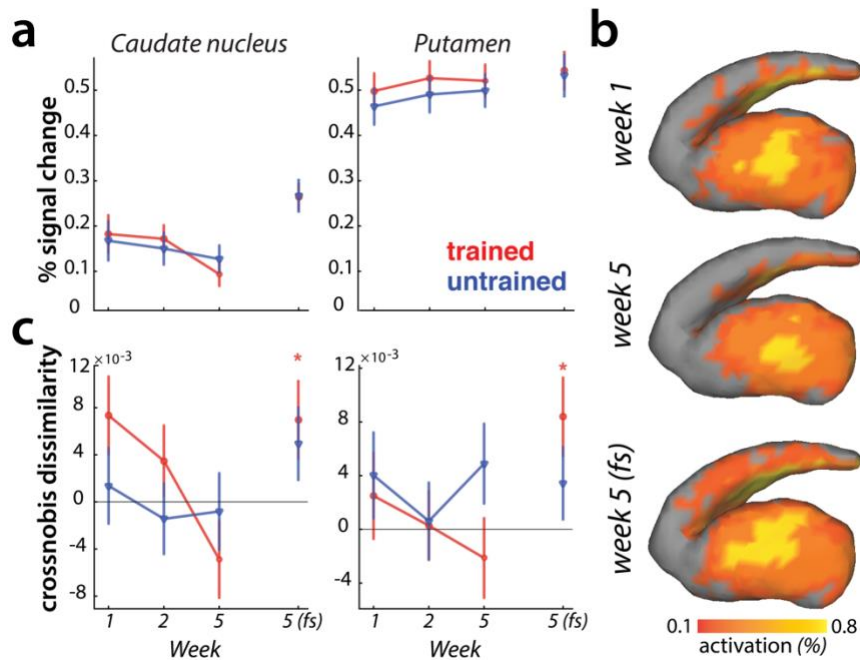
459

460 **Figure 8. Striatal changes in activation and dissimilarities with learning. a)** Overall activation (percent
461 signal change over resting baseline), for trained (red) and untrained (blue) sequences. Activation did not
462 differ across sessions, or sequence types in the striatum. Error bars indicate the standard error across
463 participants. **b)** Activation during performance of trained sequences in the striatum across weeks 1, 5
464 (paced speed) and 5 (full speed – *fs*), averaged across sequences and participants. **c)** Crossnobis
465 dissimilarities between activation patterns of sequence pairs, calculated separately for trained and
466 untrained patterns. Dissimilarities were not significantly different for trained or untrained sequences during
467 paced performance. At full speed, sequence-specific activity patterns amongst trained sequences differed
468 significantly in both caudate nucleus and pallidum (* indicates *p*<.05). Error bars indicate the standard error
469 across participants.

18

Discussion

470
471

472 Here we present a large longitudinal motor sequence learning study that allowed us to

473 systematically investigate several previously proposed fMRI signatures of motor learning,

474 including one new hypothesis concerning the change in multivariate activity patterns with learning.

475 The existing literature, with its diversity of experimental protocols and analysis approaches, does

476 currently not provide a consistent picture of learning-related changes. This inconsistency is

477 exacerbated by the fact that most papers prioritize making new claims over re-examining

478 previously established findings. Consequently, it is very hard to assess the replicability of most

479 past findings. We address this issue here by a) producing a well-powered, longitudinal data set

480 that tackles some of the methodological inconsistencies (i.e. speed matching), b) pre-registering

481 both design and hypotheses, and c) making data and analysis pipelines openly available, such

482 that other hypotheses and analyses techniques can be freely tested.

483   Our findings reveal that parietal and premotor areas show widespread decreases in overall

484 activation, as well as reorganization of sequence-specific patterns early in learning. Additionally,

485 we observed that sequence specific patterns in these areas (as well as the striatum) were more

486 distinct during full speed performance. In contrast to this set of results, none of these learning-

487 specific metrics were detected in M1, even after 5 weeks of training.

488   On the one hand, our lack of any observable change in M1 activation contradicts some

489 prior results, where increased activation in M1 was observed for matched performance after

490 learning (Karni et al., 1995; Matsuzaka, Picard, & Strick, 2007; Penhune & Doyon, 2002; Steele

491 & Penhune, 2010; Vahdat et al., 2015), and does not align with reports of M1 stimulations

492 influencing consolidation or storage of motor skills (in motor sequence tasks: Kang & Paik, 2011;

493 Nitsche et al., 2003; Reis et al., 2009; Waters-metenier, Husain, & Wiestler, 2014; in other motor

494 tasks: Classen, Liepert, Wise, Hallett, & Cohen, 1998; Galea, Vazquez, Pasricha, Orban De Xivry,

495 & Celnik, 2011; Hadipour-Niktarash, Lee, Desmond, & Shadmehr, 2007). We also found no

496 support for a combination of increases and decreases of activation with training, which would lead

497 to an overall change of the mean activity pattern (Steele & Penhune, 2010).

498   Instead, our results suggest that the pattern of neural activity in M1 does not change as

499 participants become more skilled at producing motor sequences. This is consistent with a recent

500 line of evidence demonstrating that M1 does not change activation with learning (Huang et al.,

501 2013), and primarily encodes single movement elements, rather than sequences (Yokoi,

502 Arbuckle, & Diedrichsen, 2018; Russo et al., 2019). Somewhat more surprisingly, we also

503 observed no difference in overall M1 activation during full speed performance, when performance

504   was considerably faster for trained sequences. This suggests that the activity increases related

505   to faster movement speeds are compensated for by the shorter duration spent on the task.

506   Primary somatosensory cortex in many ways paralleled the results observed in M1. We

507   observed no overall activation change, or change in the sequence-specific pattern correlation

508   across sessions. The only exception was the observed shift in the mean activation pattern across

509   sessions. One possible explanation is that feedback-related sensory activity in S1 undergoes

510   some plastic changes with learning. This is consistent with a recent study demonstrating that S1,

511   but not M1, is involved during consolidation of motor skills (Kumar, Manning, & Ostry, 2019; for a

512   review on somatosensory plasticity in motor learning see Ostry & Gribble, 2016).

513   In contrast to the limited evidence of learning-related changes in primary somatosensory

514   and primary motor areas, higher order association areas (e.g. parietal and premotor cortices)

515   displayed an array of learning-related changes. First, activation decreased in areas involved in

516   sequence execution, with larger decreases for trained as compared to untrained sequences. This

517   result contrasts with other previous studies reporting increases in activation in premotor areas

518   with learning (Grafton et al., 2002; Honda et al., 1998; Penhune & Doyon, 2002; Vahdat et al.,

519   2015). Partially responsible for these inconsistencies may be a publication bias, favoring reports

520   of signal increases over signal decreases with learning. For example, a recent metanalysis

521   reanalyzed evidence for signal increases in the main text, while moving the (matched) evidence

522   for signal decreases into the supplementary materials (Hardwick et al., 2013). Our data

523   corroborates a number of recent studies reporting reduced activation in task-evoked premotor

524   and parietal areas (Steele & Penhune, 2010; Wiestler & Diedrichsen, 2013; Wu et al., 2004).

525   The only activation increases for trained relative to untrained sequences were observed

526   in areas that were suppressed below baseline during sequence execution. This has also been

527   previously reported in a motor sequence learning study (Tamás Kincses et al., 2008), where

528   deactivation was larger during performance of trained than random sequences. These areas

529   include the precuneous, temporal parietal junction and the cingulate, regions commonly assigned

530   to the default mode network (Raichle et al., 2001; Shulman et al., 1997). This group of regions is

531   more activated during rest than during task performance, and has been associated with functions

532   such as episodic memory retrieval and attention to internal states (Andrews-Hanna, Reidler,

533   Sepulcre, Poulin, & Buckner, 2010; Gusnard, Akbudak, Shulman, & Raichle, 2001). Our

534   observation of decreased inhibition of the default mode network likely reflects central attentional

535   resources being freed up, allowing participants to engage in other mental processes (e.g.,

536   daydreaming) while performing the task. Thus, this release from initial deactivation is possibly

537   task-irrelevant, reflecting increased automaticity with learning (Shamloo & Helie, 2016).

538   Overall, changes in average activation are relatively hard to interpret, as they could reflect
539   a combination of numerous factors. As a more direct fMRI metric of plasticity, we suggest to
540   inspect changes in the sequence-specific activity patterns, since these constitute a more likely
541   fMRI correlate of the sequences-specific performance advantage observed after training. In this
542   project, this provided us with two key insights of how activation patterns reorganize in association
543   areas with learning. First, activity patterns associated with each of individual trained sequences,
544   changed to a greater extent earlier in learning, and stabilized later. This finding resonates with
545   several animal studies suggesting that the emergence of skilled behavior is associated with early
546   plasticity and later stabilization of neuronal activity patterns (Makino et al., 2017; Peters et al.,
547   2017). Here we report a similar effect in humans, and advance these findings by demonstrating
548   that this reorganization occurs at the level of sequence-specific patterns. In past studies using
549   rodent models, sequence-specific patterns could not be dissociated from the overall activity
550   pattern, as the animals were only trained on production of a single sequence. Additionally, by
551   pacing participants' speed, we were able to cleanly dissociate changes in the organization of
552   activity patterns from changes in the behavioral performance or variability. Second, activation
553   patterns became more distinct for trained sequences at full speed. This indicates that the
554   engagement of specific neuronal subpopulations for different sequences is particularly important
555   when pushing the limit of performance.

556   While our study focused on the role of cortical areas in motor sequence learning, we also
557   examined activation in the striatum, which has been suggested to play a critical role in skilled
558   performance (Graybiel & Grafton, 2015; Kawai et al., 2015; Otchy et al., 2015). In contrast to
559   previous fMRI studies (Coynel et al., 2010; Lehéricy et al., 2006; Reithler, van Mier, & Goebel,
560   2010), we did not find clear evidence for differences in overall activity, or shifts of the overall
561   activity pattern with learning. Nonetheless, we observed distinguishable striatal activation patterns
562   for different trained sequences at full speed, in line with a recent report showing distinguishable
563   striatal patterns for performance of consolidated motor sequences (Pinsard et al., 2018). While
564   by itself the finding of differential sequence-specific activity patterns is not evidence for a causal
565   role of the striatum in the production of skilled behaviors, it is a necessary condition for such a
566   functional role. Therefore, our results here are in line with the proposed involvement of the
567   striatum in motor sequence learning.

568   An important feature of our design was that we collected imaging data in the trained state,
569   both when performance was clamped to the initial speed, and when participants performed as
570   fast as possible. Previous studies have usually included only one of these two options, making
571   direct comparisons difficult (see Lutz et al., 2004 for an examination of various execution speeds

572　　on BOLD activity and Orban et al., 2010 in a motor learning context). Our results provide two
573　　important insights: first, in terms of the overall fMRI activation, the pattern of results remained the
574　　same for paced vs. full speed performance. This indicates that, in this specific case, the increased
575　　motor demands and the decreased time on task averaged out. In general, however, these two
576　　factors may not balance perfectly – therefore paced performance may be a better choice when
577　　comparing overall activation across sessions. Second, even though slow and paced performance
578　　in the trained state activated sequence-specific activation patterns, these were much stronger
579　　when performing at maximal speeds. Thus, for questions regarding the fine-grained patterns, it
580　　might be more suitable to challenge the system fully.

581　　　　　Of course, our list of inspected fMRI metrics of learning was not exhaustive. For instance,
582　　we did not investigate whether various fMRI correlates of learning predict behavioral outcomes,
583　　or how functional connectivity and network metrics change with learning, partly because of the
584　　absence of specific predictions. Pre-registration of hypothesis are especially important for these
585　　analyses, since the search space of possible tests becomes exponentially larger (e.g. correlating
586　　all possible brain metrics with all possible behavioral metrics; or using various metrics to assess
587　　inter-regional relationships). However, we hope that our dataset, upon its public release, can
588　　serve as a resource for other researchers to (re-)test novel predictions about learning related
589　　changes.

590

591　　## Conclusion

592　　The search for neural substrates of learning is a daunting task: the acquisition of longitudinal data
593　　sets is work intensive, and the large dimensionality of possible brain metrics makes the search
594　　difficult (Poldrack, 2000). Historically, the question was simplified by studying activation increases
595　　in single areas as proxies for motor 'engram' localization (Berlot, Popp, & Diedrichsen, 2018).
596　　Here we found no evidence for such activation increases; instead we observed widespread and
597　　distributed decreases in activation across cortical areas. In contrast, subtler changes in the
598　　distributed patterns of fMRI activity have the potential to provide more direct metrics of plasticity.
599　　Increased pattern reorganization (across weeks), and larger pattern separation for trained
600　　sequences was found across prefrontal, parietal, and striatal regions. These metrics may be
601　　useful as general fMRI correlates of neural reorganization beyond the domain of motor learning.

Materials and Methods

Participants

Twenty-seven volunteers participated in the experiment. One of them was excluded because field map acquisition was distorted in one of the four scans. The average age of the remaining 26 participants was 22.2 years (SD = 3.3 years), and the sample included 17 women and 9 men. All participants were right-handed and had no prior history of psychiatric or neurological disorders. They provided written informed consent to all procedures and data usage before the study started. The experimental procedures were approved by the Ethics Committee at Western University.

Apparatus

Participants performed finger sequences with their right hand on an MRI-compatible keyboard **(Fig. 2a)**, with keys numbered 1-5 for thumb-little finger. The keys had a groove for each fingertip and were not depressible. The force of isometric finger presses was measured by the force transducers (FSG-15N1A, Sensing and Control, Honeywell; dynamic range 0-25 N) mounted underneath each key with an update rate of 2 ms. A key press was recognized when the sensor force exceeded 1 N. The measured signal was amplified and sampled at 200 Hz.

Learning paradigm

Participants were trained to execute six 9-digit finger sequences over a period of five weeks **(Fig. 2a)**. They were split into two groups with trained sequences of one group constituting the untrained sequences for the other group and vice versa. Finger sequences of both groups were matched as closely as possible in terms of the starting finger, number of finger repetitions in a sequence and first-order finger transitions. This counterbalancing between the groups ensured that any of the observed results were not specific to a set of chosen trained sequences.

In the pre-training session prior to the first scan **(Fig. 2b)**, participants were acquainted with the apparatus and task performed during scanning. Sequences executed during this pre-training session were not encountered later on in the experiment.

During the training sessions, participants were trained to perform the six sequences as fast as possible. They received visual feedback for the correctness of their presses with digits turning green for a correct finger press and red for an incorrect one. After each trial, participants received points based on the accuracy and their movement time (MT – time from the first press until the last finger release in the sequence; **Fig. 2c**). Trials executed correctly and faster than participant's median MT from the previous blocks were rewarded with 1 point. If participants

23

637   performed correctly and 20% faster than the median MT from previous blocks, they received 3

638   points. If they made a mistake or performed below their median MT, they received 0 points.

639   Participants performed each sequence twice in a row: digits were written on the screen for the

640   first execution, but removed for the second execution so that participants had to perform the finger

641   sequence from memory. Training sessions were broken into several blocks, each consisting of

642   24 trials (4 trials per trained sequence), with time between blocks to rest. At the end of each block,

643   participants received feedback on their error rate, median MT and points obtained during the

644   block. If participants performed with an error of <15% and faster than the previous median MT,

645   the MT threshold was updated. This design feature was chosen to maintain participants'

646   motivation to execute the sequences as fast as possible, within the allowed error range.

647   During the behavioral test sessions **(Fig. 2d)**, participants executed sequences they were

648   trained on as well as untrained matched sequences, which were randomly interspersed. All

649   sequences were still performed twice in a row, with numbers on the screen present on both

650   executions.

651

652   ***Experimental design during scanning***

653   Participants underwent four scanning sessions **(Fig. 2d)** – with the first one before learning regime

654   started, the second after a week and two more scans after completion of the 5 training weeks.

655   Each scanning session consisted of eight functional runs. We employed an event-related design,

656   randomly intermixing execution of trained and untrained sequences. Each sequence was

657   repeated twice in a row (with digits always present on the screen), and there was a total of six

658   repetitions per sequence in every run. Each trial started with 1 second preparation time, during

659   which the sequence was presented on the screen. After that time, a 'go' signal was displayed as

660   short pink line underneath the sequence numbers. In scanning sessions 1-3, this line started

661   expanding below the written numbers, indicating the speed at which participants were required to

662   press along. In scanning session 4, only a short line was presented in front and underneath the

663   sequences. When the line disappeared, this signaled a 'go' cue for participants to execute the

664   presented sequence as fast as possible. The execution phase including the feedback on overall

665   performance lasted for 3.5 seconds, and the inter-trial interval was 0.5 seconds (see

666   supplementary figure **S2**). Each trial lasted for 5 seconds. Five periods of rest, each 10 seconds

667   long, were added randomly between trials in each run to provide a better estimate of baseline

668   activation.

669

670   ***Image acquisition***

24

671 Data was acquired on a 7-Tesla Siemens Magnetom scanner with a 32-receive channel head coil

672 (8-channel parallel transmit). Anatomical T1-weighted scan was acquired at the beginning of the

673 first scanning session, using a magnetization-prepared rapid gradient echo sequence (MPRAGE)

674 with voxel size of 0.75x0.75x0.75 mm isotropic (field of view = 208 x 157 x 110 mm [A-P; R-L; F-

675 H], encoding direction coronal). Functional data were acquired using a sequence (GRAPPA 3,

676 multi-band acceleration factor 2, repetition time [TR] = 1.0 s, echo time [TE] = 20 ms, flip angle

677 [FA] = 30 deg). We acquired 44 slices with isotropic voxel size of 2x2x2 mm. For estimating

678 magnetic field inhomogeneities, we additionally acquired a gradient echo field map. Acquisition

679 was in the transversal orientation with field of view 210 x 210 x 160 mm and 64 slices with 2.5

680 mm thickness (TR = 475 ms, TE = 4.08 ms, FA = 35 deg).

681

682 ***First-level analysis***

683 Functional data were analyzed using SPM12 and custom written MATLAB code. Functional runs

684 were corrected for geometric distortions using fieldmap data (Hutton et al., 2002), and head

685 movements during the scan (3 translations: x, y, z; 3 rotations: pitch, roll, yaw), and aligned across

686 sessions to the first run of the first session. The functional data were then co-registered to the

687 anatomical scan. No smoothing or normalization to an atlas template was performed.

688 Preprocessed data were analyzed using a general linear model (GLM; Friston et al.,

689 1994). Each of the performed sequences was defined as a separate regressor per imaging run,

690 resulting in 12 regressors per run (6 trained, 6 untrained sequences), together with intercept for

691 each of the functional runs. The regressor was a boxcar function starting at the beginning of the

692 trial and lasting for trial duration. The boxcar function was convolved with a hemodynamic

693 response function, with a time to peak of 5.5 seconds, and a manually adjusted onset to best fit

694 each participant's average evoked response. This analysis resulted in one activation estimate

695 (beta image) for each of the 12 conditions per run, in each scanning session.

696

697 ***Surface reconstruction and regions of interest***

698 We reconstructed individual subjects' cortical surfaces using FreeSurfer (Dale, Fischl, & Sereno,

699 1999). All individual surfaces were aligned to the FreeSurfer's Left-Right symmetric template

700 (workbench, 164k nodes) via spherical registration. To detect sequence representation across

701 the cortical surface, we used a surface-based searchlight approach (Oosterhof, Wiestler,

702 Downing, & Diedrichsen, 2011), where for each node we selected a circular region of 120 voxels

703 in the grey matter. The resulting analyses (dissimilarities between sequence-specific activity

704 patterns, see below) was assigned to the center node. As a slightly coarser alternative to

705 searchlights, we performed regular tessellation of cortical surface into 162 tessels per
706 hemisphere. This allowed us to fit correlation models (see below) across the cortical surface,
707 while not being as computationally intensive as searchlight analyses.

708 We defined four regions of interest to cover primary somato-motor regions as well as
709 secondary associative regions. M1 was defined using probabilistic cytoarchitectonic map (Fischl
710 et al., 2008) by including nodes with the highest probability of belonging to Brodmann area (BA)
711 4, while excluding nodes more than 2.5 cm from the hand knob (Yousry et al., 1997). Similarly,
712 S1 was defined as nodes related to hand representation in BA 1, 2 and 3. Additionally, we included
713 dorsal premotor cortex (PMd) as the lateral part of the middle frontal gyrus. The anterior part of
714 the superior parietal lobule (SPLa) was defined to include anterior, medial and ventral intraparietal
715 sulcus. We also defined caudate nucleus and putamen as striatal regions of interest. The
716 definition was carried out in each subject using FSL's subcortical segmentation.

717

718 ***Changes in overall activation***
719 We calculated the average percent signal change for trained and untrained sequences (averaged
720 across the 6 trained and 6 untrained sequences) relative to the baseline for each voxel. The
721 resulting volume map was projected to the surface for each subject, and a group statistical *t*-map
722 was generated across subjects. Statistical maps were thresholded at $p<.01$, uncorrected, and the
723 family-wise error corrected *p*-value for the size of the peak activation and activation cluster size
724 was determined using a permutation test. Specifically, we ran 1000 simulations where we
725 randomly flipped the sign of the contrast for subjects (chosen at random out of 2_{26} possible
726 permutations). The rationale behind this is that under the null hypothesis, there should be no
727 difference between the two conditions, and the sign of each contrast should be interchangeable.
728 As for the data, we thresholded the statistical map from each permutation, and recorded the peak
729 *t*-value (across the map) and the size of the largest cluster. The real data was then compared
730 against this distribution to assess the probability of the observed *t*-value and cluster-size under
731 the null hypothesis.

732 Additionally, we assessed changes in percent signal in predefined regions of interest (M1,
733 S1, PMd, SPLa). This was performed in the native volume space of each subject. To do so, we
734 averaged the percent signal change of voxels belonged to a defined region per subject and
735 quantified activation changes across subjects using ANOVAs and t-tests across subjects.

736 Besides overall activation, we also examined *relative* changes in elicited activation for
737 trained sequences across sessions. This was done by normalizing (z-scoring) the percent signal
738 change surface maps across voxels, separately for each subject. Normalization was applied both

26

739   map-wise (for **Fig. 5b**), as well as for each of the pre-defined ROIs separately (for cross-sections

740   in **Fig. 5c**). Statistical assessment of the difference between relative evoked activation pattern for

741   trained vs. untrained sequence was carried out by calculating cosine angle dissimilarities between

742   the mean evoked patterns. Cosine angle dissimilarity was chosen because it is not sensitive to

743   overall magnitude in activation, and therefore assesses the difference in the relative activation

744   distribution.

745

746   ***Dissimilarities between sequence-specific activity patterns***

747   To evaluate which cortical areas display sequence-specific encoding, we performed a searchlight

748   analysis calculating the dissimilarities between evoked beta patterns of individual sequences.

749   Beta patterns were first multivariately prewhitened (standardized by voxels' residuals and

750   weighted by the voxel covariance matrix), which has been found to increase the reliability of

751   dissimilarity estimates (Walther et al., 2016). We then calculated the cross-validated squared

752   Mahalanobis dissimilarities (i.e. crossnobis dissimilarities) between evoked sequence patterns

753   (66 dissimilarity pairs for 6 trained and 6 untrained sequences). These dissimilarities were then

754   averaged overall, as well as separately for pairs within trained sequences, and within untrained

755   sequences. This metric was used both for searchlight analysis and calculation of metric within

756   predefined regions (cortical and striatal). The cortex surface maps contrasting dissimilarities

757   between trained and untrained sequences were corrected for multiple comparisons using

758   permutations, as described above for percent signal change surface maps.

759

760   ***Pattern component analyses: modelling sequence-specific correlation across sessions***

761   Correspondence of sequence-specific patterns across sessions was quantified using pattern

762   component modelling (PCM; Diedrichsen et al., 2017). This framework is superior at estimating

763   correlations than simply performing Pearson's correlation on raw activity patterns, or even in a

764   crossvalidated fashion. The main problem with estimating correlations on data is that activation

765   patterns are biased by noise, which varies across scanning sessions, and would therefore

766   underestimate the true correlation. PCM separately models the noise and signal component, and

767   can in this way combat the issue more than simply performing crossvalidation would. We

768   designed 30 correlation models with correlations between 0 and 1 in equal step sizes and

769   assessed the group likelihood of the observed data under each model.

770   Subsequent group inferences were performed using crossvalidated approach on

771   assessing individual log-Bayes factors (model evidence). A crossvalidated approach was used to

772   ensure that our choice of 'best-fitting models' and the evidence associated was independent and

773   did not involve double-dipping. Specifically, we used n-1 subjects to determine the best-fitting

774   models for trained and untrained patterns and recorded the log-Bayes factors for those two

775   correlation models on the left-out subject. This was repeated across all subjects and a t-test was

776   performed on the recorded log-Bayes factors (i.e. out-of-sample model evidences). The same

777   evaluation was performed for pre-defined regions of interest (**Fig. 6b**), as well as a regular

778   tessellation across the cortical surface (**Fig. 6c**).

779

780

781   **Acknowledgements**

786

787

788   **Author contributions**

789   EB, NJP and JD designed the experiment; EB and NJP programmed the experiment;

790   EB and NJP collected the data; EB analyzed the data; EB prepared figures; EB drafted

791   manuscript; EB, NJP and JD edited and revised the manuscript.

792

793

794   **Competing interests**

795   The authors declare no competing interests.

28

796 **Supplementary Figures**

797
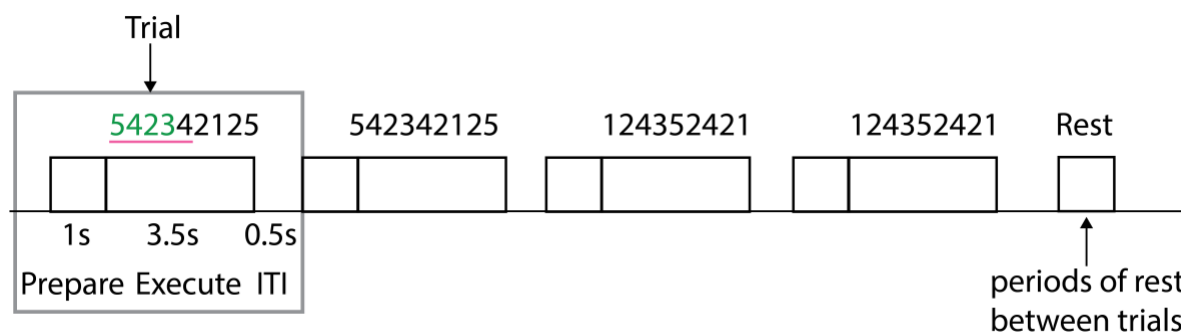
798
799 **Figure S2. Experimental trial structure during scanning sessions.** Each trial consisted of a preparation
800 period, execution period and inter-trial-interval (ITI), during which the feedback was presented on
801 correctness of the trial. Each sequence was presented twice in a row. Periods of rest were added in-
802 between the trials.

week 1   week 2   week 5

activation difference (t_{25})

untrained > trained   trained > untrained

-5   -2.5   2.5   5

Week 2:

#	peak t_{25}	P_{t_unc}	P_{t_corr}	cm^2	$P_{cluster}$
1	2.85	.010	.34	0.5	.073
2	4.28	$3.13e^{-4}$	.003	3.4	.003
3	4.03	$5.88e^{-4}$	.010	2.3	.008
4	3.95	$7.16e^{-4}$	.012	3.9	<.001
5	4.13	$4.59e^{-4}$	.006	2.5	.003
6	4.23	$3.50e^{-4}$	.004	7.1	<.001

Week 5:

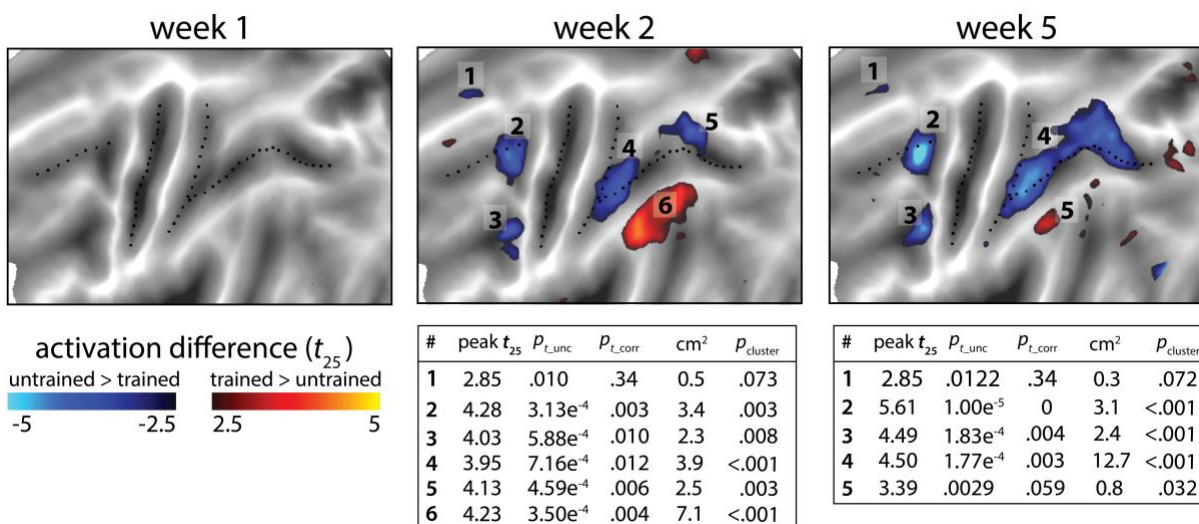| #	peak t_{25}	P_{t_unc}	P_{t_corr}	cm^2	$P_{cluster}$
1	2.85	.0122	.34	0.3	.072
2	5.61	$1.00e^{-5}$	0	3.1	<.001
3	4.49	$1.83e^{-4}$	.004	2.4	<.001
4	4.50	$1.77e^{-4}$	.003	12.7	<.001
5	3.39	.0029	.059	0.8	.032

Figure S4. Statistical maps for the trained vs. untrained contrasts on elicited activation in each session. Trained>untrained is shown in red, untrained>trained in blue. Maps were thresholded at a $t_{25}=\pm2.5$, $p<.01$ uncorrected for a two-sided t-test. Tables show peak t-value and size (in cm_2) for each super-threshold cluster (indicated by numbers) for maps of week 2 and 5. p_{t_unc} is the uncorrected p-value for the peak of each cluster. Family-wise error corrected p-values were determined using permutation testing for the peak t-value (p_{t_corr}) and cluster size ($p_{cluster}$).

Trained vs. untrained max correlation
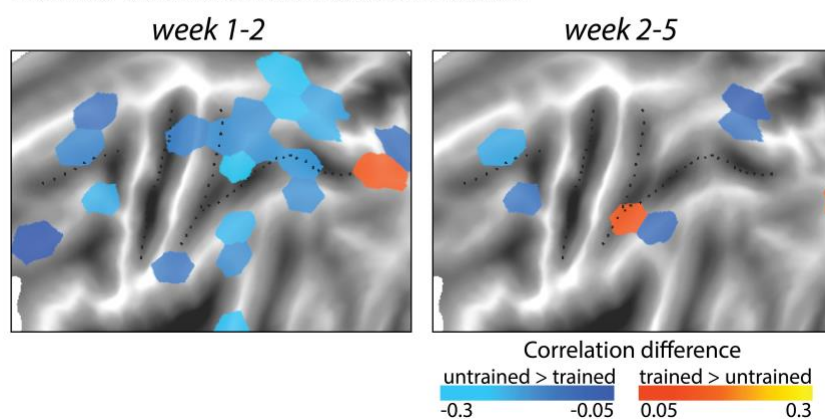
810

811 **Figure S6. Difference between correlation of winning model for trained and untrained sequences.**
812 Difference between the correlations of the winner models for trained and untrained sequences, separately
813 for week 1-2 and week 2-5. Blue indicates a lower correlation across weeks for trained than untrained
814 patterns of activity. The correlation difference values are plotted in tessels where the difference in model
815 evidence was significant, as based on the cross-validated *t*-test (for two-sided *p*<.05).

week 5 (paced)

week 5 (full speed)

a — Activation

Week 5 (paced):

#	peak t_{25}	P_{t_unc}	P_{t_corr}	cm²	$p_{cluster}$
1	2.85	.012	.34	0.3	.072
2	5.61	$1.00e^{-5}$	<.001	3.1	<.001
3	4.49	$1.83e^{-4}$	.004	2.4	<.001
4	4.50	$1.77e^{-4}$	.003	12.7	<.001
5	3.39	.0029	.059	0.8	.032

Week 5 (full speed):

#	peak t_{25}	P_{t_unc}	P_{t_corr}	cm²	$p_{cluster}$
1	3.55	.0020	.047	11.1	.028
2	6.30	$1.68e^{-6}$	<.001	1.0	<.001
3	5.36	$2.57e^{-3}$	<.001	13.4	<.001

activation difference (t_{25}) — trained > untrained (5 / 2.5) / untrained > trained (−2.5 / −5)

b — Dissimilarity

Week 5 (paced):

#	peak t_{25}	P_{t_unc}	P_{t_corr}	cm²	$p_{cluster}$
1	3.85	$9.40e^{-4}$	.266	1.4	.047
2	3.41	.0029	.52	5.1	.006

Week 5 (full speed):

#	peak t_{25}	P_{t_unc}	P_{t_corr}	cm²	$p_{cluster}$
1	2.77	.012	.85	2.2	.023
2	3.48	.0024	.44	1.9	.023
3	4.43	$2.11e^{-4}$	.068	5.0	.006

crossnobis dissimilarity (t_{25}) — trained > untrained (4 / 1.7) / untrained > trained (−1.7 / −4)

816

817 **Figure S7. Statistical maps for trained vs. untrained contrasts in week 5 (paced) and 5* (full speed)**
818 **sessions.** Trained>untrained is shown in red, untrained>trained in blue. **a)** Statistical contrast for average
819 activation. Maps were thresholded at a t_{25}=± 2.5, p<.01 uncorrected for a two-sided t-test. Tables show
820 peak t-value and size (in cm²) for each super-threshold cluster. p_{t_unc} is the uncorrected p-value for the peak
821 of each cluster. Family-wise error corrected p-values were determined using permutation testing for the
822 peak t-value (p_{t_corr}) and cluster size ($p_{cluster}$). **b)** Statistical contrast for average dissimilarity of sequence-
823 specific activity pattern. Map was thresholded at t_{25}=± 1.7, p<.05, uncorrected. Statistical quantification
824 using permutation tests is in the table below each map.

32

References:

Andrews-Hanna, J. R., Reidler, J. S., Sepulcre, J., Poulin, R., & Buckner, R. L. (2010). Functional-Anatomic Fractionation of the Brain's Default Network. *Neuron*, *65*(4), 550–562. http://doi.org/10.1016/j.neuron.2010.02.005

Ashby, F. G., Turner, B. O., & Horvitz, J. C. (2010). Cortical and basal ganglia contributions to habit learning and automaticity. *Trends in Cognitive Sciences*, *14*(5), 208–215. http://doi.org/10.1016/j.tics.2010.02.001

Bassett, D. S., Wymbs, N. F., Porter, M. A., Mucha, P. J., Carlson, J. M., & Grafton, S. T. (2010). Dynamic reconfiguration of human brain networks during learning. *Proceedings of the National Academy of Sciences*, *108*(18), 7641–7646. http://doi.org/10.1073/pnas.1018985108

Bassett, D. S., Yang, M., Wymbs, N. F., & Grafton, S. T. (2015). Learning-Induced Autonomy of Sensorimotor Systems. *Nature Neuroscience*, *18*(5), 744–751. http://doi.org/10.1038/nn.3993

Berlot, E., Popp, N., & Diedrichsen, J. (2017, August 29). Neural correlates of long-term motor sequence learning. Retrieved from osf.io/etnqc

Berlot, E., Popp, N. J., & Diedrichsen, J. (2018). In search of the engram, 2017. *Current Opinion in Behavioral Sciences*, *20*. http://doi.org/10.1016/j.cobeha.2017.11.003

Beukema, P., Diedrichsen, J., & Verstynen, T. D. (2019). Binding During Sequence Learning Does Not Alter Cortical Representations of Individual Actions. *The Journal of Neuroscience*, *39*(35), 6968–6977. http://doi.org/10.1523/jneurosci.2669-18.2019

Classen, J., Liepert, J., Wise, S. P., Hallett, M., & Cohen, L. G. (1998). Rapid plasticity of human cortical movement representation induced by practice. *Journal of Neurophysiology*, *79*(2), 1117–1123. http://doi.org/10.1152/jn.1998.79.2.1117

Coynel, D., Marrelec, G., Perlbarg, V., Pélégrini-Issac, M., Van de Moortele, P. F., Ugurbil, K., … Lehéricy, S. (2010). Dynamics of motor-related functional integration during motor sequence learning. *NeuroImage*, *49*(1), 759–766. http://doi.org/10.1016/j.neuroimage.2009.08.048

Dale, A. M., Fischl, B., & Sereno, M. I. (1999). Cortical Surface-Based Analysis. *NeuroImage*, *9*(2), 179–194. http://doi.org/10.1006/nimg.1998.0395

Dayan, E., & Cohen, L. G. (2011). Neuroplasticity subserving motor skill learning. *Neuron*, *72*(3), 443–454. http://doi.org/10.1016/j.neuron.2011.10.008

Diedrichsen, J., Yokoi, A., & Arbuckle, S. A. (2017). Pattern component modeling: A flexible approach for understanding the representational structure of brain activity patterns. *NeuroImage*. http://doi.org/10.1016/j.neuroimage.2017.08.051

Doyon, J., Song, A. W., Karni, A., Lalonde, F., Adams, M. M., & Ungerleider, L. G. (2002). Experience-dependent changes in cerebellar contributions to motor sequence learning. *Proceedings of the National Academy of Sciences of the United States of America*, *99*(2), 1017–22. http://doi.org/10.1073/pnas.022615199

Fischl, B., Rajendran, N., Busa, E., Augustinack, J., Hinds, O., Yeo, B. T. T., … Zilles, K. (2008). Cortical folding patterns and predicting cytoarchitecture. *Cerebral Cortex*, *18*(8), 1973–1980. http://doi.org/10.1093/cercor/bhm225

Floyer-Lea, A., & Matthews, P. M. (2005). Distinguishable brain activation networks for short- and long-term motor skill learning. *Journal of Neurophysiology*, *94*(1), 512–518. http://doi.org/10.1152/jn.00717.2004

Friston, K. J., Worsley, K. J., Poline, J.-P. B., Frith, C. D., Frackowiak, R. S. J., Holmes, a. P., … Frackowiak, R. S. J. (1994). Statistical parametric maps in functional imaging: A general linear approach. *Human Brain Mapping*, *2*(4), 189–210. http://doi.org/10.1002/hbm.460020402

Galea, J. M., Vazquez, A., Pasricha, N., Orban De Xivry, J. J., & Celnik, P. (2011). Dissociating

875   the roles of the cerebellum and motor cortex during adaptive learning: The motor cortex
876   retains what the cerebellum learns. *Cerebral Cortex*, *21*(8), 1761–1770.
877   http://doi.org/10.1093/cercor/bhq246
878 Grafton, S. T., Hazeltine, E., & Ivry, R. (1995). Functional mapping of sequence learning in
879   normal humans. *Journal of Cognitive Neuroscience*, *7*(4), 497–510.
880   http://doi.org/10.1162/jocn.1995.7.4.497
881 Grafton, S. T., Hazeltine, E., & Ivry, R. B. (2002). Motor sequence learning with the
882   nondominant left hand: A PET functional imaging study. *Experimental Brain Research*,
883   *146*(3), 369–378. http://doi.org/10.1007/s00221-002-1181-y
884 Graybiel, A. M. (2016). The Basal Ganglia. *Principles of Neural Science*, 982–998.
885   http://doi.org/10.1007/978-3-319-42743-0
886 Graybiel, A. M., & Grafton, S. T. (2015). The striatum: Where skills and habits meet. *Cold Spring
887   Harbor Perspectives in Biology*, *7*(8), 1–14. http://doi.org/10.1101/cshperspect.a021691
888 Gusnard, D. A., Akbudak, E., Shulman, G. L., & Raichle, M. E. (2001). Medial prefrontal cortex
889   and self-referential mental activity: Relation to a default mode of brain function.
890   *Proceedings of the National Academy of Sciences of the United States of America*, *98*(7),
891   4259–4264. http://doi.org/10.1073/pnas.071043098
892 Hadipour-Niktarash, A., Lee, C. K., Desmond, J. E., & Shadmehr, R. (2007). Impairment of
893   retention but not acquisition of a visuomotor skill through time-dependent disruption of
894   primary motor cortex. *Journal of Neuroscience*, *27*(49), 13413–13419.
895   http://doi.org/10.1523/JNEUROSCI.2570-07.2007
896 Hardwick, R. M., Rottschy, C., Miall, R. C., & Eickhoff, S. B. (2013). A quantitative meta-analysis
897   and review of motor learning in the human brain. *NeuroImage*, *67*, 283–297.
898   http://doi.org/10.1016/j.neuroimage.2012.11.020
899 Hazeltine, E., Grafton, S. T., & Ivry, R. (1997). Attention and stimulus characteristics determine
900   the locus of motor-sequence encoding. A PET study. *Brain*, *120*(1), 123–140.
901   http://doi.org/10.1093/brain/120.1.123
902 Hikosaka, O., Sakai, K., Lu, X., Nakahara, H., Rand, M. K., Nakamura, K., … Doya, K. (1999).
903   Parallel neural networks for learning sequential procedures. *Trends in Neurosciences*,
904   *22*(10), 464–471. http://doi.org/10.1016/S0166-2236(99)01439-3
905 Honda, M., Deiber, M. P., Ibáñez, V., Pascual-Leone, A., Zhuang, P., & Hallett, M. (1998).
906   Dynamic cortical involvement in implicit and explicit motor sequence learning. A PET study.
907   *Brain*, *121*(11), 2159–2173. http://doi.org/10.1093/brain/121.11.2159
908 Huang, Y., Zhen, Z., Song, Y., Zhu, Q., Wang, S., & Liu, J. (2013). Motor Training Increases the
909   Stability of Activation Patterns in the Primary Motor Cortex. *PLoS ONE*, *8*(1).
910   http://doi.org/10.1371/journal.pone.0053555
911 Hutton, C., Bork, A., Josephs, O., Deichmann, R., Ashburner, J., & Turner, R. (2002). Image
912   distortion correction in fMRI: A quantitative evaluation. *NeuroImage*, *16*(1), 217–240.
913   http://doi.org/10.1006/nimg.2001.1054
914 Jenkins, I. H., Brooks, D. J., Nixon, P. D., Frackowiak, R. S. J., & Passingham, R. E. (1994).
915   Motor sequence learning: A study with positron emission tomography. *Journal of
916   Neuroscience*, *14*(6), 3775–3790. http://doi.org/10.1523/jneurosci.14-06-03775.1994
917 Kang, E. K., & Paik, N. J. (2011). Effect of a tDCS electrode montage on implicit motor
918   sequence learning in healthy subjects. *Experimental and Translational Stroke Medicine*,
919   *3*(1), 2–7. http://doi.org/10.1186/2040-7378-3-4
920 Karni, A., Meyer, G., Jezzard, P., Adams, M. M., Turner, R., & Ungerleider, L. G. (1995).
921   Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature*,
922   *377*(6545), 155–8. http://doi.org/10.1038/377155a0
923 Karni, A., Meyer, G., Rey-Hipolito, C., Jezzard, P., Adams, M. M., Turner, R., & Ungerleider, L.
924   G. (1998). The acquisition of skilled motor performance: fast and slow experience-driven
925   changes in primary motor cortex. *Proceedings of the National Academy of Sciences of the*

34

926   *United States of America*, *95*(3), 861–8. http://doi.org/10.1073/pnas.95.3.861

927 Kawai, R., Markman, T., Poddar, R., Ko, R., Fantana, A. L., Dhawale, A. K., … Ölveczky, B. P.
928   (2015). Motor Cortex Is Required for Learning but Not for Executing a Motor Skill. *Neuron*,
929   *86*(3), 800–812. http://doi.org/10.1016/j.neuron.2015.03.024

930 Kumar, N., Manning, T. F., & Ostry, D. J. (2019). Somatosensory cortex participates in the
931   consolidation of human motor memory. *PLoS Biology*, *17*(10), e3000469.
932   http://doi.org/10.1371/journal.pbio.3000469

933 Lashley, K. (1950). In search of the engram.

934 Lehéricy, S., Bardinet, E., Tremblay, L., Van de Moortele, P. F., Pochon, J. B., Dormont, D., …
935   Ugurbil, K. (2006). Motor control in basal ganglia circuits using fMRI and brain atlas
936   approaches. *Cerebral Cortex (New York, N.Y. : 1991)*, *16*(2), 149–61.
937   http://doi.org/10.1093/cercor/bhi089

938 Lehéricy, S., Benali, H., Van de Moortele, P.-F., Pélégrini-Issac, M., Waechter, T., Ugurbil, K., &
939   Doyon, J. (2005). Distinct basal ganglia territories are engaged in early and advanced
940   motor sequence learning. *Proceedings of the National Academy of Sciences of the United*
941   *States of America*, *102*(35), 12566–12571. http://doi.org/10.1073/pnas.0502762102

942 Lutz, K., Koeneke, S., Wüstenberg, T., & Jäncke, L. (2004). Asymmetry of cortical activation
943   during maximum and convenient tapping speed. *Neuroscience Letters*, *373*(1), 61–66.
944   http://doi.org/10.1016/j.neulet.2004.09.058

945 Ma, L., Wang, B., Narayana, S., Hazeltine, E., Chen, X., Robin, D. A., … Xiong, J. (2010).
946   Changes in regional activity are accompanied with changes in inter-regional connectivity
947   during 4 weeks motor learning. *Brain Research*, *1318*, 64–76.
948   http://doi.org/10.1016/j.brainres.2009.12.073

949 Makino, H., Ren, C., Liu, H., Kim, A. N., Kondapaneni, N., Liu, X., … Komiyama, T. (2017).
950   Transformation of Cortex-wide Emergent Properties during Motor Learning. *Neuron*, *94*(4),
951   880–890.e8. http://doi.org/10.1016/j.neuron.2017.04.015

952 Matsuzaka, Y., Picard, N., & Strick, P. L. (2007). Skill representation in the primary motor cortex
953   after long-term practice. *Journal of Neurophysiology*, *97*(2), 1819–1832.
954   http://doi.org/10.1152/jn.00784.2006

955 Mattar, M., Wymbs, N. F., Bock, A. S., Aguirre, G. K., Grafton, S. T., & Bassett, D. S. (2016).
956   Predicting future learning from baseline network architecture. *bioRxiv*, *172*(November
957   2017), 56861. http://doi.org/10.1101/056861

958 Nitsche, M. A., Schauenburg, A., Lang, N., Liebetanz, D., Exner, C., Paulus, W., & Tergau, F.
959   (2003). Facilitation of implicit motor learning by weak transcranial direct current stimulation
960   of the primary motor cortex in the human. *Journal of Cognitive Neuroscience*, *15*(4), 619–
961   626. http://doi.org/10.1162/089892903321662994

962 Oosterhof, N. N., Wiestler, T., Downing, P. E., & Diedrichsen, J. (2011). A comparison of
963   volume-based and surface-based multi-voxel pattern analysis. *NeuroImage*, *56*(2), 593–
964   600. http://doi.org/10.1016/j.neuroimage.2010.04.270

965 Orban, P., Peigneux, P., Lungu, O., Albouy, G., Breton, E., Laberenne, F., … Doyon, J. (2010).
966   The multifaceted nature of the relationship between performance and brain activity in motor
967   sequence learning. *NeuroImage*, *49*(1), 694–702.
968   http://doi.org/10.1016/j.neuroimage.2009.08.055

969 Ostry, D. J., & Gribble, P. L. (2016). Sensory Plasticity in Human Motor Learning. *Trends in*
970   *Neurosciences*, *39*(2), 114–123. http://doi.org/10.1016/j.tins.2015.12.006

971 Otchy, T. M., Wolff, S. B. E., Rhee, J. Y., Pehlevan, C., Kawai, R., Kempf, A., … Ölveczky, B. P.
972   (2015). Acute off-target effects of neural circuit manipulations. *Nature*, *528*(7582), 358–
973   363. http://doi.org/10.1038/nature16442

974 Penhune, V. B., & Doyon, J. (2002). Dynamic cortical and subcortical networks in learning and
975   delayed recall of timed motor sequences. *Journal of Neuroscience*, *22*(4), 1397–406.
976   http://doi.org/22/4/1397 [pii]

Penhune, V. B., & Steele, C. J. (2012). Parallel contributions of cerebellar, striatal and M1 mechanisms to motor sequence learning. *Behavioural Brain Research*, *226*(2), 579–591. http://doi.org/10.1016/j.bbr.2011.09.044

Peters, A. J., Lee, J., Hedrick, N. G., Neil, K. O., & Komiyama, T. (2017). Reorganization of corticospinal output during motor learning. *Nature Publishing Group*, (July). http://doi.org/10.1038/nn.4596

Picard, N., Matsuzaka, Y., & Strick, P. L. (2013). Extended practice of a motor skill is associated with reduced metabolic activity in M1. *Nature Neuroscience*, *16*(9), 1340–7. http://doi.org/10.1038/nn.3477

Pinsard, B., Boutin, A., Gabitov, E., Lungu, O., Benali, H., & Doyon, J. (2018). Consolidation alters motor sequence-specific distributed representations. *bioRxiv*, 376053. http://doi.org/10.1101/376053

Poldrack, R. A. (2000). Imaging brain plasticity: Conceptual and methodological issues - A theoretical review. *NeuroImage*, *12*(1), 1–13. http://doi.org/10.1006/nimg.2000.0596

Poldrack, R. A., Sabb, F. W., Foerde, K., Tom, S. M., Asarnow, R. F., Bookheimer, S. Y., & Knowlton, B. J. (2005). The neural correlates of motor skill automaticity. *Journal of Neuroscience*, *25*(22), 5356–5364. http://doi.org/10.1523/JNEUROSCI.3880-04.2005

Raichle, M. E., MacLeod, A. M., Snyder, A. Z., Powers, W. J., Gusnard, D. A., & Shulman, G. L. (2001). A default mode of brain function. *Proceedings of the National Academy of Sciences of the United States of America*, *98*(2), 676–682. http://doi.org/10.1073/pnas.98.2.676

Reis, J., Schambra, H. M., Cohen, L. G., Buch, E. R., Fritsch, B., Zarahn, E., … Krakauer, J. W. (2009). Noninvasive cortical stimulation enhances motor skill acquisition over multiple days through an effect on consolidation. *Proceedings of the National Academy of Sciences of the United States of America*, *106*(5), 1590–1595. http://doi.org/10.1073/pnas.0805413106

Reithler, J., van Mier, H. I., & Goebel, R. (2010). Continuous motor sequence learning: Cortical efficiency gains accompanied by striatal functional reorganization. *NeuroImage*, *52*(1), 263–276. http://doi.org/10.1016/j.neuroimage.2010.03.073

Russo, A. A., Bittner, S. R., Perkins, S. M., Seely, J. S., London, B. M., Lara, A. H., ... & Abbott, L. F. (2018). Motor cortex embeds muscle-like commands in an untangled population response. *Neuron*, *97*(4), 953-966.

Russo, A. A., Khajeh, R., Bittner, S. R., Perkins, S. M., Cunningham, J. P., Abbott, L. F., & Churchland, M. M. (2019). Neural trajectories in the supplementary motor area and primary motor cortex exhibit distinct geometries, compatible with different classes of computation. *bioRxiv*, 650002.

Seidler, R. D., Purushotham, A., Kim, S. G., Ugurbil, K., Willingham, D., & Ashe, J. (2005). Neural correlates of encoding and expression in implicit sequence learning. *Experimental Brain Research*, *165*(1), 114–124. http://doi.org/10.1007/s00221-005-2284-z

Seidler, R. D., Purushotham, A., Kim, S. G., Uğurbil, K., Willingham, D., & Ashe, J. (2002). Cerebellum activation associated with performance change but not motor learning. *Science*, *296*(5575), 2043–2046. http://doi.org/10.1126/science.1068524

Shamloo, F., & Helie, S. (2016). Changes in default mode network as automaticity develops in a categorization task. *Behavioural Brain Research*, *313*, 324–333. http://doi.org/10.1016/j.bbr.2016.07.029

Shulman, G. L., Fiez, J. A., Corbetta, M., Buckner, R. L., Miezin, F. M., Raichle, M. E., & Petersen, S. E. (1997). Common blood flow changes across visual tasks: II. Decreases in cerebral cortex. *Journal of Cognitive Neuroscience*, *9*(5), 648–663. http://doi.org/10.1162/jocn.1997.9.5.648

Steele, C. J., & Penhune, V. B. (2010). Specific Increases within Global Decreases : A Functional Magnetic Resonance Imaging Investigation of Five Days of Motor Sequence Learning, *30*(24), 8332–8341. http://doi.org/10.1523/JNEUROSCI.5569-09.2010

Tamás Kincses, Z., Johansen-Berg, H., Tomassini, V., Bosnell, R., Matthews, P. M., &

1028   Beckmann, C. F. (2008). Model-free characterization of brain functional networks for motor
1029   sequence learning using fMRI. *NeuroImage*, *39*(4), 1950–1958.
1030   http://doi.org/10.1016/j.neuroimage.2007.09.070
1031 Toni, I., Krams, M., Turner, R., & Passingham, R. E. (1998). The time course of changes during
1032   motor sequence learning: A whole- brain fMRI study. *NeuroImage*, *8*(1), 50–61.
1033   http://doi.org/10.1006/nimg.1998.0349
1034 Ungerleider, L. G., Doyon, J., & Karni, A. (2002). Imaging brain plasticity during motor skill
1035   learning. *Neurobiology of Learning and Memory*, *78*(3), 553–564.
1036   http://doi.org/10.1006/nlme.2002.4091
1037 Vahdat, S., Lungu, O., Cohen-Adad, J., Marchand-Pauvert, V., Benali, H., & Doyon, J. (2015).
1038   Simultaneous Brain-Cervical Cord fMRI Reveals Intrinsic Spinal Cord Plasticity during
1039   Motor Sequence Learning. *PLoS Biology*, *13*(6), e1002186.
1040   http://doi.org/10.1371/journal.pbio.1002186
1041 Walther, A., Nili, H., Ejaz, N., Alink, A., Kriegeskorte, N., & Diedrichsen, J. (2016). Reliability of
1042   dissimilarity measures for multi-voxel pattern analysis. *NeuroImage*, *137*, 188–200.
1043   http://doi.org/10.1016/j.neuroimage.2015.12.012
1044 Waters-metenier, S., Husain, M., & Wiestler, T. (2014). Bihemispheric Transcranial Direct
1045   Current Stimulation Enhances Effector-Independent Representations of Motor Synergy and
1046   Sequence Learning, *34*(3), 1037–1050. http://doi.org/10.1523/JNEUROSCI.2282-13.2014
1047 Wiestler, T., & Diedrichsen, J. (2013). Skill learning strengthens cortical representations of
1048   motor sequences. *eLife*, *2013*(2), 1–20. http://doi.org/10.7554/eLife.00801
1049 Wu, T., Kansaku, K., & Hallett, M. (2004). How Self-Initiated Memorized Movements Become
1050   Automatic: A Functional MRI Study. *Journal of Neurophysiology*, *91*(4), 1690–1698.
1051   http://doi.org/10.1152/jn.01052.2003
1052 Wymbs, N. F., & Grafton, S. T. (2015). The human motor system supports sequence-specific
1053   representations over multiple training-dependent timescales. *Cerebral Cortex*, *25*(11),
1054   4213–4225. http://doi.org/10.1093/cercor/bhu144
1055 Xiong, J., Ma, L., Wang, B., Narayana, S., Duff, E. P., Egan, G. F., & Fox, P. T. (2009). Long-
1056   term motor training induced changes in regional cerebral blood flow in both task and
1057   resting states. *NeuroImage*, *45*(1), 75–82. http://doi.org/10.1016/j.neuroimage.2008.11.016
1058 Yin, H. H., Mulcare, S. P., Hilário, M. R. F., Clouse, E., Holloway, T., Davis, M. I., … Costa, R.
1059   M. (2009). Dynamic reorganization of striatal circuits during the acquisition and
1060   consolidation of a skill. *Nature Neuroscience*, *12*(3), 333–341.
1061   http://doi.org/10.1038/nn.2261
1062 Yokoi, A., Arbuckle, S. A., & Diedrichsen, J. (2018). The role of human primary motor cortex in
1063   the production of skilled finger sequences. *Journal of Neuroscience*, *38*(6), 1430–1442.
1064   http://doi.org/10.1523/JNEUROSCI.2798-17.2017
1065 Yokoi, A., & Diedrichsen, J. (2019). Neural Organization of Hierarchical Motor Sequence
1066   Representations in the Human Neocortex. *Neuron*, *103*(6), 1178–1190.e7.
1067   http://doi.org/10.1016/j.neuron.2019.06.017
1068 Yousry, T. A., Schmid, U. D., Alkadhi, H., Schmidt, D., Peraud, A., Buettner, A., & Winkler, P.
1069   (1997). Localization of the motor hand area to a knob on the precentral gyrus. A new
1070   landmark. *Brain*, *120*(1), 141–157. http://doi.org/10.1093/brain/120.1.141
1071
1072