The effects of habits on motor skill learning

Nicola J. Popp[1*], Atsushi Yokoi[2,3], Paul Gribble[1,4,5] & Jörn Diedrichsen[1,6,7]

[1]The Brain and Mind Institute, University of Western Ontario, Canada,

[2]Graduate School of Frontier Biosciences, Osaka University, Japan,

[3]Center for Information and Neural Networks (CiNet), NICT, Japan,

[4]Department of Physiology & Pharmacology, University of Western Ontario, Canada,

[5]Haskins Laboratories, USA,

[6]Department of Statistical and Actuarial Sciences, University of Western Ontario, Canada,

[7]Department of Computer Science, University of Western Ontario, Canada

Acknowledgments

Correspondence address

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

Author contributions

N.J.P. and J.D. designed the experiment. N.J.P, A.Y. and J.D. programmed the experiment. N.J.P collected the data and analyzed the data together with J.D.. N.J.P., J.D., and P.G. wrote the manuscript.

Conflict of interest

The authors declare no conflict of interest.

1 **Abstract**

2 Skill learning involves the formation of stable motor patterns. In musical and athletic

3 training, however, the formation of bad habits can impede the attainment of high levels of

4 performance. We developed a novel experimental paradigm to induce a specific motor

5 habit in a sequence production task and investigated how it affected subsequent

6 optimization over a 3-week training period. Participants initially practiced small segments

7 of 2 to 3 finger movements, which were then combined to form longer sequences. This

8 initial training induced a persistent chunking behavior, with shorter intervals between

9 presses within a chunk and longer presses at chunk boundaries. By aligning these chunks

10 with biomechanically easy or difficult transitions, we were able to induce chunking habits

11 that were either beneficial or detrimental to performance. We found that participants

12 overcame detrimental habits using two kinds of optimization processes. Failure to change

13 detrimental chunk structures predicted lower levels of final performance.

14 **Introduction**

15 Humans are capable of developing astonishing feats of motor skill, such as performing

16 athletics, playing music and dancing. But what does it take to become a motor expert?

17 The first obvious factor is practice: it is often estimated that 10,000 hours of training are

18 necessary to develop high-level motor skills[1,2]. Perhaps the motor system simply needs

19 to acquire a large amount of information over a long period of time to optimize the

20 muscle commands that are necessary to achieve skilled movement. According to this

21 view, given enough practice, motor expertise should automatically emerge[1,2].

22 Simply practicing for many hours will not automatically lead to expert

23 performance, however. There are numerous examples in which motor skill acquisition is

24 slow or fails altogether[3]. This is commonly attributed to the formation of habits[4,5]:

25 automatic and highly entrenched behavioral patterns that resist change through

26 retraining[4–14]. While the automatic nature of habits can be beneficial to performance by

27 reducing cognitive load[3,15,16], the rigid nature of habits can potentially be detrimental to

28 improvement[4–6]. Despite being a highly relevant phenomenon across many domains,

29 the influence of motor habits on skill learning has not been investigated in a controlled

30   laboratory study. How long does it take for a motor habit to become ingrained? How

31   stable are motor habits over time? Can a bad habit be overcome with practice?

32   To investigate habit formation in motor skill learning, we introduce a novel

33   experimental paradigm that enables us to induce specific motor habits and to test

34   whether and how participants can overcome these habits with subsequent practice. As

35   an experimental model of motor skill acquisition, we used the discrete sequence

36   production task (DSP), in which participants perform a series of single finger presses as

37   fast as possible while having full knowledge of the sequence[17,18]. Learning in this task

38   depends on both cognitive and motor processes[19,20]. Initial performance relies strongly

39   on forming a declarative memory of the sequence and can, therefore, be sculpted

40   through explicit instructions[21–24]. Later stages of learning increasingly rely on the

41   optimization of execution-related processes that are involved in producing the fast and

42   co-articulated finger presses. As such, the DSP task provides a model of the rich

43   interplay between cognitive and motor processes involved in skill acquisition.

44   An important cognitive influence on sequence performance is "chunking" – the

45   process by which participants separate a long sequence into smaller subsets[22,25].

46   Chunking has been shown to aid memorization and performance by reducing memory

47   capacity demands [25–29]. The structure of the declarative memory representation of a

48   sequence may then constrain subsequent motor optimization processes[30,31]; it has been

49   suggested that sequential movements may be optimized within a chunk, but not across

50   chunk boundaries[32]. Thus, we hypothesized that depending on the way a long

51   sequence is cognitively chunked, the resulting habit could either facilitate or impede

52   subsequent skill learning. Hence, we use chunking as a tool to impose habits onto

53   participants' behavior and investigate how these subsequently change with practice.

54   We trained 3 groups of participants to perform the same set of 7 sequences,

55   each consisting of 11 isometric keypresses. Training occurred on 15 separate days,

56   spread over 3 weeks. The experiment was designed to induce a specific chunk

57   structure at the beginning of training, and then test how this chunk structure affected

58   subsequent improvement through practice (optimization). In the induction phase (Fig.

59   1a), participants (groups 1 & 2) first practiced small 2-3 digit chunks and then learned

3

60   the sequences, each of which consisted of four of the pre-trained chunks. The two

61   groups differed in how the sequences were broken up into chunks (Fig. 1c & suppl.

62   Table 2). We imposed chunk structures that were designed to be either aligned or

63   misaligned with biomechanically easy or difficult transitions within the sequence and

64   therefore were predicted to be beneficial or detrimental to subsequent movement

65   speed. To test for patterns of spontaneous chunking, the third (control) group was

66   trained on the same sequences but we did not impose a specific chunk structure during

67   the initial part of training.

68   In the subsequent optimization phase (Fig. 1a), participants were instructed to

69   improve their performance through practice. During this phase, we did not make any

70   further mention of chunks (see Methods). Because the sequences were identical for all

71   groups, we could make strong inferences about whether performance during this

72   subsequent training was dictated by biomechanical demands (which were identical

73   across groups) or whether optimization was affected by the chunk structure imposed

74   during the induction phase. Furthermore, using a Bayesian model, we estimated

75   changes in chunk structure that were independent of performance changes. This

76   allowed us to investigate three questions: First, can explicit instructions at the beginning

77   of training lead to stable motor habits? Second, do particular ways of chunking impede

78   or enhance participants' ability to reach skilled performance? Finally, what are the

79   optimization processes that allow participants to overcome "bad" habits through

80   practice?

Results

82   Over 15 days we trained 32 participants to produce sequences of 11 isometric

83   keypresses from memory on a keyboard-like device. Participants were rewarded with

84   pleasant sounds and points for executing the sequences as fast as possible while

85   keeping the proportion of incorrect key presses in each block of trials below 15%.

86   We manipulated how participants memorized the sequences by splitting the

87   sequences up into several chunks, each composed of 2-3 keypresses. The idea was to

88   test whether the different ways of chunking (hereafter "chunk structures") imposed in the

89   induction phase (Fig. 1a) would affect how participants were able to optimize their

4

90   performance in the subsequent two weeks of training. On the first day of the

91   experiment, we trained participants on eleven different 2-3 press chunks. Participants

92   learned to produce each chunk in response to a visually presented letter. For example,

93   the letter A corresponded to the chunk "3,2,1" (middle finger, index finger, thumb). At

94   the end of the first day, participants could reliably produce the chunks from memory with

95   an average accuracy of 92.7%. On days 2-4, we combined these smaller chunks to

96   form 7 different 11-press sequences. Each sequence was also associated with a visual

97   cue in the form of a single character symbol (e.g. $ or %). At the end of day 4,

98   participants were able to recall all sequences from memory using the sequence cue with

99   an accuracy of 92.5%. In the optimization phase (days 5-14), participants then tried to

100   improve their performance by decreasing their movement time (MT - time elapsed from

101   first to last press of the sequence). In this phase, no further instructions in terms of

102   chunks were given.

103   All participants trained on the same 7 sequences, but we imposed different chunk

104   structures, depending on the group that the participants were assigned to. This allowed

105   us to dissociate the effects of the initial instruction from chunking that emerges from the

106   biomechanical requirements of the sequence. To compare the formation of "good" and

107   "bad" habits, we designed chunk structures that were predicted to be either beneficial or

108   detrimental to performance. These predictions were based on a separate experiment

109   (see Methods), in which we trained 7 participants on all possible 2 and 3 keypress

110   combinations over the course of 3 days and measured how fast they could execute

111   these short sequences. The results showed that the transition between two adjacent

112   fingers can be performed faster than two repeated presses of the same finger ($t_{(6)} =$

113   13.965, $p = 8.404e\text{-}06$; see Fig. 1b). Given that the 2-3 press sequences hardly taxed

114   the cognitive system, these results can be taken as a characterization of the

115   biomechanical constraints of our task.

116   We used these results to design two different ways of separating the sequences

117   into chunks. In one case, chunk boundaries were **aligned** with digit transitions that were

118   performed more slowly (as measured in the independent dataset) – i.e. they were

119   preferentially placed between digit repetitions (Fig. 1c). The time that is required to

120   perform these difficult finger transition can therefore simultaneously be used to recall a

121   next chunk, which should benefit overall performance. In the **misaligned** chunk

122   structure, chunk boundaries were placed at digit transitions that can be performed

123   quickly, thereby breaking up transitions between adjacent fingers or runs (Fig. 1c).

124   Participants would, therefore, have to slow down their performance at these fast

125   transitions to recall the next chunk, which we predicted should hurt their overall speed.

126   Each participant learned 3 of the 7 sequences with a misaligned chunk structure and 3

127   sequences with an aligned chunk structure, with the assignment counterbalanced

128   across groups. For the last remaining sequence, both ways of chunking were predicted

129   to be equally fast, as both possible chunk structures were relatively well aligned with the

130   biomechanical requirements ("neutral chunk structure", Fig. 1d & Supp. Table 1).
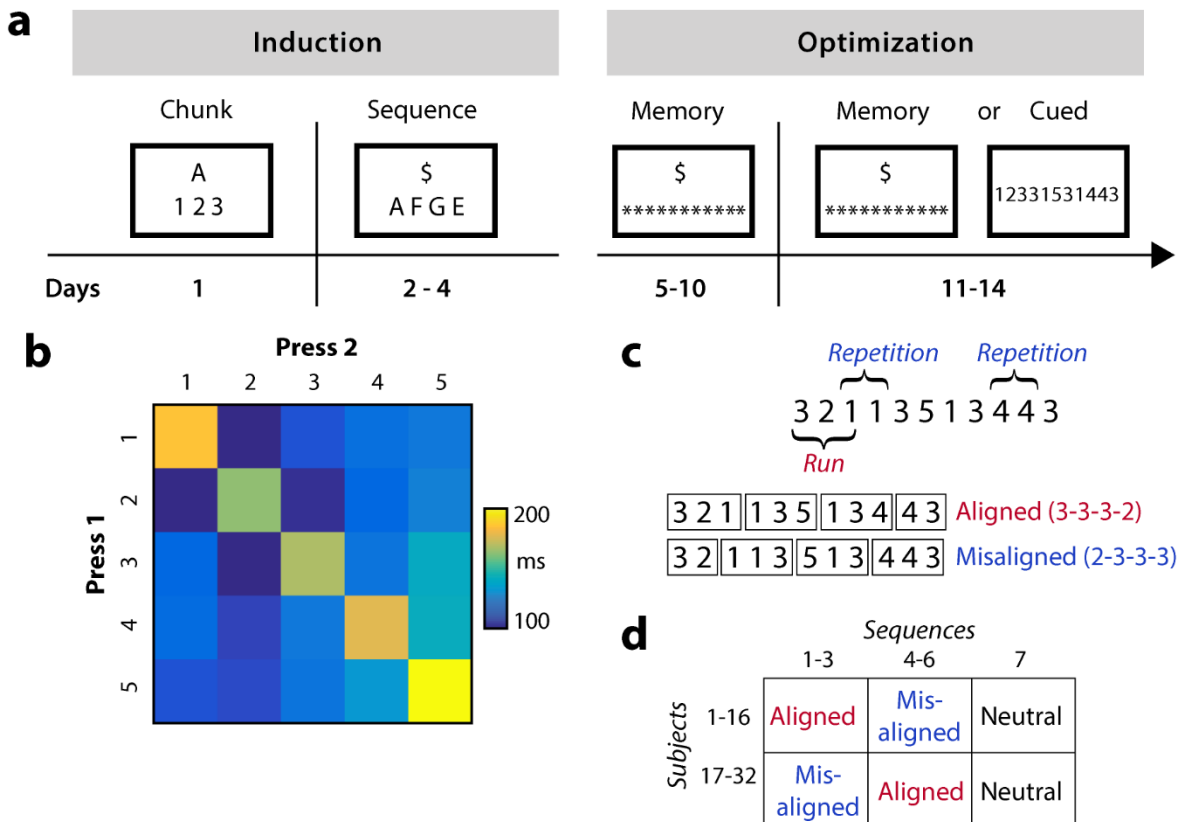
Figure 1. Experimental procedure. (*a*) *Experimental timeline depicting the training at each stage. In the induction phase participants memorized chunks and sequences. In the optimization phase participants trained to perform these sequences as fast as possible from memory. In the last week of training, half of the participants were directly cued with the sequence, while the other half performed the sequences from memory. (**b**) Data from an independent dataset, in which participants performed all possible combinations of 2 and 3-digit transitions. Matrix indicates the median inter-press interval (IPI) to produce the transition between pairs of keypresses. (**c**) Top: Example sequence containing a 3-digit run and two digit repetitions. Bottom: The sequence was separated using two possible chunk structures. In the aligned structure, the chunk boundaries fell between repetitions, in the misaligned structure the chunk boundary broke up the run. (**d**) The assignment of chunk structures to sequences was counterbalanced between participants.*

7

Chunk induction induces stable motor habits

131 To assess whether the experimentally imposed chunk structure influenced participants'

132 motor behavior, we examined inter-press time intervals (IPIs). An increased IPI is

133 commonly taken as a sign of the beginning of a new chunk, as the cognitive processes

134 (memory recall, action selection) involved in switching from one chunk to another

135 require additional time[24,33]. Hence, we would expect our participants to exhibit shorter

136 IPIs between keypresses that belonged to one of the chunks imposed during day 1

137 (within-chunk IPIs) and larger IPIs for the boundaries between chunks (between-chunk

138 IPIs). Consistent with this idea, we found significantly longer between-chunk IPIs

139 compared to within-chunk IPIs in the first four days of training (Fig. 2a: days 2-4: $t_{(31)} =$

140 7.728, p = 5.098e-09).

142 After day 4, we ceased to cue sequences using the alphabetic letters associated

143 with the chunks. Instead, participants had to recall the sequences from memory using

144 only the symbolic sequence cues (e.g. "$"). From this point forward, no further reference

145 was made to the chunk structure imposed in the induction phase. Across days 5-10, the

146 difference between the within- and between-chunk IPIs remained stable; $t_{(15)}$ = 5.985, p

147 = 2.351e-08 (Fig. 2a). Importantly, this difference cannot be attributed the

148 biomechanical difficulty of the finger transitions. The within-chunk IPIs for one group

149 were the between-chunk IPIs for the other group and vice versa; IPIs that were within-

150 chunk for all participants (e.g. the first and last IPI of a sequence) were excluded from

151 this analysis. In summary, even though after day 4, sequences were cued only using a

152 single symbol, participants persisted in performing the sequences in a manner

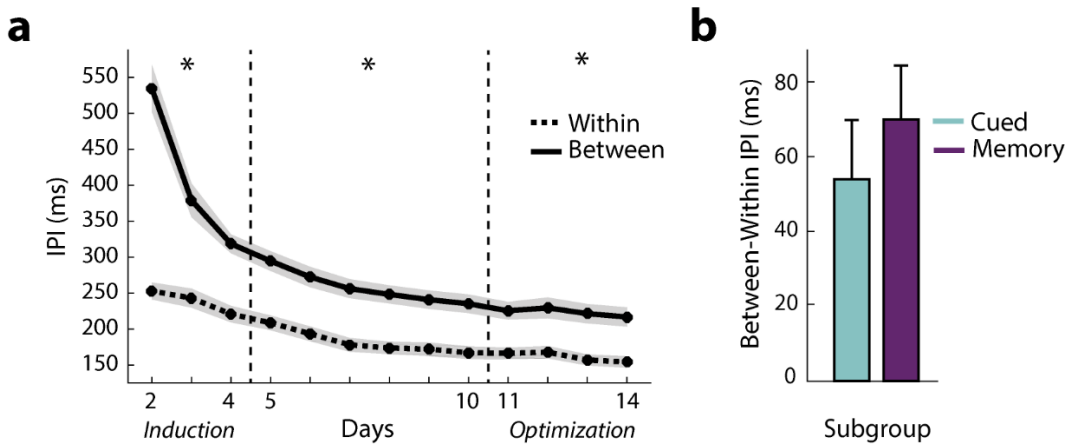153 consistent with the chunk structures imposed early in training.

Figure 2. Within vs. between inter-press intervals (IPIs). **(a)** A stable difference between within and between IPIs can be observed over the course of the experiment. Asterisks indicate significant differences in the corresponding week (separated by dashed lines). **(b)** Average between – within IPIs in the last week of training, split by whether participants had to recall the sequences from memory or were cued with the sequence numbers. Error bars denote between-subject standard error.

154   In the last four days of training (days 11-14) we tested whether the persistence of
155   the instructed chunk structure reflected a motor habit or whether it reflected memory
156   recall. We split each experimental group into two subgroups: half of the participants
157   continued to perform the sequences from memory, while the other half were cued using
158   the numbers (Fig. 1a) that indicated the necessary keypresses, therefore removing any
159   memory demands. Both the memory ($t_{(15)}$ = 4.8651, p = 2.0591e-Fig. 2b) and the cued
160   subgroup ($t_{(15)}$ = 3.403 , p = 0.004) showed a significant difference between the within-
161   and between-chunk IPIs and there was no reliable difference between the two
162   subgroups in this effect ($t_{(30)}$ = -0.749, p = 0.460). Thus, removing the requirement for
163   memory recall in the last four days of training did not abolish chunking. Because none of
164   the subsequent analyses showed any significant difference between the two subgroups,
165   we will report their combined results for the remainder of the article. Overall, these
166   results suggest the explicit chunk training early in learning established a stable motor
167   habit that outlasted 10 days of subsequent practice.

Changes in misaligned vs. aligned chunk structure

Next, we examined whether participants retained the chunk structure to the same degree for aligned and misaligned chunking patterns. We split the IPI data by whether the intervals were taken from the sequences that were instructed with the aligned or misaligned chunk structure. While the difference between the within- and between-chunk IPIs for the instructed aligned chunk structure was stable over the entire 3 weeks of training, the difference for the instructed misaligned structure disappeared in the last four days of training (Fig. 3a). The three-way interaction between day x within/between x instruction (aligned or misaligned) was significant ($F_{(24,744)} = 12.493$, $p < 1e-16$). Thus, in the last four days of training participants seemed to diverge from the misaligned chunk structure while consistently keeping the aligned chunk structure.

Influence of chunking instruction on performance

We predicted that chunking instructions that are misaligned with biomechanical constraints would hinder performance. Indeed, in the first two weeks of training (days 2-10) the sequences that were instructed with the misaligned chunk structure were performed slower than the sequences instructed with the aligned chunk structure (one-sample t-test of differences in performance: days 2-4: $t_{(31)} = 2.693$, $p = 0.006$, days 5-10: $t_{(31)} = 2.313$, $p = 0.014$; Fig. 3b). Hence, we were able to induce differences in performance, purely based on our training instructions. This is an important finding because it shows that we were not only able to manipulate *how* participants performed a sequence but also *how well* they could perform it.

Interestingly, the difference in performance between the sequences that were instructed with the aligned compared to the misaligned chunk structure was no longer significant in the last four days of training (days 10-14: $t_{(31)} = 0.764$, $p = 0.225$; Fig. 3b). This coincided with the fact that in the last four days, participants no longer showed a systematic difference between the within- and between-chunk IPIs for the sequences that were instructed with the misaligned chunk structure. This suggests that participants were able to overcome the induced performance deficit by abandoning the non-optimal chunk structure.
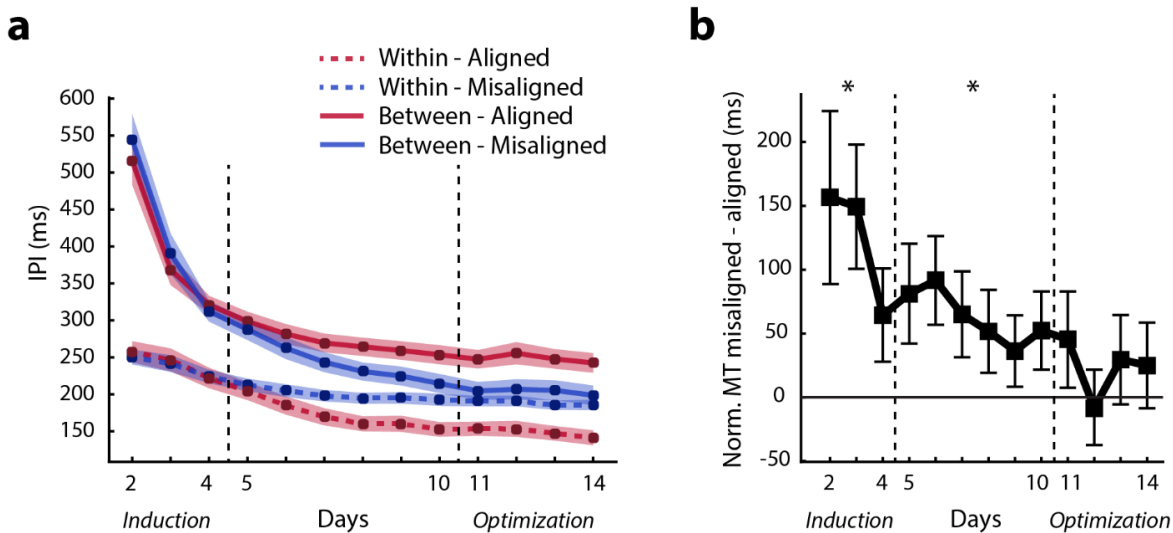
10

Figure 3. Change in chunk structure and performance for aligned and misaligned instructed sequences. (a) *Within- or between-chunk IPIs over days, separated by whether we they were in the aligned or misaligned instructed sequences.* **(b)** *Differences in movement time (MT) between sequences instructed with an aligned and misaligned way of chunking. Asterisk indicates a significant difference from 0 (no difference). Error bars denote between-subject standard error.*

Tracking changes in chunking

197
198 The IPI analysis presented in Figure 3a shows that participants changed their chunk
199 structure more for the misaligned than for the aligned chunk structure. What we cannot
200 discern from this analysis, however, is how participants restructured their chunking
201 behavior and whether they completely abandoned the misaligned chunk structure. To
202 get a clearer understanding of how participants changed their chunk structure, we used
203 a Bayesian model that allows us to estimates the probability of each possible chunk
204 structure on a trial-by-trial basis[34]. The state in this Hidden Markov Model indicated
205 which of the 1023 possible chunk structures was present on each trial. Using the
206 expectation-maximization (EM) algorithm[35,36], we simultaneously estimated the 9 free
207 parameters of the model (for details see Methods), as well as the posterior probability
208 for each possible chunk structure on each trial. We accounted for the effects of
209 biomechanical difficulty by regressing out an average IPI pattern of each possible finger

11

210   transition from individual IPIs before modeling. Importantly, our version of the algorithm

211   included a model for separate learning-related changes to the within- and between-

212   chunk intervals (Fig. 4a). Our method, therefore, allowed us to estimate participants'

213   chunk structure independent of the overall speed of performance. We confirmed this

214   independence using simulated data (see methods).

215   Figure 4b shows an example of an individual participant and sequence. The

216   participant chunked the sequence according to the initial instructions at first, then

217   inserted 1 or 2 additional chunk boundaries, and at the end of training performed the

218   sequence as a single chunk. To summarize these changes across participants and

219   sequences, we used a distance measure that characterizes differences between chunk

220   structures. To compute this distance, we simply counted the number of chunk

221   boundaries that differed between the two chunk structures – i.e. the number of chunks

222   that needed to be split or merged to transform one chunk structure into the other (Fig.
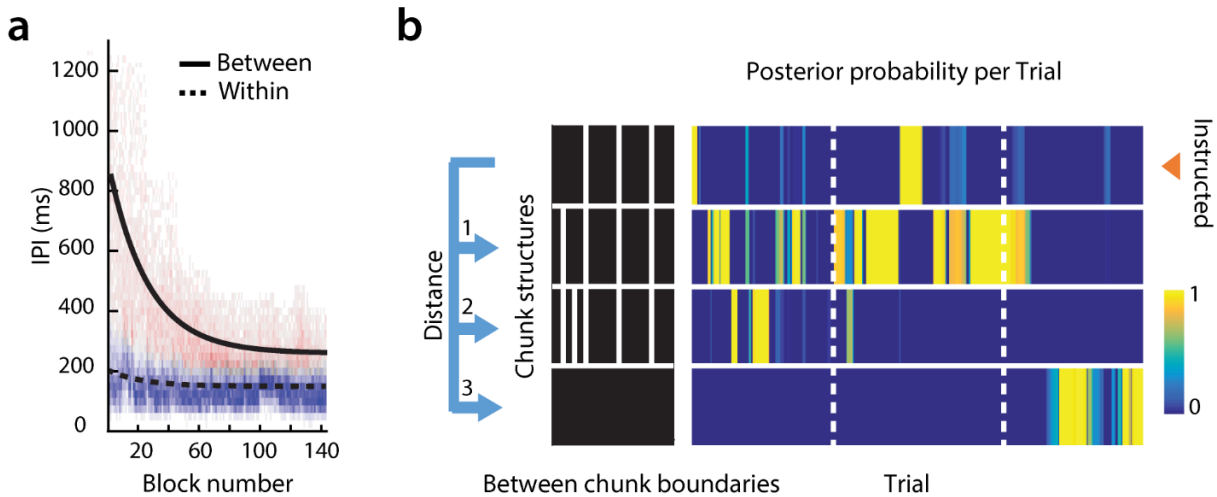
223   4b - distance).

Figure 4. Probabilistic chunking model fitted to example participant data. (a)
The change of within- and between-chunk IPIs were modeled using two separate
exponential functions across training. The density plot shows individual IPIs, with the
color indicating the probability to be a between- (pink) or within-chunk interval (blue).
(b) *Example of the posterior probability of 4 possible chunk structures (out of 1023)*
for a single participant and a single sequence over the course of the experiment. The
black box indicates chunk structures with white lines referring to chunk boundaries
within the 11-press sequence (max. 10 boundaries). The colored box indicates the
posterior probability of a given chunk structure for each trial - with yellow indicating
higher probabilities. The dashed vertical lines indicate the boundaries between
training weeks. The first row refers to the instructed chunk structure. The 2nd-3rd rows
differ from the instructed structure by an additional chunk boundary each. In the last
row the sequence is performed as a single chunk. The distance between two
structures indicates how many chunk boundaries have changed.

Movement towards a single chunk structure

To visualize learning-related changes in chunking, we projected the estimated chunk
structure into a 2-dimensional space (Fig. 5a). On the horizontal axis, we plotted the
expected distance of participants' chunk structure to the single-chunk structure. Given
the nature of our distance metric, this measure simply counts the number of chunk

13

229   boundaries. On the vertical axis, we plotted how close the estimated chunk structure

230   was to the aligned and misaligned chunk structure.

231   Previous literature has suggested that participants group smaller chunks together

232   with training, resulting in fewer chunk boundaries[25,28,32,37–39]. To test this idea, we

233   estimated the expected distance to a single chunk for each participant averaged across

234   sequences (the neutral sequence was excluded). Interestingly, on the 2nd day of

235   practice participants separated sequences into more chunks than the 3 chunks we

236   instructed (Fig. 5a, $t_{(31)}$ = 4.224, p = 0.0002). This tendency continued on day 3, where

237   participants tended to subdivide the sequences into even smaller chunks compared to

238   the day before (Fig. 5b; day 2 vs. 3: $t_{(31)}$ = 2.023, p = 0.052). After day three the

239   estimated number of chunk boundaries decreased as shown by a significant effect of

240   day in a repeated measures ANOVA ($F_{(11,341)}$ = 11.710, p < 1e-16). However, even in

241   the last phase of training, participants performed the sequences with an average of 2.9

242   chunk boundaries (we instructed 3 chunk boundaries). Thus, while there was a clear

243   tendency towards merging chunks, participants did not perform the sequence as a

244   single chunk, even after 3 weeks of practice.

245   Participants abandoned the misaligned chunk structure to a greater degree

246   Next, we probed how much participants diverged from our initial instructions. Our

247   analysis showed that participants slowly changed their chunk structure for both aligned

248   and misaligned instructed sequences with training. The average distance to the

249   instructed chunk structure increased systematically over time (repeated measures

250   ANOVA, effect of day, $F_{(12,372)}$ = 7.055, p < 1e-16, Fig. 5c). Thus, even the aligned

251   chunk structure underwent some changes with practice.

252   Consistent with our analysis of the IPIs (Fig. 3a), we observed that participants

253   abandoned the misaligned chunk structure to a greater degree than the aligned chunk

254   structure (Day x Instruction interaction: $F_{(12, 372)}$ = 5.610, p < 1e-16). The misaligned

255   instructed sequences showed a larger distance to the instructed chunk structure than

256   the aligned instructed sequences in the last four days of training: $t_{(31)}$ = 2.294, p = 0.029

257   (Fig. 5c). Additionally, we found a significant Day x Instruction interaction ($F_{(12, 372)}$ =

258   2.215, p = 0.011) for the distance to a single chunk (Fig. 5b), suggesting a greater drive

14

259   towards performing the sequence as a single chunk for the misaligned compared to the

260   aligned chunked sequences over time. Together these results indicate that participants

261   changed their chunking behavior for the misaligned instructions more quickly than for

262   the aligned instructions.

263   Despite the divergence from the misaligned chunk structure with training,

264   participants did not overcome the influence of the instruction completely. Even at the

265   end of training, the sequences that were instructed with the aligned chunk structure

266   showed a significantly smaller distance to the aligned chunk structure than to the

267   misaligned chunk structure ($t_{(31)}$ = 5.592, p < 1e-16). This observation was also true for

268   the sequences that were instructed with the misaligned chunk structure; the average

269   distance between the participants' estimated chunk structure and the misaligned chunk

270   structure was significantly smaller than to the aligned chunk structure ($t_{(31)}$ = 6.962, p <

271   1e-16). This finding clearly shows that even the misaligned instructions had a lasting

272   influence on the participants' motor behavior, even after three weeks of practice.

15
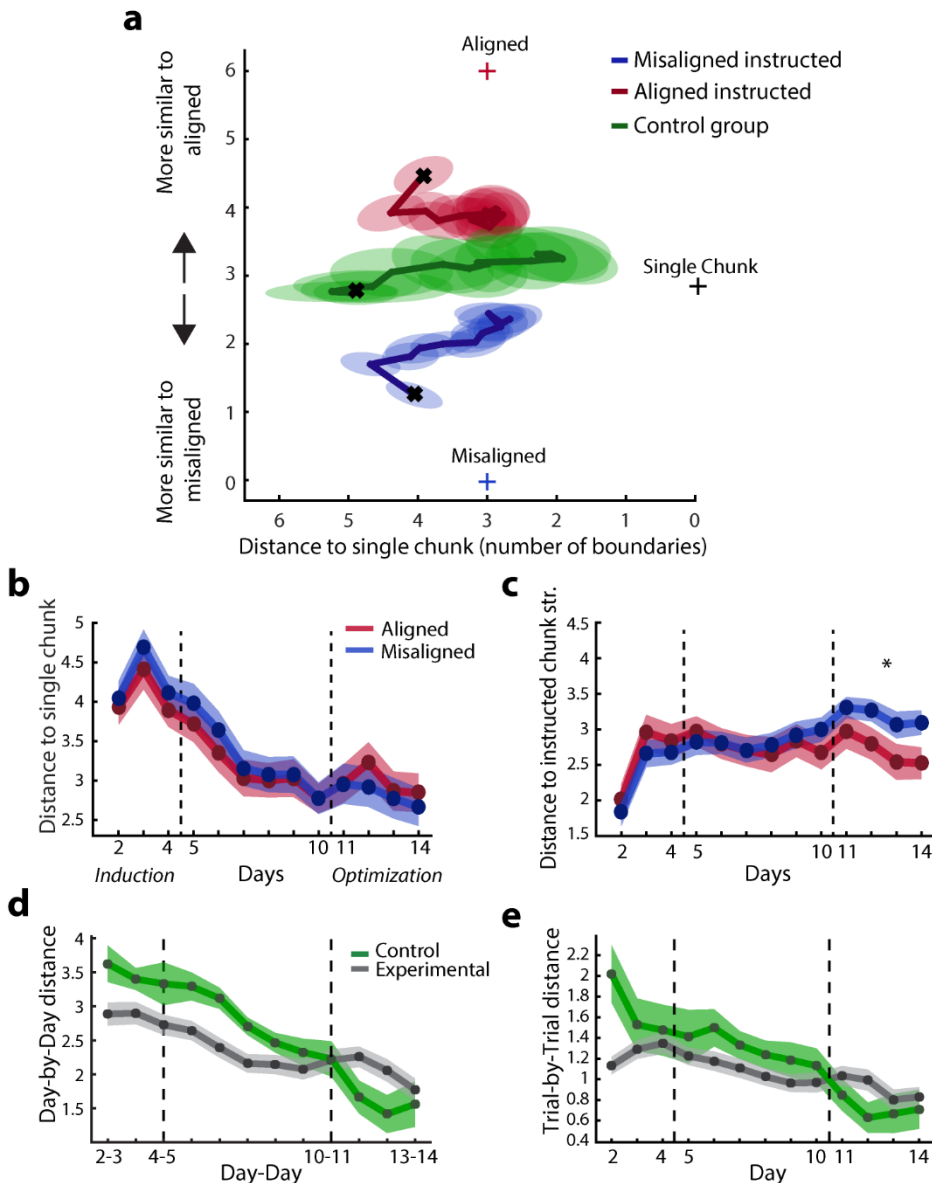
***Figure 5. Changes in chunk structure with learning**. **(a)** The average chunk structure over 13 days of practice for the two instructed chunk structures of the experimental groups and the control group. X-axis indicates the distance to the single-chunk structure, which counts the number of chunk boundaries. Y-axis represents how closely participants performed the sequence in accordance with the aligned or misaligned chunk structure. The "x" marks day 2. Ellipses denote the between-subject standard error. **(b)** Distance from a single chunk for aligned and misaligned instructed sequences across days. **(c)** Distance from the instructed chunk structure. **(d)** The distance between the average chunk structure between consecutive days. **(e)** The distance of chunk structures between consecutive trials of the same sequence within each day. Error bars indicate between-subject standard error.*

16

Chunk structure "crystallizes" with training

Would longer training allow participants to completely overcome the influence of the instruction? Would they end up performing all sequences as a single chunk? Although experiments with longer training are necessary to provide a definitive answer, our data indicate that this process, if indeed occurring, may take a very long time. We observed that the amount of change in the chunk structure for each sequence reduced dramatically in the last week of training, suggesting that a stable motor habit formed. This phenomenon is akin to the development of an invariant temporal and spectral structure in bird-song learning, a process that has been termed "crystallization"[40]. To estimate crystallization, we calculated the distance between the chunk structures from one day to the next (Fig. 5d) and within each day from one trial to the next (Fig. 5e). The analysis was performed separately for each sequence and participant. Overall, the day-to-day distance decreased dramatically over time (effect of day in repeated measurement ANOVA: $F_{(11,330)} = 18.794$, $p < 1e\text{-}16$). The trial-by-trial change in chunk structure across days showed the same pattern ($F_{(12,456)} = 13.245$, $p < 1e\text{-}16$). Therefore, participants appeared to settle onto a stable pattern in the last four days of training. Consequently, additional training would likely only lead to very slow changes in their chunk structure.

In summary, our analysis provides a clearer picture of how chunking changes with learning. Firstly, in line with previous research[25,28,32,37,38] participants gradually moved towards performing the sequence as a single chunk by dividing the sequence into fewer chunks. Secondly, participants diverged from the instructions over time with a quicker deviation from the misaligned chunk structure. Nevertheless, they did not completely overcome the initial instruction, nor did they perform the sequences as a single chunk at the end of training. Considering that the chunk structure crystallized in the last four days of training, these results demonstrate the formation of a stable motor habit that is still influenced by the initial instruction.

Spontaneously emerging chunk structures

In addition to the two experimental groups, we also tested a control group (N=8) to investigate how participants would spontaneously chunk the sequences. The control

17

303 group did not receive any explicit chunk training in the induction phase. Rather,

304 participants were presented with the entire 11-digit sequences on the first day and had

305 to memorize them without any reference to chunks starting on the second day (see

306 methods for details). In agreement with the experimental groups, the control group

307 showed a tendency to subdivide the sequences into smaller chunks in the beginning

308 and then slowly combine them into larger chunks. The distance to a single chunk

309 structure decreased significantly over days (repeated measures ANOVA, effect of day:

310 $F_{(12,84)}$ = 17.977, p < 1e-16), and reached a level that was statistically not different from

311 the experimental participants on the last day of training ($t_{(38)}$ = -0.940, p = 0.353).

312 Interestingly, the control group performed the sequences on the first day closer to the

313 misaligned chunk structure than to the aligned chunk structure ($t_{(7)}$ = -2.799, p = 0.027).

314 With training, however, participants moved closer to the aligned chunk structure, as

315 indicated by a significant change in the difference between the distance to the aligned

316 and misaligned chunk structure (repeated measures ANOVA, $F_{(12,84)}$ = 5.303, p < 1e-

317 16). The control group also showed clear crystallization over time. Compared to the

318 experimental groups, control participants showed a higher day-to-day and trial-by-trial

319 expected distance in the beginning of training, which then reduced more quickly (group

320 x day interaction; day-to-day: $F_{(11,330)}$ = 3.780, p = 4.003e-05; trial-by-trial: $F_{(12,456)}$ =

321 4.254, p = 2.167e-06). In summary, the control group showed similar behavioral

322 patterns to the experimental participants, indicating that similar processes of habit

323 formation are also at play in the absence of an explicit habit induction.

324 **Two optimization processes correlate with faster performance**

325 Having obtained a more complete picture of changes in chunking, we can now examine

326 to what degree these changes helped the participants to improve their performance. We

327 first asked whether performing the sequences closer to a single chunk would correlate

328 with performance. Within each participant, we estimated the relationship between the

329 MT of a specific sequence and the distance to the single chunk structure in the last four

330 days of training. For each participant, we regressed the MT for 6 sequences (excluding

331 the neutral sequence) against the average distance to the single chunk structure (Fig.

332 6a). The majority of the participants showed a positive relationship between the number

333 of chunks and MT: a one-sample t-test indicated that the individual slopes were

18

334 significantly greater than 0 (Fig. 6a, $t_{(31)}$ = 6.104, p = 4.560e-07). This significant

335 relationship was also found for the control participants (Fig. 6b, $t_{(7)}$ = 3.429, p = 0.006).

336 This finding provides clear evidence that performing the sequences with fewer chunks

337 correlates with improved performance.

338 Secondly, we investigated whether performing the sequences in alignment with

339 the biomechanical constraints would lead to fast performance. We regressed the MT for

340 6 sequences (excluding the neutral sequence) against the average distance to the

341 aligned chunk structure in the last four days of training. On average the individual

342 slopes again were significantly greater than 0, both for the experimental (Fig. 6c; $t_{(31)}$ =

343 2.220, p = 0.017), and for the control group (Fig. 6d, $t_{(7)}$ = 2.720, p = 0.015). These

344 results suggest that finding a particularly good way of chunking (for the same number of

345 chunk boundaries) also improves performance.
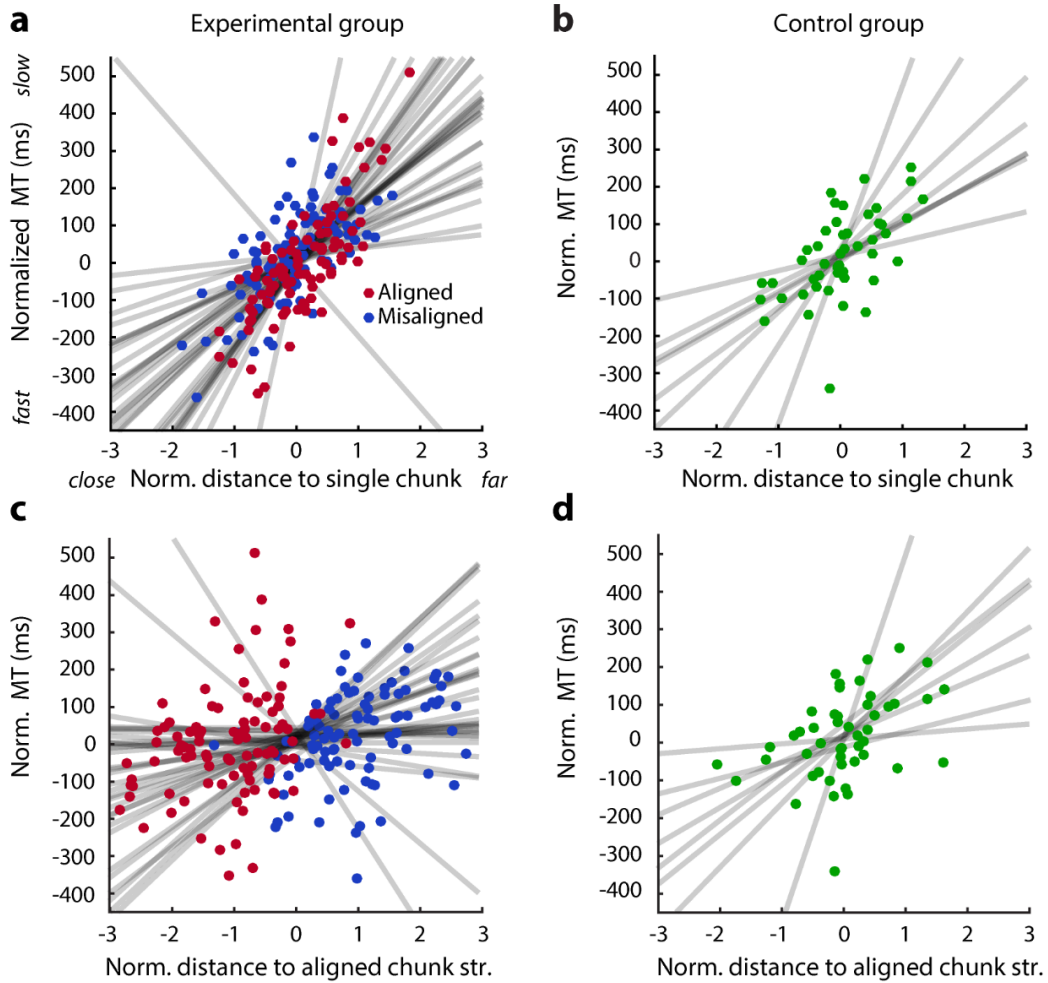
Figure 6. Relationship between the distance to the aligned/single chunk structure and MT. (a) *Scatterplot between the normalized (per subj.) distance to a single chunk and normalized MT in the last four days of practice. A separate regression line is fitted to the 6 sequences for each subject. Red dots indicate sequences with aligned instructions, blue dots sequences with misaligned chunking instructions.* **(b)** *Same as a but for the control group.* **(c&d)** *same as a & b but for the normalized distance to the aligned chunk structure.*

346   To visualize the relationship between the chosen chunk structure and the MT in

347   the last four days of training, we plotted the MT and chunk structure for each sequence

348   and participant in the 2-dimensional space defined in Fig. 5a (Fig. 7). The diameter of

349   each circle visually represents the value of the MT. This visualization clearly shows the

350   performance benefit from being closer to a single chunk as well as being closer to the

351   aligned chunk structure.

352   Overall, these results suggest that the two optimization processes - joining

353   chunks and aligning the remaining chunk boundaries with biomechanical constraints -

354   positively influence participants' speed with training. These analyses also showed that

355   sequences for which participants could not develop a better way of chunking were

356   performed substantially slower.

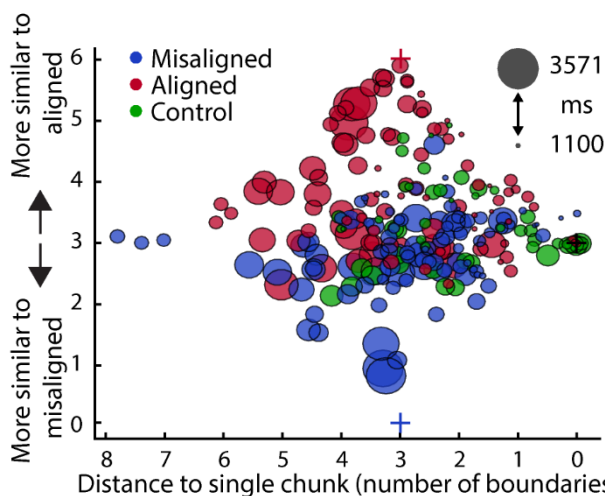***Figure 7. Relationship between chunking and speed (days 11-14).*** *The x-axis indicates the distance to a single chunk and the y-axis the relative distance to the two instructed chunk structures. Each data point indicates the average chunk structure and MT of a single sequences and participant in the last four days of training. The diameter of each circle represents the MT with larger circles indicating slower performance.*

Discussion

357   Our study utilizes a novel experimental paradigm to investigate how motor habits are

359   formed and how they may be overcome. We influenced the structure of the initial

360   declarative memory representation of the sequences by manipulating how participants

361   memorized them. By experimentally imposing two different chunk structures on the

362   same physical sequence, we could make causal inferences on the effects of cognitive

363   chunking on motor skill development. This is an important advance over previous

364   observational studies[28,32,41], which had no experimental control over how participants

365   chose to chunk the sequence.

366   Our data demonstrate that stable motor habits can be induced through cognitive

367   instruction and practice of these instructions during the initial phase of training. While

368   participants diverged from the instruction over time, they did not completely overcome

369   the initial chunk structure we imposed. Additionally, we found that participants' chunk

370   structure crystallized over the course of training, making it unlikely that the influence of

371   the initial instruction would disappear completely with longer training. Moreover, we

372   found that the motor habit remained stable, even when we changed from a memory-

373   guided to a stimulus-guided task. Thus, the initial instruction led to the formation of

374   motor habits that changed over time but were still clearly present after three weeks of

375   training.

376   Based on the independent measure of the speed of digit transitions, we designed

377   two alternative ways of chunking: One that was aligned and one that was misaligned

378   with the biomechanical difficulty of the transitions. Indeed, we could show that this

379   manipulation influenced performance in the first two weeks of practice. This

380   demonstrates that the structure of the declarative memory representation can either

381   facilitated or impede performance.

382   Our results also indicated that this bad habit could be overcome through practice:

383   Participants changed the misaligned chunk structure more rapidly than the aligned

384   chunk structure, and on the group level the performance detriment imposed by the

385   misaligned instruction was no longer significant in the last week of training. We

386   identified two ways by which participants overcame the limitation induced by the bad

387   habit.

388   First, participants joined chunks together, decreasing the amount of additional

389   time spent on chunk boundaries. While previous research has shown that the size of

390   chunks increases with training, it is debated whether this tendency affects performance

391   [28,29,42,43]. Using a Bayesian model that allowed us to assess chunk structure

392   independent of performance, we could demonstrate a clear positive relationship

393   between chunk concatenation and execution speed. This effect was not only seen in the

394   experimental group, but also in the control group which developed a chunking strategy

395   without explicit instructions. However, our results also indicate that participants did not

396   merge the sequence into a single chunk after 3 weeks of training, but still subdivided

397   each sequence into 3-4 chunks. This suggests that the length or duration of motor

398   actions that can be joined and optimized as a single element may be limited[32].

399   Given this constraint, participants also optimized performance by placing the

400   necessary chunk boundaries in a biomechanical efficient manner. Consistent with our

401   prediction based on the difficulty of individual digit transitions, placing chunk boundaries

402   at digit transitions that take more time to execute resulted in faster performance for the

403   full sequence. This optimization process was also observable in the control group that

404   memorized and practiced sequences on their own terms.

405   Conversely, we observed that sequences that were not chunked in line with

406   either of the two optimization strategies were performed slower. Therefore, if a more

407   beneficial way of chunking was not found, participants still showed a clear detriment,

408   suggesting that other learning mechanisms cannot fully make up for a persistent bad

409   habit. Considering that participants' behavior became highly invariant in the last week of

410   practice, we predict that some bad habit will remain and continue to influence

411   participants' performance even after prolonged training.

412   In many motor tasks there are numerous possible strategies that can lead to

413   excellent performance. Examining Figure 7, one can observe that the shortest MTs

414   were achieved anywhere in the space between the aligned and single chunk structure.

23

415   Occasionally, good performance was also reached in other positions in chunk space.

416   Participants adopted quite idiosyncratic chunk structures for each sequence at the end

417   of training. This suggests that there may be some inter-individual variability in which

418   technique works best for reaching a high level of performance. While we based our

419   biomechanical constraint estimates on a representative sample, it might not perfectly

420   reflect the constraints experienced by each participant. Alternatively, a number of ways

421   of chunking may work approximately equally well, such that the cost of changing an

422   established habit may outweigh the small benefit that could be gained from changing

423   the structure. A similar observation can be made in sports, where even top-ranked

424   athletes use slightly different techniques to reach similar performance levels. This

425   variation may reflect individual biomechanical differences but also differences in

426   instruction and training combined with subsequent habit formation.

427   While our experimental design enabled us to manipulate participants' habits in a

428   laboratory setting, sequence learning only captures a specific aspect of motor skill

429   acquisition. Nevertheless, similar persistence of habits has been observed in other

430   motor learning paradigms[45]. In bimanual coordination, for instance, Park et al.[46] showed

431   that acquired habits stayed remarkably stable even over 8 years of not performing the

432   task.

433   The establishment of a novel experimental paradigm to study habit formation in

434   skill learning is in itself important. In the future, it will allow us to explore how we can

435   encourage learners to abandon or change a current habit. While our attempt at

436   accelerating this process by changing the task from a memory-based to a stimulus-

437   based task was ultimately not successful, there are many other techniques that would

438   be possible. In many disciplines, teachers have developed ways to help students

439   overcome bad habits. For instance, the Hanon exercise is a piano technique to help

440   students play difficult passages of a musical piece by breaking up the learned musical

441   piece into new chunks to explore different rhythms. Playing a passage slower than

442   intended has also been suggested to break bad habits[47]. Overall, a clear message from

443   the diverse literature on learning piano is to diversify training and to practice with

444   diligence to prevent injury and bad habits from forming[48]. This suggests that changes in

24

445   context and providing students with new ways of exploring their skill might aid

446   performance and the abandonment of bad habits.

447   The current study yields new insights into how habits are formed during motor

448   skill learning and how the nature of the habit influences its stability. Furthermore, we

449   obtained insight into the optimization processes that are involved in overcoming a

450   detrimental habit. Our novel experimental paradigm now allows the further study of how

451   we can aid the abandonment of bad habits.

452   **Methods**

453   **Participants**

454   Forty neurologically healthy participants were recruited for the study (30 females; ages:

455   19 to 33). Thirty-two were randomly split into two experimental groups and the

456   remaining eight participants were assigned to the control group. All participants were

457   right-handed based on the Edinburgh Handedness Inventory and completed informed

458   consent. On average, participants had received 4.68 ($\pm$ 5.55) years of musical training,

459   with 55% percent having played the piano for more than half a year. While the amount

460   of musical training positively influenced performance speed, it did not affect chunk

461   behavior. The study protocol was approved by the ethics board of the University of

462   Western Ontario.

463   **Apparatus**

464   A custom-built five-finger keyboard was used. The keys were not depressible, but were

465   equipped with a force transducer (FSG-15N1A, Sensing and Control, Honeywell)

466   underneath each key which reliably measured participants' isometric force production

467   with a repeatability of <0.02N and a dynamic range of 16N [49–51]. The signal was

468   amplified and sampled at 200 Hz.

469   **Discrete sequence production task**

470   We used a discrete sequence production task (DSP), in which participants had to

471   execute sequences of 2, 3, or 11 keypresses as fast as possible while keeping their

472   error rate under 15% within each block. A trial was termed erroneous if participants

473   pressed a wrong key anywhere within the sequence. No pause between presses was

474   required and thus some co-articulation between fingers emerged with faster execution.

475   A finger press was detected when the produced force of a given finger crossed a

476   threshold of 3N. Subsequently, a finger was detected as released when the force of the

477   same finger fell below 1.5N. In order for a subsequent finger to be registered as pressed

478   the previous finger had to be released. This rule prevented participants to press with

479   more than 2 fingers at once. The force magnitude applied to each key was represented

480   by 5 lines on an LCD monitor, with the height of the line representing the force in the

481   corresponding finger. A white asterisk (memory-guided conditions) or digits (cued

482   condition) for each finger press was presented above the lines. Immediately after the

483   press threshold was reached, participants received visually and auditory feedback. If the

484   press was correct, the white cue changed its color to green and a sound was presented.

485   If the press was incorrect, the cue turned red and a lower-pitch sound was presented.

486   After each trial, participants received points based on their performance. Correct

487   sequences that were performed faster than the movement time (MT) threshold were

488   rewarded with 1 point. MTs that were 20% faster than the threshold were rewarded with

489   3 points. Incorrect presses or MTs exceeding the threshold resulted in 0 points. At the

490   end of each block, participants received feedback on their error rate, median MT, points

491   obtained during the block, and total points obtained during the session. In order to

492   maintain motivation, we adjusted the MT threshold necessary to obtain points. After

493   each block, we lowered the threshold by 500ms if participants performed with an error

494   rate of 15% or lower and had a median MT faster than the current threshold.

495   **Study design**

496   To impose a particular way of chunking onto participants' behavior, we first had

497   participants memorize and perform smaller 2-3 press chunks. These chunks were then

498   combined to form the training sequences. All participants were trained on the same 7

499   sequences, each consisting of 11 digit presses (see suppl. Table 1). Each participant

500   completed 15 training sessions in total: one session per day across a 3-week period

501   (excluding weekends). Each session lasted approximately 1 hour, excluding the two

502   initial sessions and the last session which took 2 hours.

503   To be able to verify that the chunking behavior was influenced by our instruction,

504   we used two different ways of chunking. We split each sequence either into one 2-digit

505   and three 3-digit chunks (2-3-3-3, misaligned) or into three 3-digit chunks and one 2-

506   digit chunk (3-3-3-2, aligned). Each participant practiced half of the sequences with one

507   chunk structure and the other half of the sequences with the other chunk structure. This

508   resulted in two groups of participants, which received training in two different sets of

509   chunks (suppl. Table 2). The control group did not receive any explicit chunking

510   instructions.

511   *Days 1-4: Chunk induction & initial sequence learning*

512   *Experimental group:* At the beginning of training the experimental groups were pre-

513   trained on a specific set of eleven 2- or 3-digit chunks (2 two-press chunks and 9 three-

514   press chunks). Participants received one of two different sets of chunks (suppl. Table

515   2). Each chunk was consistently associated with a letter of the alphabet (A-K).

516   Participants were explicitly told to learn this association. Each chunk was presented

517   twice in succession. In half of the blocks, on the first trial of each chunk presentation,

518   the numbers corresponding to the finger presses accompanied the letter on the screen

519   while on the second trial participants had to recall the presses solely based on the letter

520   (numbers were interchanged with stars). This trial order was reversed on every second

521   block. To ensure that participants had memorized the chunks we added speeded recall

522   blocks at the end of days 1 and 2. After practicing the 2-3 press chunks on day 1 and at

523   the beginning of day 2, participants trained on the seven 11-press sequences. Each

524   sequence was associated with a symbol (e.g. $; suppl. Table 1). Each sequence was

525   presented twice in succession and participant had to perform the sequences from

526   memory using the sequence cue on one trial or with the help of the chunk letters on the

527   next trial. We tested participants' sequence knowledge with a self-paced recall block at

528   the end of days 2-4.

529   *Control group:* The control group did not receive any chunk pre-training but trained

530   directly on the seven 11-press sequences. On the first day, the control participants

531   practiced the sequences using the digits presented on the screen. We matched the

532   amount of training across groups by ensuring that all participants were required to

27

533   produce the same number of finger presses. On the first day, the control participants

534   were not aware that they would have to memorize the sequences later on. On days 2-4

535   they then were instructed to memorize the sequences using the same sequence cues

536   as the experimental groups and were subsequently tested on their sequence

537   knowledge. The rest of the experimental design was identical for all groups.

538   *Days 5-10: Memory Recall*

539   For the days 5-10 of training participants practiced exclusively on the entire eleven-

540   press sequences and chunks were no longer mentioned or trained on. Each sequence

541   was presented twice in succession and participants had to recall the sequence from

542   memory on both trials using the sequence cue.

543   *Days 11-14: Memory recall or cued presentation*

544   During the last four days of training, half of the experimental participants performed the

545   sequences from memory while for the other half and for the control participants we

546   removed the sequence cue and presented participants with the actual numbers that

547   corresponded to the sequences (Fig. 1a). Participants completed an additional

548   generalization test on day 15. The results of this test is not reported in this article.

549   *Aligned vs. misaligned chunk structures*

550   To determine how to design our sequences and chunk structures to aid or impede

551   performance we needed to estimate the biomechanical difficulty of performing all

552   possible two and three finger transitions on our device. We trained a separate set of 7

553   participants for 3 days to perform each possible combination of 2 and 3 digit transitions

554   (e.g. 12 or 123). The required digit combination was presented on the screen, and

555   participants were instructed to execute the presses as fast as possible while trying to

556   keep their error rate below 15% in each block.

557   Among all possible 2-digit transitions, finger repetitions (e.g. 11) were executed

558   the slowest while transitions between neighboring fingers (e.g. 12) were especially fast

559   (Fig. 1b). Using these results, we designed the sequences such that they would include

560   both fast transitions (runs e.g. 123) and slow finger repetitions (113; suppl. Table 1).

561   Depending on which chunk structure was instructed, these transitions would either fall

28

562   on a chunk boundary or lie within a chunk. We counterbalanced this within and between

563   participants, meaning that each participant trained on 3 sequences with the aligned

564   chunk structure and 3 with the misaligned chunk structure (suppl. Table 1). One control

565   sequence was added which included a within-chunk run for both groups.

566   **Statistical Analysis**

567   We recorded and analyzed the force traces for each of the fingers. For each trial, we

568   calculated the reaction time (RT, time between presentation and first crossing of the

569   threshold), movement time (MT, time between first press and last release) and inter-

570   press-intervals (IPIs; time between force peaks of two consecutive presses). All

571   analyses were performed using custom-written code for MATLAB (the MathWorks). We

572   excluded trials that contained one or more incorrect presses from our analyses, as well

573   as trials with an MT or a press with an IPI that was three standard deviations above the

574   mean. The data were analyzed using mixed-effects analysis of variance (mixed

575   ANOVA), Pearson's correlation and paired and one sample t-tests. All t-tests were two-

576   sided. A probability threshold of $p < 0.05$ for the rejection of the null hypothesis was used

577   for all statistical tests. For the regression analyses as well as for calculating the MT

578   difference between the sequences with misaligned and aligned instruction we

579   normalized the data for each participant by subtracting the mean performance for each

580   day due to a wide range of performance speeds.

581   *Probabilistic model for estimating chunk structure*

582   We used an extended version of a Bayesian model of chunking behavior,

583   developed by Acuna et al.[34]. The algorithm uses a Hidden Markov Model to estimate the

584   posterior probability that a specific chunk structure is active on a given trial. As we had

585   10 digit transitions, each of which could either coincide with a chunk boundary or not,

586   we had to consider $2^{10}-1 = 1023$ possible chunk structures. Between trials, the hidden

587   Markov process could either stay in the same chunk structure with probability *p* or jump

588   to any other chunk structure with probability $(1-p)/1022$. The IPIs were modeled as a

589   Gaussian random variable, with a different mean and variance, depending on whether

590   the transition was within- or between-chunks. In contrast to Acuna et al.[34], where

591   learning effects were removed in a preprocessing step using a single exponential, we

29

592   modeled the learning within our model with two separate exponentials for the IPI mean.

593   This captured the faster reduction in the between- compared to the within-chunk

594   intervals (Fig. 2a). The inclusion of separate learning curves for within- and between-

595   chunk IPIs into the model allowed us to estimate participants' chunk structure

596   independently of the overall performance speed. This is an important advance over

597   previous methods that used a constant cutoff value to distinguish between within- and

598   between chunk intervals. For these methods, faster performance would automatically

599   decrease the number of chunk boundaries detected. To confirm that our algorithm did

600   not show this bias, we simulated artificial data using parameter estimates for individual

601   participants. We simulated sequences that switched between 4 different chunk

602   structures, each of which contained 4 chunks. Even though IPIs decreased by about

603   300ms with learning, the estimated average number of chunks remained stable across

604   the entire simulated experiment (average distance to single chunk: 3.35).

605   The model did not use errors and IPIs covariance structure, as these did not

606   relate systematically to the imposed chunk structure even early in training. We used an

607   Expectation-Maximization algorithm to simultaneously estimate the posterior probability

608   of each chunk structure for each trial, as well as the 9 parameters of the model: 3

609   parameters each for the exponential curve for the within- and between chunk IPIs, 1

610   variance parameter for each, and the transition probability *p*.

611   As a preprocessing step, we regressed the IPIs for each subject against the

612   average biomechanical profile, which was estimated as the average IPI profile for all

613   possible 2 digit-presses from our pilot experiment (Fig. 1b). The fitted values were

614   removed from the IPIs. Removing temporal regularities that could be modeled with

615   biomechanics alone should result in chunking estimates that more closely reflect

616   cognitive and learning influences. Qualitatively comparable results were also obtained

617   using the raw IPIs, without biomechanical factors removed.

618   *Expected distance*

619   We quantified how much participants changed their chunking behavior over time by

620   calculating the expected distance between two estimated chunk structures. The

621   distance between two chunk structures, $d(i,j)$, was defined as how many of the 10

30

622   transitions would have to change from a chunk boundary to a non-boundary (and vice

623   versa) to transform one structure into the other (for an example, see Fig. 4b). A distance

624   of 0 would indicate no change and the average distance between two randomly chosen

625   chunk structures is 6. Because we could not be 100% certain which chunk structure

626   participants adopted in each trial, we calculated the expected distance. For this, we first

627   calculated a 1023 X 1023 matrix containing the distances between any chunk structure

628   *i*, and chunk structure *j*. From the posterior probability distribution, we could then derive

629   how likely each of these chunk structure changes was, *p(i,j)*. The expected value of the

630   distance was then simply calculated as

631    $E(d) = \sum_{i=1}^{1023} \sum_{j=1}^{1023} p(i,j)d(i,j).$

632   *Code availability*

633   Bayesian algorithm code available on GitHub:

634   https://github.com/jdiedrichsen/chunk_inference.

635   Custom MATLAB code is available from the corresponding author on reasonable

636   request.

637   *Data availability*

638   The datasets generated during and analyzed during the current study are available from

639   the corresponding author on reasonable request.

Supplementary Table 1. Sequences and chunk structures for the two experimental groups.
Displayed are the 7 sequences used together with the associated sequence cue. The black vertical lines indicate the chunk boundaries that were imposed. Chunk structures were either aligned with the biomechanical requirements (red) or misaligned (blue). The last sequence (green) was included as a control sequence that was chunked either with a 3-3-3-2 or 2-3-3-3 structure but performance wise should lead to similar speeds as for both chunk structures the boundaries were placed at biomechanically slow transitions. This sequence was not included in the analyses.

Sequence	Group 1	Group 2
?	1 2 3|5 1 3|3 2 1|3 4	1 2|3 5 1|3 3 2|1 3 4
+	4 3 5|5 1 2|3 5 3|3 2	4 3|5 5 1|2 3 5|3 3 2
%	4 3 5|5 1 2|3 5 3|3 2	4 3|5 5 1|2 3 5|3 3 2
$	3 2|1 1 3|5 1 2|3 4 3	3 2 1|1 3 5|1 2 3|4 3
&	3 2|1 1 3|5 1 3|4 4 3	3 2 1|1 3 5|1 3 4|4 3
@	3 4|1 1 3|5 1 3|4 4 3	3 4 1|1 3 5|1 3 4|4 3
# (control sequence)	3 4|1 2 3|5 1 2|3 4 3	3 4 1|2 3 5|1 2 3|4 3

\|Chunk Boundary ▬Aligned ▬Misaligned

32

Supplementary Table 2. Instructed chunks and sequences for the two experimental groups. *The tables depict the finger presses (1-5) that were associated with the chunk cues (A-K) and the chunks that were associated with the sequences cues (symbols).*

Chunks Group 1

Chunk	Presses		
A	3	2	
B	3	4	
C	1	1	3
D	1	2	3
E	3	2	1
F	3	4	3
G	3	5	3
H	4	3	5
I	4	4	3
J	5	1	2
K	5	1	3

Chunks Group 2

Chunk	Presses		
A	1	2	
B	4	3	
C	1	2	3
D	1	3	4
E	1	3	5
F	2	3	5
G	3	2	1
H	3	3	2
I	3	4	1
J	3	5	1
K	5	5	1

Sequences Group 1

Seq.	Chunks			
?	D	K	E	B
+	H	K	E	B
%	H	J	G	A
$	A	C	J	F
&	A	C	K	I
@	B	C	K	I
#	B	D	J	F

Sequences Group 2

Seq.	Chunks			
?	A	J	H	D
+	B	K	H	D
%	B	K	F	H
$	G	E	C	B
&	G	E	D	B
@	I	E	D	B
#	I	F	C	B

640   References

641   1.   Ericsson, K. A. *et al.* The role of deliberate practice in the acquisition of expert

642   performance. *Psychol. Rev.* **100,** 363–406 (1993).

643   2.   Hayes, J. R. *The Complete Problem Solver*. (Taylor & Francis, 2013).

644   3.   Haith, A. M. & Krakauer, J. W. The multiple effects of practice: skill, habit and reduced

645   cognitive load. *Curr. Opin. Behav. Sci.* **20,** 196–201 (2018).

646   4.   Hélie, S., Waldschmidt, J. G. & Ashby, F. G. Automaticity in rule-based and information-

647   integration categorization. *Attention, Perception, Psychophys.* **72,** 1013–1031 (2010).

648   5.   Hardwick, R. M., Forrence, A. D., Krakauer, J. W. & Haith, A. M. Skill Acquisition and Habit
649   Formation as Distinct Effects of Practice. *bioRxiv* (2017). doi:10.1101/201095

650   6.   Ashby, F. G., Ell, S. W. & Waldron, E. M. Procedural learning in perceptual categorization.
651   *Mem. Cogn.* **31,** 1114–1125 (2003).

652   7.   Jager, W. Breaking ' bad habits ': a dynamical perspective on habit. *Hum. Decis. Mak.*
653   *Environ. Perception. Underst. Assist. Hum. Decis. Mak. Real-life Settings. Lib. Americum*
654   *Charles Vlek.* 149–160 (2003).

655   8.   Graybiel, A. M. Habits, rituals, and the evaluative brain. *Annu. Rev. Neurosci.* **31,** 359–387
656   (2008).

657   9.   Robbins, T. W. & Costa, R. M. Habits. *Curr. Biol.* **27,** R1200–R1206 (2017).

658   10. Seger, C. A. & Spiering, B. J. A critical review of habit learning and the Basal Ganglia.
659   *Front. Syst. Neurosci.* **5,** 1–9 (2011).

660   11. Dezfouli, A. & Balleine, B. W. Habits, action sequences and reinforcement learning. *Eur. J.*
661   *Neurosci.* **35,** 1036–1051 (2012).

662   12. Dolan, R. J. & Dayan, P. Goals and habits in the brain. *Neuron* **80,** 312–325 (2013).

663   13. Smith, K. S. & Graybiel, A. M. Using optogenetics to study habits. *Brain Res.* **1511,** 102–
664   114 (2013).

665   14. Graybiel, A. M. & Grafton, S. T. The Striatum: Where Skills and Habits Meet. *Cold Spring*
666   *Harb. Perspect. Biol.* **7,** a021691 (2015).

667   15. Moors, A. & De Houwer, J. Automaticity: A theoretical and conceptual analysis. *Psychol.*
668   *Bull.* **132,** 297–326 (2006).

669   16. Hélie, S. & Cousineau, D. The cognitive neuroscience of automaticity: Behavioral and
670   brain signatures. *Cogn. Sci. Hauppauge* **6,** 35–53 (2011).

671 17. Abrahamse, E. L., Ruitenberg, M. F. L., de Kleine, E. & Verwey, W. B. Control of
672   automated behavior: insights from the discrete sequence production task. *Front. Hum.*
673   *Neurosci.* **7,** 1–16 (2013).

674 18. Verwey, W. B. Concatenating familiar movement sequences: The versatile cognitive
675   processor. *Acta Psychol. (Amst).* **106,** 69–95 (2001).

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

678 20. Wong, A. L., Lindquist, M. A., Haith, A. M. & Krakauer, J. W. Explicit knowledge enhances
679   motor vigor and performance: motivation versus practice in sequence tasks. *J.*
680   *Neurophysiol.* **114,** 219–232 (2015).

681 21. Kleine, E. De & Verwey, W. B. Representations underlying skill in the discrete sequence
682   production task: effect of hand used and hand position. *Psychol. Res. Psychol. Forsch.* **73,**
683   685–694 (2009).

684 22. Verwey, W. B. & Dronkert, Y. Practicing a Structured Continuous Key-Pressing Task:
685   Motor Chunking or Rhythm Consolidation? *J. Mot. Behav.* **28,** 71–79 (1996).

686 23. Verwey, W. B., Abrahamse, E. L. & Jiménez, L. Segmentation of short keying sequences
687   does not spontaneously transfer to other sequences. *Hum. Mov. Sci.* **28,** 348–361 (2009).

688 24. Verwey, Verwey, W. B., Abrahamse, E. L. & de Kleine, E. Cognitive processing in new and
689   practiced discrete keying sequences. *Front. Psychol.* **1,** 32 (2010).

690 25. Verwey, W. B. Buffer loading and chunking in sequential keypressing. *J. Exp. Psychol.*
691   *Hum. Percept. Perform.* **22,** 544–562 (1996).

692 26. Miller, G. A. The magical number seven, plus or minus two: some limits on our capacity
693   for processing information. *Psychol. Rev.* **63,** 81–97 (1956).

694 27. Halford, G. S., Wilson, W. H. & Phillips, S. Processing capacity defined by relational
695   complexity: implications for comparative, developmental, and cognitive psychology.

696   *Behav. Brain Sci.* **21,** 803–864 (1998).

697   28.   Wymbs, N. F. F., Bassett, D. S. S., Mucha, P. J. J., Porter, M. A. A. & Grafton, S. T. T.
698   Differential Recruitment of the Sensorimotor Putamen and Frontoparietal Cortex during
699   Motor Chunking in Humans. *Neuron* **74,** 936–946 (2012).

700   29.   Solopchuk, O., Alamia, A., Olivier, E., Ze, A. & Zénon, A. Chunking improves symbolic
701   sequence processing and relies on working memory gating mechanisms. *Learn. Mem.* **23,**
702   108–112 (2016).

703   30.   Bo, J. & Seidler, R. D. Visuospatial working memory capacity predicts the organization of
704   acquired explicit motor sequences. *J. Neurophysiol.* **101,** 3116–3125 (2009).

705   31.   Seidler, R. D., Bo, J. & Anguera, J. A. Neurocognitive contributions to motor skill learning:
706   The role of working memory. *J. Mot. Behav.* **44,** 445–453 (2012).

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

709   33.   Verwey, W. B. Evidence for a multistage model of practice in a sequential movement
710   task. *J. Exp. Psychol. Hum. Percept. Perform.* **25,** 1693–1708 (1999).

711   34.   Acuna, D. E. *et al.* Multifaceted aspects of chunking enable robust algorithms. *J.*
712   *Neurophysiol.* **112,** 1849–1856 (2014).

713   35.   Welch, L. R. Hidden Markov Models and the Baum-Welch Algorithm. *IEEE Inf. Theory Soc.*
714   *Newsl.* **53,** 1,10-13 (2003).

715   36.   Dempster, A. P., Laird, N. M. & Rubin, D. B. Maximum likelihood from incomplete data
716   via the EM algorithm. *J. R. Stat. Soc. Ser. B Methodol.* **39,** 1–38 (1977).

717   37.   Sakai, K., Kitaguchi, K. & Hikosaka, O. Chunking during human visuomotor sequence
718   learning. *Exp. Brain Res.* **152,** 229–242 (2003).

719   38.   Kuriyama, K., Stickgold, R. & Walker, M. P. Sleep-dependent learning and motor-skill

720   complexity. *Learn. Mem.* **11,** 705–713 (2004).

721   39.   Bassett, D. S. *et al.* Dynamic reconfiguration of human brain networks during learning.
722   *Proc. Natl. Acad. Sci.* **108,** 7641–7646 (2011).

723   40.   Brainard, M. S. & Doupe, A. J. What songbirds teach us about learning. *Nature* **417,** 351–
724   358 (2002).

725   41.   Wright, D. L., Rhee, J.-H. J.-H. H., Vaculin, A., Rhee, J.-H. J.-H. H. & Vaculin, A. Offline
726   improvement during motor sequence learning is not restricted to developing motor
727   chunks. *J. Mot. Behav.* **42,** 317–324 (2010).

728   42.   Song, S. & Cohen, L. Impact of conscious intent on chunking during motor learning.
729   *Learn. Mem.* **21,** 449–51 (2014).

730   43.   Verstynen, T. *et al.* Dynamic Sensorimotor Planning during Long-Term Sequence
731   Learning: The Role of Variability, Response Chunking and Planning Errors. *PLoS One* **7,**
732   e47336 (2012).

733   44.   Jiménez, L. Taking patterns for chunks: Is there any evidence of chunk learning in
734   continuous serial reaction-time tasks? *Psychol. Res.* **72,** 387–396 (2008).

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

737   46.   Park, S.-W., Dijkstra, T. M. H. & Sternad, D. Learning to never forget—time scales and
738   specificity of long-term memory of a motor skill. *Front. Comput. Neurosci.* **7,** 1–13 (2013).

739   47.   Chang, C. C. *Fundamentals of Piano Practice*. *Brain* (CreateSpace Independent Publishing
740   Platform, 2004).

741   48.   Sadnicka, A., Kornysheva, K., Rothwell, J. C. & Edwards, M. J. A unifying motor control
742   framework for task-specific dystonia. *Nat. Rev. Neurol.* **14,** 116–124 (2018).

743   49.   Wiestler, T. & Diedrichsen, J. Skill learning strengthens cortical representations of motor

744   sequences. *Elife* **2,** 1–20 (2013).

745   50.   Wiestler, T., Waters-Metenier, S. & Diedrichsen, J. Effector-Independent Motor Sequence

746   Representations Exist in Extrinsic and Intrinsic Reference Frames. *J. Neurosci.* **34,** 5054–

747   5064 (2014).

748   51.   Yokoi, A. *et al.* Restricted transfer of learning between unimanual and bimanual finger

749   sequences. *J. Neurophysiol.* **117,** 1043–1051 (2017).

750