1

1

2

3

4

5   **The Evolution of Chunks in Sequence Learning**

6

7

8   Laure Tosatto[1,2], Joël Fagot[1,2], Dezso Nemeth[3,4,5] and Arnaud Rey[1,2]

9   [1]Aix Marseille Univ, CNRS, LPC, Marseille, France

10   [2]Aix Marseille Univ, ILCB, Aix-en-Provence, France

11   [3] Université Claude Bernard Lyon 1, Lyon Neuroscience Research Center, Lyon, France

12   [4] Institute of Psychology, Eotvos Lorand University, Budapest, Hungary

13   [5] Institute of Cognitive Neuroscience and Psychology, Hungarian Academy of Sciences,

14   Budapest, Hungary

15

16

17

18

19

20

21

22

23   Correspondence concerning this article should be addressed to Laure Tosatto,

24   Laboratoire de Psychologie Cognitive, CNRS, Aix-Marseille Université, 3, place Victor Hugo,

25   Case D, 13331 Marseille Cedex 03, France. Email: laure.tosatto@univ-amu.fr

26   **Abstract**

27   Chunking mechanisms are central to several cognitive processes and notably to the acquisition

28   of visuo-motor sequences. Individuals segment sequences into chunks of items to perform

29   visuo-motor tasks more fluidly, rapidly, and accurately. However, the exact dynamics of

30   chunking processes in the case of extended practice remain unclear. Using an operant

31   conditioning device, eighteen Guinea baboons (*Papio papio*) produced a fixed sequence of nine

32   movements during 1,000 trials by pointing to a moving target on a touch screen. Response times

33   analyses revealed a specific chunking pattern of the sequence for each baboon. More

34   importantly, we found that these patterns evolved during the course of the experiment, with

35   chunks becoming progressively fewer and longer. We identified two chunk reorganization

36   mechanisms: the *recombination* of preexisting chunks and the *concatenation* of two distinct

37   chunks into a single one. These results provide new evidence on chunking mechanisms in

38   sequence learning and challenge current models of associative and statistical learning.

39   *Keywords:* chunking, sequence learning, statistical learning, associative learning, non-

40   human primate

The Evolution of Chunks in Sequence Learning

A key mechanism allowing our cognitive system to compress information and increase short term memory capacity is the formation of chunks (Mathy & Feldman, 2012; Miller, 1956). Chunking is defined as the process of associating and grouping several items together into a single processing unit (Gobet & al, 2001; Gobet, Lloyd-Kelly and Lane, 2016). In coherence with the limits encountered by our cognitive system (Cowan, 1988, 2017), the storage capacity of the chunks themselves seems limited (3 or 4 items per chunk, Allen & Coyne, 1988; Chase & Simon, 1973; Johnson, 1970). In the field of perceptual-motor learning, chunking has been considered as the main motor sequence integration mechanism (Diedrichsen & Kornysheva, 2015; Wymbs et al. 2012).

Perceptual-motor sequence learning is commonly described as the process by which a sequence of movements is acquired and executed with increased speed and accuracy (Willingham, 1998). Current sequence learning paradigms use a sequential button-press task: subjects are presented with a series of visual stimuli organized in a sequence and are asked to press a corresponding response button (e.g., Cohen, Ivry, & Keele, 1990; Nemeth et al., 2010; Verwey, 2001). A faster and more accurate performance over time reflects learning of the sequence. Typically, researchers have reported robust effects of sequence's length on subjects' response times (RTs): for short sequences (<4 keypresses), a decrease in the successive RTs can be observed when the sequence is learned; for longer sequences though, an initial decrease in RTs for the first 3 to 4 keypresses is followed by a longer RT on the next position. Then, RTs start decreasing again for the next 3 to 4 keypresses (e.g., Bo & Seidler, 2009; Verwey & Eikelboom, 2003). Similar results have been observed in non-human subjects (Terrace, 2002; Scarf et al., 2018). This phenomenon has been interpreted as the long sequence being spontaneously segmented into shorter motor chunks reflecting the sequence organization in memory (Sakai et al., 2003). Long temporal gaps between responses are assumed to mark chunk

66   boundaries (Abrahamse et al., 2013; Bottary et al., 2016). Yet, few studies have been interested

67   in the evolution of chunks in the case of extended practice (e.g., Ramkumar et al., 2016; Song

68   & Cohen, 2014; Wymbs et al., 2012). The present study aims at collecting new evidence on the

69   evolution of chunking mechanisms in a perceptual-motor sequence learning task.

70   One of the main issues in most sequence learning studies is the inconsistent transition

71   probabilities (TPs) between items of the sequence. Indeed, in an experimental set up where only

72   a few response buttons are available to the subject (e.g., Grafton, Hazeltine and Ivry, 2002;

73   Verwey, 2001; Willingham, 1999), some stimuli are necessarily presented multiple times

74   within the same sequence. For instance, in a sequence such as A-B-A-C-D-B-C, A is either

75   followed by B or C, therefore the probability of B given A is 0.5, whereas D is always followed

76   by B thus the probability of B given D is 1. This heterogeneity of TPs may constrain the strength

77   of the connections formed between the successive elements of the sequence and affect the

78   resulting chunking pattern.

79   Another characteristic of previous studies is, apart from a few exceptions (e.g., Conway

80   & Christiansen, 2001), that most sequence learning studies have been conducted with human

81   participants. However, even if implicit instructions are provided, humans almost systematically

82   use their inner language and develop strategies to perform the task (Rey et al., 2019). This may

83   affect both their performance and their chunking processes, and it is hard to tease apart these

84   verbal and explicit influences from associative and chunking mechanisms.

85   To avoid these difficulties, we first chose to test non-human primates for neutralizing

86   the possible effect of language-based strategies. Second, we used sequences composed of items

87   only occurring once per sequence, i.e., with a transitional probability between elements equal

88   to 1, using the serial pointing task proposed by Minier et al. (2016). In this task, participating

89   monkeys had to track and touch target locations in a 3 x 3 matrix of crosses displayed on the

90   screen. They were exposed to the same sequence of 9 locations that always appeared in the

91   same order upon a thousand trials. RTs for each location were recorded, providing us with

92   detailed information on the temporal dynamics of sequence processing. This extended practice

93   on a long single repeated sequence allowed us to collect new evidence on the formation and the

94   evolution of chunks over time.

Method

Participants

97   Thirteen female and five male Guinea baboons (*Papio papio*, age range 2.8—23.7 years)

98   from the CNRS primate facility in Rousset (France), living in a social group of 25 individuals,

99   were tested in this study. Water was provided *ad libitum* during the test, and the monkeys

100   received their normal food ration of fruits every day at 5 PM. For practical reasons, we stopped

101   the experiment after 18 monkeys completed all scheduled trials.
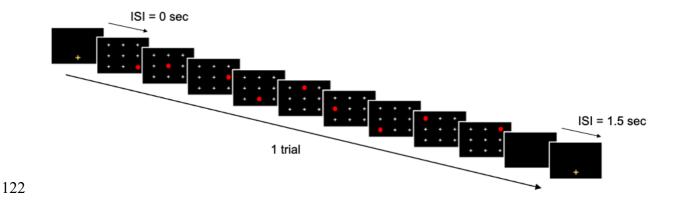
Materials

Apparatus

104   This experiment was conducted with a computer-learning device based on the voluntary

105   participation of baboons (for details, see Fagot & Bonté, 2010). Baboons implanted with a RFID

106   microchip had free access to 10 automatic operant conditioning learning devices equipped with

107   touch screens. Each time a monkey entered a test chamber, it was identified by its microchip,

108   and the system resumed the trial list where the subject left it at its previous visit. The experiment

109   was controlled by E-prime (Version 2.0, Psychology Software Tools, Pittsburgh, PA, USA).

Task and stimuli

111   The screen was divided into nine equidistant predetermined locations represented by

112   white crosses on a black background, virtually labeled as position 1 to 9 (see Figure 1). A trial

113   began with the presentation of a yellow fixation cross at the bottom of the screen. Once pressed,

114   the fixation cross disappeared and the nine white crosses were displayed, one of them being

115   replaced by the target, a red circle. When the target was touched, it was immediately replaced

116   by the cross. The red circle then replaced the next position in the sequence until it was touched,

117   and a new position was displayed. Reward was provided at the end of a sequence of nine

118   touches.

119

120   **Figure 1**

121   *Experimental display and stimuli presentation*

122

123
124
125   To learn the task, baboons initially received random trials that were rewarded after three

126   touches. Then, the number of touches in a trial was progressively increased up to nine. If

127   baboons touched an inappropriate location (incorrect trial) or failed to touch the screen within

128   5,000ms after the red circle's appearance (aborted trial), a green screen was displayed for

129   3,000ms as a marker of failure. Aborted trials were not retained and therefore presented again,

130   while incorrect trials were not. The time elapsed between the appearance of the red circle and

131   the baboon's touch on this circle was recorded as the RT.

132   ***Design of the sequences***

133   To control the motor difficulty of the transitions to be produced in the sequence, a

134   random phase of sequence production was first conducted, where thirteen baboons performed

135   random sequences of six positions for 1,000 trials. Based on these random trials, a baseline

136   measure for all possible transitions from one location to another was computed by calculating

137   mean RTs for each transition (e.g., from position 1 to 9), leading to a 9x9 matrix of mean RTs

138   calculated over the entire group of baboons (see Appendix A).

139   Based on these baseline measures, we designed two sequences of nine serial positions

140   for which each transition T was faster (or equally fast) to produce than the next one (i.e.,

141   T1≤T2≤…≤T8; see Appendix B). This way, a decrease in RT for a given transition can be

142   interpreted as the anticipation (or learning) of that position from the previous one.

Procedure

144   Baboons were either presented with Sequence 1 (N=8) or Sequence 2 (N=10) and had

145   to produce it for 1000 successive trials. RTs for each position of the sequence were recorded

146   for all the trials.

<div align="center">

Results

</div>

148   On average, baboons required 2.9 days to complete the 1,000 trials, with a mean of

149   339.7 trials per day and a mean accuracy level of 92.5% (*SD*=6.2%). Incorrect trials were

150   removed from the dataset (7.8%). A recursive trimming procedure excluded RTs greater or

151   smaller than 2.5 standard deviations from the subject's mean for each of the nine possible

152   positions (24.4%)[1]. RTs for each of the nine positions and for the 1,000 trials were divided into

153   10 Blocks of 100 trials.

General sequence learning

155   Learning of the sequence was estimated on mean RTs by a repeated measures one-way

156   ANOVA with Block (1-10) as the within factor. The effect of Block was highly significant

157   (Block 1, *M*=452.8, *SD*=45.3; Block 10, *M*=400.1, *SD*=56.3), $F(1,9)=30.43$, $p<.001$, $\eta^2=0.642$,
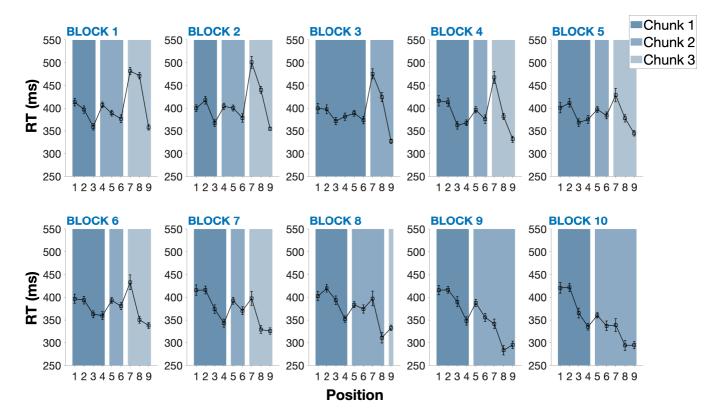
---

[1] Inspection of the response times distribution revealed that a majority of responses were produced around 500ms. A smaller group of RTs appeared around 1,000 ms and was likely due to situations in which baboon's response was not recorded by the computer, because their hands were dirty. In this situation, they had to touch the screen again, and longer RTs were recorded (that are on average twice longer compared to the first responses). This is why we have adopted this recursive trimming procedure. After applying this procedure, there was still a mean number of 77.9 remaining RTs for each position per participant and per block of 100 trials (the minimum number of remaining RTs for one position being 32).

158 indicating that mean RTs decreased throughout the blocks of trials and that monkeys learned

159 the sequence.

Chunking of the sequence

161 To study the chunking pattern of the sequence by monkeys, we adopted the following

162 reasoning. We considered successive positions A and B to be part of the same chunk as long as

163 the transition time from one position to the next did not correspond to a significant increase in

164 RT. Therefore, if $RT_A < RT_B$ then A and B do not belong to the same chunk, and the AB transition

165 marks a chunk boundary (Kennerly et al., 2004). This reasoning was applied on each Block and

166 for each monkey. Statistical significance was assessed through paired-sample t-tests for each

167 pair of successive positions. Each time the RT of a pair's second position was significantly

168 higher than the first position, it marked a chunk boundary. Figure 2 illustrates this method on

169 the performances of one monkey.

170

171 **Figure 2**

172 *Evolution of the chunking pattern for one individual (Atmosphere) throughout the task*

173

174 *Note.* Mean RT per position across the 10 blocks of trials for one baboon (Atmosphere) showing the evolution of

175 the chunking pattern. This individual initially parses the sequence into three chunks of three positions in the first

176 three blocks of trials. A reorganization starts in Block 4, as Position 4 is integrated into the first chunk. Another

177 reorganization occurs in Block 8, where the 2nd and 3rd chunks are progressively concatenated (error bars represent
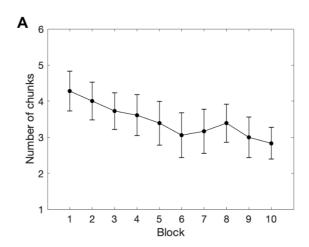
178 95% confidence intervals).

179

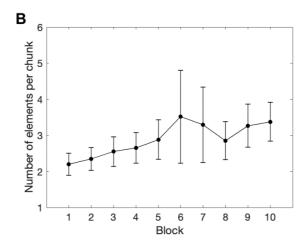180 **Evolution of chunks**

181 With this method, we were able to quantify the number of chunks and their average size

182 produced on each block by each monkey. For example, the monkey from Figure 2 produced 3

183 chunks and an average chunk size of 3 in Block 1. In Block 10, the number of chunks was 2

184 and the average chunk size was 4.5. Figure 3 reports the average number of chunks and chunk

185 size for all monkeys across the 10 blocks of the experiment. For example, we found that the

186   mean chunk size was 2.2 positions (*CI*=0.31; *Min*=1; *Max*=5) in the first block and 3.375

187   (*CI*=0.53; *Min*=1; *Max*=8) in the last block. Two repeated measures one-way ANOVA were

188   conducted to test the effect of Block on the mean number of chunks and the average chunk size.

189   This analysis revealed that the number of chunks significantly decreased across blocks,

190    $F(1,9)=9.421$, $p<.001$, $\eta^2=.357$, and that the average chunk size significantly increased across

191   blocks, $F(1,9)=4.794$, $p<.01$, $\eta^2=.22$.

192

193   **Figure 3**

194   *Evolution of chunks across blocks*

195

196   *Note*. A. Mean number of chunks per block. B. Mean chunk size (i.e. number of elements per chunk) per block

197   (error bars represent 95%confidence intervals).

198   The evolution of these two indicators can be accounted for by two reorganization

199   mechanisms: the *recombination* of pre-existing chunks (50.4% of the reorganizations, observed

200   for 15 monkeys) and the *concatenation* of two chunks into a single one (49.6% of the

201   reorganizations, observed for 17 monkeys). An illustration of these reorganizations is provided

202   in Figure 2: from Block 1 to Block 3, the first six positions of the sequence were parsed into

203   two equivalent chunks of three positions, indicated by the significant increase in RTs between

204   Position 3 and 4. Starting Block 4, the chunks are recombined with Position 4 being slowly

205   integrated to the first chunk as its difference with Position 3 disappears and a significant

206   increase in RTs is now appearing between Position 4 and 5. The concatenation of two chunks

207   into a single one is also illustrated in Figure 2: while Chunks 2 and 3 were separated by a

208   significant increase in RT between Position 6 and Position 7 until Block 8, this difference

209   disappeared in Block 9, and the two chunks were then grouped into one. Table 1 provides a

210   detailed description of the chunks appearing in Figure 2. For each block of trials, it counts the

211   number of chunks, their respective size, and each occurrence of a concatenation or a

212   recombination. At the group level, Table 2 provides the total number of concatenations and

213   recombinations obtained for each block and for all monkeys[2]. A repeated measure ANOVA

214   with Block and Mechanisms (concatenation vs. recombination) did not reveal any significant

215   effect (all ps > .05).

216

---

[2] See Supplementary material for a description of the patterns of responses for each monkey over blocks. The same information as the one displayed in Figure 2 and Table 1 for the monkey "Atmosphere" is provided for all monkeys from the present experiment (see Figures and Tables S1 to S17). These data clearly show that each monkey displays a specific chunking pattern that follows a specific evolution over blocks of trials.

217   **Table 1**

218   *Summary table of the reorganizations observed throughout the task for one baboon*

219   *(Atmophere).*

	Chunk size			Nb of chunks	Concatenation	Recombinaison
	C1	C2	C3			
Block 1	3	3	3	3	-	-
2	3	3	3	3	-	-
3	3	3	3	3	-	-
4	4	2	3	3	-	1
5	4	2	3	3	-	1
6	4	2	3	3	-	-
7	4	2	3	3	-	-
8	4	4	1	3	-	1
9	4	5	-	2	1	-
10	4	5	-	2	-	-
Total					1	3

220   *Note.* Numbers in columns 2-4 refer respectively to the size of the first (C1), second (C2) and third (C3) chunk (if

221   any) in the sequence.

222
223   **Table 2**

224   *Total number of concatenations and recombinations per block.*

Block	Concatenations	Recombinations
2	11	9
3	8	7
4	6	7
5	6	4
6	8	5
7	3	6
8	1	10
9	8	3
10	5	6
Total	56	57

225

226   **Additional analyses**

227   Our data show that a chunking pattern emerges and evolves over time for each monkey

228   (i.e., Figures S1 to S17). The averaged data on Block 1 indicates that a first chunking pattern

229   rapidly emerges after a few dozen of trials. However, we checked that monkeys produced the

230   same ascending pattern of RTs during their first 10 trials on the repeated sequence than the one

231   we computed from the random baseline trials (see Figure S18). This ensures that our sequences

232   were correctly designed and that chunking mechanisms are rapidly efficient.

233   We also compared the mean response times computed on every trial and averaged over

234   blocks in the random and repeated sequence conditions. The result indicates that chunking

235   mechanisms qualitatively and quantitatively influenced the monkey's behavior by significantly

236   increasing their global efficacy in this visuo-motor task (see Figure S19).

237   **Discussion**

238   Three main findings were obtained in the present study. First, we found that non-human

239   primates spontaneously segmented long sequences into short chunks. Second, with extended

240   practice, chunks became longer and fewer. Third, based on these observations, we assumed this

241   decrease in the number of chunks and this increase in chunks' size was due to two types of

242   reorganizations: the recombination of several preexisting chunks and the concatenation of two

243   distinct chunks into a single one.

244   Our first finding is consistent with previous studies on sequence learning in humans as

245   participants were found to initially segment the sequence into small chunks (e.g., Nissen &

246   Bullemer, 1987; Verwey, 1996; 2001; 2003, Verwey et al., 2002). Chunks typically contained

247   2 to 4 items, sometimes 5 (Sakai et al., 2003), which is similar to what we observed in the first

248   block of the experiment. With extended practice, the mean chunk size could reach up to 8

249   successive positions. However, these large chunks are very rare in our data and in the literature,

250   although Kennerly, Sakai and Rushworth (2004) found in humans a mean chunk size of 7.83

251   items for sequences of 12 items. This difference is likely due to motor constrains that vary

252   across experimental paradigms and may facilitate the development of longer chunks. This first

253   result indicates that chunking is a fundamental mechanism of sequence learning in both human

254 and non-human primates, and that the cognitive system spontaneously forms strong

255 associations between repeated co-occurring events (e.g., Perruchet & Vinter, 1998; 2002; Rey

256 et al., 2019). However, the results also show that there is an initial limit to the number of

257 associations that can be formed successively.

258 Our second finding indicates that extended practice allows the baboons to exceed these

259 limits, as we found that chunks can be reorganized into longer segments. This feature is not

260 systematic as we observed reorganizations in 17 monkeys out of 18. Ramkumar et al. (2016)

261 suggested that to limit the cost of computation, learning new sequences of movements starts

262 with many short chunks. With practice, the execution of short chunks becomes more efficient,

263 which reduces the computation's complexity. This increase in efficiency for short chunks would

264 promote more complex computations leading to the development of longer chunks. However,

265 the precise neural mechanisms supporting these changes over time are still unclear, as are the

266 conditions that promote switching from one pattern to another.

267 Our third main finding is related to the evolution of chunks and the two reorganization

268 mechanisms that allow chunks to become longer: the concatenation and recombination of

269 chunks. Abrahamse et al. (2013) define the concatenation mechanism as the process by which

270 two successive chunks are performed more fluidly and the temporal gap between them

271 decreases. This description is consistent with our findings but, to our knowledge, there is no

272 other report of two successive chunks becoming one and the temporal gap between them

273 disappearing. As for recombinations, they correspond to the emergence of a new segmentation

274 pattern across chunks. As shown in Table 2, the occurrence of concatenation or recombination

275 does not seem related to a specific stage of learning. In both cases, one can assume that these

276 modifications in the chunking pattern is certainly related to the increase in efficiency in the

277 realization of some chunks (Ramkumar et al., 2026). This would then modify the stability of

278 the preexisting chunking pattern and favor either the concatenation of two chunks into a longer

279   one or local rearrangements leading to recombinations and to more functional chunking

280   patterns.

281   Interestingly, current models of chunking only partially account for the reorganizations

282   we found. Most models, like the Competitive Chunking model (Servan-Schreiber & Anderson,

283   1990) or the PARSER model (Perruchet & Vinter, 1998) assume that repeated sequences lead

284   to the formation of chunks that are stored in memory. They also assume that two co-occurring

285   chunks can lead to a new and larger chunk by concatenating two previously smaller chunks.

286   Recombination, though, is not a mechanism implemented in any model of chunking. The

287   present data, therefore, represent a strong challenge for current models of statistical and

288   associative learning.

289   To summarize, these data indicate that when non-verbal non-human animals are

290   repeatedly exposed to a long sequence of 9 elements, associations are formed initially between

291   a limited number of co-occurring elements (i.e., from 2 to 4 elements). With extended practice,

292   these patterns of associations can become longer. The reorganization mechanisms supporting

293   these evolutions are not yet accounted for by any model of associative learning but

294   understanding the dynamics of these mechanisms represents a strong challenge for the future

295   generation of computational models of associative and statistical learning.

296   **Context**

297   We continuously learn and encode the statistical regularities appearing in our

298   environment. Previous empirical and theoretical studies have proposed that chunking and

299   associative learning mechanisms are good candidates to account for our ability to encode

300   sequential regularities. We were interested in the present study in determining the dynamics of

301   these chunking mechanisms when the sequential information was composed of a long, repeated

302   sequence. This sequence was processed by a non-human primate species that does not have

303   access to language recoding abilities. Consistent with previous theoretical accounts, we found

304 that the long sequence was initially parsed into small chunks. After extensive training, some of

305 these small chunks were concatenated leading to longer chunks and to a more efficient

306 processing of the sequence. However, we also found other types of chunk reorganizations, that

307 we called "recombinations", that are not predicted by current models of statistical and

308 associative learning. These results will certainly provide new constraints for elaborating the

309 next generations of computational models accounting for chunking mechanisms.

310 **Acknowledgements**

317 **Open Practices Statements**

318 Data from the experiment are available on Open Science Framework at

319 https://osf.io/xcw95/.

320 **Conflict of Interest**

321 LT, JF, DN and AR declare that they have no conflict of interest.

322 **Animal rights**

323 This research adhered to the applicable French rules for ethical treatment of research

324 animals and received ethical approval from the French Ministry of Education (approval

325 APAFIS#2717-2015111708173794 10 v3).

326

327
328 **References**

329 Abrahamse, E., Ruitenberg, M., De Kleine, E., & Verwey, W. B. (2013). Control of automated

330   behavior: insights from the discrete sequence production task. *Frontiers in Human*

331   *Neuroscience*, *7*, 82.

332 Allen, P. A., & Coyne, A. C. (1988). Age differences in primary organization or processing

333   variability? Part I: An examination of age and primary organization. *Experimental*

334   *Aging Research*, *14*(3), 143-149.

335 Bo, J., Borza, V., & Seidler, R. D. (2009). Age-related declines in visuospatial working memory

336   correlate with deficits in explicit motor sequence learning. *Journal of*

337   *Neurophysiology*, *102*(5), 2744-2754.

338 Bottary, R., Sonni, A., Wright, D., & Spencer, R. M. (2016). Insufficient chunk concatenation

339   may underlie changes in sleep-dependent consolidation of motor sequence learning in

340   older adults. *Learning & Memory*, *23*(9), 455-459.

341 Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, *4*(1), 55-81.

342 Cohen, A., Ivry, R. I., & Keele, S. W. (1990). Attention and structure in sequence

343   learning. *Journal of Experimental Psychology: Learning, Memory, and*

344   *Cognition*, *16*(1), 17.

345 Conway, C. M., & Christiansen, M. H. (2001). Sequential learning in non-human

346   primates. *Trends in Cognitive Sciences*, *5*(12), 539-546.

347 Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their

348   mutual constraints within the human information-processing system. *Psychological*

349   *Bulletin*, *104*(2), 163.

350 Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic*

351   *Bulletin & Review*, *24*(4), 1158-1170.

352 Diedrichsen, J., & Kornysheva, K. (2015). Motor skill learning between selection and

353   execution. *Trends in Cognitive Sciences*, *19*(4), 227-233.

354 Fagot, J., & Bonté, E. (2010). Automated testing of cognitive performance in monkeys: Use of

355   a battery of computerized test systems by a troop of semi-free-ranging baboons (Papio

356   papio). *Behavior Research Methods*, *42*(2), 507-516.

357 Gobet, F., Lane, P. C., Croker, S., Cheng, P. C., Jones, G., Oliver, I., & Pine, J. M. (2001).

358   Chunking mechanisms in human learning. *Trends in Cognitive Sciences*, *5*(6), 236-243.

359 Gobet, F., Lloyd-Kelly, M., & Lane, P. C. (2016). What's in a name? The multiple meanings of

360   "Chunk" and "Chunking". *Frontiers in Psychology*, *7*, 102.

361 Grafton, S. T., Hazeltine, E., & Ivry, R. B. (2002). Motor sequence learning with the

362   nondominant left hand. *Experimental Brain Research*, *146*(3), 369-378.

363 Johnson, N. F. (1970). The role of chunking and organization in the process of recall.

364   In *Psychology of Learning and Motivation* (Vol. 4, pp. 171-247). Academic Press.

365 Kennerley, S. W., Sakai, K., & Rushworth, M. F. S. (2004). Organization of action sequences

366   and the role of the pre-SMA. *Journal of Neurophysiology*, *91*(2), 978-993.

367 Mathy, F., & Feldman, J. (2012). What's magic about magic numbers? Chunking and data

368   compression in short-term memory. *Cognition*, *122*(3), 346-362.

369 Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our

370   capacity for processing information. *Psychological Review*, *63*(2), 81.

371 Minier, L., Fagot, J., & Rey, A. (2016). The temporal dynamics of regularity extraction in non-

372   human primates. *Cognitive Science*, *40*(4), 1019-1030.

373 Nemeth, D., Janacsek, K., Londe, Z., Ullman, M. T., Howard, D. V., & Howard, J. H. (2010).

374   Sleep has no critical role in implicit motor sequence learning in young and old

375   adults. *Experimental Brain Research*, *201*(2), 351-358.

376   Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from

377   performance measures. *Cognitive Psychology*, *19*(1), 1-32.

378   Perruchet, P., & Vinter, A. (1998). PARSER: A model for word segmentation. *Journal of*

379   *Memory and Language*, *39*(2), 246-263.

380   Perruchet, P., & Vinter, A. (2002). The self-organizing consciousness. *Behavioral and Brain*

381   *Sciences*, *25*(3), 297-330.

382   Rey, A., Minier, L., Malassis, R., Bogaerts, L., & Fagot, J. (2019). Regularity extraction across

383   species: Associative learning mechanisms shared by human and non-human

384   primates. *Topics in Cognitive Science*, *11*(3), 573-586.

385   Ramkumar, P., Acuna, D. E., Berniker, M., Grafton, S. T., Turner, R. S., & Kording, K. P.

386   (2016). Chunking as the result of an efficiency computation trade-off. *Nature*

387   *Communications, 7(1)*, 1-11.

388   Sakai, K., Kitaguchi, K., & Hikosaka, O. (2003). Chunking during human visuomotor sequence

389   learning. *Experimental Brain Research*, *152*(2), 229-242.

390   Scarf, D., Smith, C. D., Jaswal, V. K., Magnuson, J. S., & Terrace, H. (2018). Chunky Monkey?

391   The spontaneous temporal chunking of simultaneous chains by Humans (Homo

392   Sapiens) and rhesus monkeys (Macaca mulatta). In A. Ibáñez-Contreras & B.

393   Hernandez-Godinez (Eds), *Studies of Rhesus Monkeys and their Behaviors* (p. 39-58).

394   Nova Science Publishers.

395   Servan-Schreiber, E., & Anderson, J. R. (1990). Learning artificial grammars with competitive

396   chunking. *Journal of Experimental Psychology: Learning, Memory, and*

397   *Cognition*, *16*(4), 592.

398   Song, S., & Cohen, L. (2014). Impact of conscious intent on chunking during motor

399   learning. *Learning & Memory*, *21*(9), 449-451.

400   Terrace H.S. (2002) The Comparative Psychology of Chunking. In S.B. Fountain, M.D.

401   Bunsey, J.H. Danks, & M.K. McBeath (eds), *Animal Cognition and Sequential Behavior*

402   (p. 23-55). Springer, Boston, MA.

403   Verwey, W. B. (1996). Buffer loading and chunking in sequential keypressing. *Journal of*

404   *Experimental Psychology: Human Perception and Performance*, *22*(3), 544.

405   Verwey, W. B. (2001). Concatenating familiar movement sequences: The versatile cognitive

406   processor. *Acta Psychologica*, *106*(1-2), 69-95.

407   Verwey, W. B., & Eikelboom, T. (2003). Evidence for lasting sequence segmentation in the

408   discrete sequence-production task. *Journal of Motor Behavior*, *35*(2), 171-181.

409   Verwey, W. B., Lammens, R., & van Honk, J. (2002). On the role of the SMA in the discrete

410   sequence production task: a TMS study. *Neuropsychologia*, *40*(8), 1268-1276.

411   Willingham, D. B. (1998). A neuropsychological theory of motor skill learning. *Psychological*

412   *Review*, *105*(3), 558.

413   Willingham, D. B. (1999). Implicit motor sequence learning is not purely perceptual. *Memory*

414   *& Cognition*, *27*(3), 561-572.

415   Wymbs, N. F., Bassett, D. S., Mucha, P. J., Porter, M. A., & Grafton, S. T. (2012). Differential

416   recruitment of the sensorimotor putamen and frontoparietal cortex during motor

417   chunking in humans. *Neuron*, *74*(5), 936-946.

418

419 **Appendix A**

420 *Mean response times over the entire group of baboons for each of the 72 possible transitions*

421 *calculated from the 1000 random trials.*

1st position in Transition	2nd position in Transition								
	1	2	3	4	5	6	7	8	9
1	-	519	573	495	482	509	521	497	543
2	569	-	553	513	474	511	523	491	509
3	558	519	-	513	472	488	544	493	512
4	551	517	560	-	464	509	522	482	546
5	549	504	552	501	-	483	535	479	527
6	567	515	546	507	484	-	533	483	511
7	555	504	558	475	463	516	-	484	541
8	554	512	540	485	448	472	512	-	507
9	546	512	540	514	460	464	550	485	-

422 *Note.* All transitions are in milliseconds (ms) and correspond to the time elapsed between the

423 disappearance of the red circle from the 1st position of the Transition and the monkey's touch on the 2nd position

424 of the Transition.

425

426 **Appendix B**

427 *Selected sequences and corresponding mean transition times (based on the baseline*

428 *acquisition, see Appendix A)*

Sequence	Position									Mean transition time (ms)							
	1	2	3	4	5	6	7	8	9	T1	T2	T3	T4	T5	T6	T7	T8
Seq. 1 (N=8)	9	5	6	8	2	4	7	1	3	460	483	483	512	513	522	555	573
Seq. 2 (N=10)	7	5	6	8	2	4	9	1	3	463	483	483	512	513	522	555	573

429