



HAL
open science

The Evolution of Chunks in Sequence Learning

Laure Tosatto, Joël Fagot, Dezso Nemeth, Arnaud Rey

► **To cite this version:**

Laure Tosatto, Joël Fagot, Dezso Nemeth, Arnaud Rey. The Evolution of Chunks in Sequence Learning. *Cognitive Science*, 2022, 46 (4), pp.e13124. 10.1111/cogs.13124 . hal-03640218

HAL Id: hal-03640218

<https://hal.science/hal-03640218>

Submitted on 13 Apr 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29

The Evolution of Chunks in Sequence Learning

Laure Tosatto^{a,b}, Joël Fagot^{a,b,c}, Dezso Nemeth^{d,e,f}, and Arnaud Rey^{a,b}

- a. Aix Marseille Univ, CNRS, LPC, Marseille, France
- b. Aix Marseille Univ, ILCB, Aix-en-Provence, France
- c. Station de Primatologie - Celphedia, CNRS UAR846, Rousset, France
- d. Université Claude Bernard Lyon 1, Lyon Neuroscience Research Center, Lyon, France
- e. Institute of Psychology, Eotvos Lorand University, Budapest, Hungary
- f. Institute of Cognitive Neuroscience and Psychology, Hungarian Academy of Sciences, Budapest, Hungary

Running head: EVOLUTION OF CHUNKS

Correspondence concerning this article should be addressed to Laure Tosatto, Laboratoire de Psychologie Cognitive, CNRS, Aix-Marseille Université, 3, place Victor Hugo, Case D, 13331 Marseille Cedex 03, France. Email: laure.tosatto@univ-amu.fr.

30

Abstract

31 Chunking mechanisms are central to several cognitive processes and notably to the
32 acquisition of visuo-motor sequences. Individuals segment sequences into chunks of items to
33 perform visuo-motor tasks more fluidly, rapidly, and accurately. However, the exact dynamics
34 of chunking processes in the case of extended practice remain unclear. Using an operant
35 conditioning device, eighteen Guinea baboons (*Papio papio*) produced a fixed sequence of nine
36 movements during 1,000 trials by pointing to a moving target on a touch screen. Response times
37 analyses revealed a specific chunking pattern of the sequence for each baboon. More
38 importantly, we found that these patterns evolved during the course of the experiment, with
39 chunks becoming progressively fewer and longer. We identified two chunk reorganization
40 mechanisms: the *recombination* of preexisting chunks and the *concatenation* of two distinct
41 chunks into a single one. These results provide new evidence on chunking mechanisms in
42 sequence learning and challenge current models of associative and statistical learning.

43

44 *Keywords:* chunking, sequence learning, statistical learning, associative learning, non-human
45 primate

46 **1. Introduction**

47 A key mechanism allowing our cognitive system to compress information and increase
48 short term memory capacity is the formation of chunks (Mathy & Feldman, 2012; Miller, 1956).
49 Chunking is defined as the process of associating and grouping several items together into a
50 single processing unit (Gobet & al, 2001; Gobet, Lloyd-Kelly and Lane, 2016). In coherence
51 with the limits encountered by our cognitive system (Cowan, 1988, 2017), the storage capacity
52 of the chunks themselves seems limited (3 or 4 items per chunk, Allen & Coyne, 1988; Chase
53 & Simon, 1973; Johnson, 1970). In the field of perceptual-motor learning, chunking has been
54 considered as the main motor sequence integration mechanism (Diedrichsen & Kornysheva,
55 2015; Wymbs et al., 2012).

56 Perceptual-motor sequence learning is commonly described as the process by which a
57 sequence of movements is acquired and executed with increased speed and accuracy
58 (Willingham, 1998). Current sequence learning paradigms use a sequential button-press task:
59 participants are presented with a series of visual stimuli organized in a sequence and are asked
60 to press a corresponding response button (e.g., Cohen, Ivry, & Keele, 1990; Nemeth et al., 2010;
61 Verwey, 2001). A faster and more accurate performance over time reflects learning of the
62 sequence. Typically, researchers have reported robust effects of sequence's length on
63 participants' response times (RTs): for short sequences (<4 keypresses), a decrease in the
64 successive RTs can be observed when the sequence is learned; for longer sequences though, an
65 initial decrease in RTs for the first 3 to 4 keypresses is followed by a longer RT on the next
66 position. Then, RTs start decreasing again for the next 3 to 4 keypresses (e.g., Bo & Seidler,
67 2009; Verwey & Eikelboom, 2003). Similar results have been observed in non-human
68 individuals (Terrace, 2002; Scarf et al., 2018). This phenomenon has been interpreted as the
69 long sequence being spontaneously segmented into shorter motor chunks reflecting the

70 sequence organization in memory (Sakai et al., 2003). Long temporal gaps between responses
71 are assumed to mark chunk boundaries (Abrahamse et al., 2013; Bottary et al., 2016).

72 In the case of extended practice, several studies have investigated the evolution of
73 chunks and have reported a decrease over time of the number of chunks (e.g., Acuna et al.,
74 2015; Bera et al., 2021; Ramkumar et al., 2016). One mechanism proposed to account for this
75 phenomenon is the *concatenation* of chunks (Verwey, 2001; Wright et al., 2010).
76 Concatenation is described as independent chunks being executed more fluidly with practice
77 and with a decrease of the temporal gap between them leading to a single and longer chunk
78 (Abrahamse et al., 2013). However, the precise nature of the mechanisms allowing RTs to
79 decrease and independent chunks to merge into a larger functional unit remain unclear. The
80 present study aims at collecting new evidence on the mechanisms underlying the evolution of
81 chunking mechanisms in a perceptual-motor sequence learning task.

82 One characteristic of previous sequence learning studies is the pattern of transition
83 probabilities (TPs) between items of the sequence. In an experimental set up where only a few
84 response buttons are available to the participant (e.g., Grafton, et al., 2002; Verwey, 2001;
85 Willingham, 1999), some stimuli are necessarily presented multiple times within the same
86 sequence. For instance, in a sequence such as A-B-A-C-D-B-C, A is either followed by B or C,
87 therefore the probability of B given A is 0.5, whereas D is always followed by B thus the
88 probability of B given D is 1. This heterogeneity of TPs may constrain the strength of the
89 connections formed between the successive elements of the sequence and affect the resulting
90 chunking pattern (Aslin, 2017; Lazartigues et al. 2021).

91 Another characteristic of previous studies is, apart from a few exceptions (e.g., Conway
92 & Christiansen, 2001), that most sequence learning studies have been conducted with human
93 participants. However, even if implicit instructions are provided, humans may use their inner
94 language and develop strategies to perform the task (Rey et al., 2019). This may affect both

95 their performance and their chunking processes, and it is hard to tease apart these verbal and
96 explicit influences from associative and chunking mechanisms.

97 To address these two concerns, we used the serial pointing task developed by Minier et
98 al. (2016). In this task, participating monkeys had to touch a target red circle that could appear
99 at 9 possible locations, defined by a 3 x 3 matrix of crosses displayed on a touch screen. They
100 were repeatedly exposed to the same sequence of 9 locations that always appeared in the same
101 order upon a thousand trials. The repeated sequence was composed of the nine possible
102 locations presented in a specific order (e.g., 9-5-6-8-2-4-7-1-3) and leading to a transitional
103 probability between locations equal to 1. RTs for each location were recorded, providing us
104 with detailed information about the temporal dynamics of sequence processing. This extended
105 practice on a long single repeated sequence allowed us to collect new evidence on the formation
106 and the evolution of chunks over time.

107 **2. Method**

108 *2.1. Participants*

109 Thirteen female and five male Guinea baboons (*Papio papio*, age range 2.8—23.7 years)
110 from the CNRS primate facility in Rousset (France), living in a social group of 25 individuals,
111 were tested in this study. Water was provided *ad libitum* during the test, and the monkeys
112 received their normal food ration of fruits every day at 5 PM. For practical reasons, we stopped
113 the experiment after 18 monkeys completed all scheduled trials.

114 *2.2. Materials*

115 *2.2.1. Apparatus*

116 This experiment was conducted with a computer-learning device based on the voluntary
117 participation of baboons (for details, see Fagot & Bonté, 2010). Baboons implanted with a RFID
118 microchip had free access to 10 automatic operant conditioning learning devices equipped with
119 touch screens. Each time a monkey entered a test chamber, it was identified by its microchip,

145 mean RTs for each transition (e.g., from position 1 to 9), leading to a 9x9 matrix of mean RTs
146 calculated over the entire group of baboons (see Appendix A). This baseline measure
147 corresponds to the mean time baboons need to move their hand from one location to another,
148 when the transition is unpredictable (i.e., random).

149 Based on these baseline measures, we designed two sequences of nine serial locations
150 for which each transition T was faster (or equally fast) to produce than the next one (i.e.,
151 $T_1 \leq T_2 \leq \dots \leq T_8$; see Appendix B). This way, a decrease in RT for a given transition can be
152 interpreted as the anticipation (or learning) of that position from the previous one.

153 *2.3. Procedure*

154 Baboons were either presented with Sequence 1 (N=8 baboons) or Sequence 2 (N=10
155 baboons) and had to produce it for 1000 successive trials. RTs for each position of the sequence
156 were recorded for all the trials.

157 **3. Results**

158 On average, baboons required 2.9 days to complete the 1,000 trials, with a mean of
159 339.7 trials per day and a mean accuracy level of 92.5% ($SD=6.2\%$). Incorrect trials were
160 removed from the dataset (7.8%). A recursive trimming procedure excluded RTs greater or
161 smaller than 2.5 standard deviations from the participant's mean for each of the nine possible
162 positions (24.4%)¹. RTs for each of the nine positions and for the 1,000 trials were divided into
163 10 Blocks of 100 trials.

164 *3.1. General sequence learning*

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).

165 Learning of the sequence was estimated on mean RTs by a repeated measures one-way
166 ANOVA with Block (1 to 10) as the within factor. The effect of Block was highly significant
167 (Block 1, $M=452.8$, $SD=45.3$; Block 10, $M=400.1$, $SD=56.3$), $F(1,9)=30.43$, $p<.001$, $\eta^2=0.642$,
168 indicating that mean RTs decreased throughout the blocks of trials and that monkeys learned
169 the sequence.

170 3.2. *Chunking of the sequence*

171 To study the chunking pattern of the sequence by monkeys, we adopted the following
172 reasoning. We considered successive positions A and B to be part of the same chunk as long as
173 the transition time from one position to the next did not correspond to a significant increase in
174 RT. Therefore, if $RT_A < RT_B$ then A and B do not belong to the same chunk, and the AB transition
175 marks a chunk boundary (Kennerly et al., 2004). This reasoning was applied on each Block and
176 for each monkey. Statistical significance was assessed through paired-sample t-tests for each
177 pair of successive positions. Each time the RT of a pair's second position was significantly
178 higher than the first position, it marked a chunk boundary. Figure 2 illustrates this method on
179 the performances of one monkey.

180 < insert Figure 2 here >

181 3.3. *Evolution of chunks*

182 With this method, we were able to quantify the number of chunks and their average size
183 produced on each block by each monkey. For example, the monkey from Figure 2 produced 3
184 chunks and an average chunk size of 3 in Block 1. In Block 10, the number of chunks was 2
185 and the average chunk size was 4.5. Figure 3 reports the average number of chunks and chunk
186 size for all monkeys across the 10 blocks of the experiment. For example, we found that the
187 mean chunk size was 2.2 positions ($CI=0.31$; $Min=1$; $Max=5$) in the first block and 3.375
188 ($CI=0.53$; $Min=1$; $Max=8$) in the last block. Two repeated measures one-way ANOVA were
189 conducted to test the effect of Block on the mean number of chunks and the average chunk size.

190 This analysis revealed that the number of chunks significantly decreased across blocks,
191 $F(1,9)=9.421, p<.001, \eta^2=.357$, and that the average chunk size significantly increased across
192 blocks, $F(1,9)=4.794, p<.01, \eta^2=.22$.

193 The evolution of these two indicators can be accounted for by two reorganization
194 mechanisms: the *recombination* of pre-existing chunks (50.4% of the reorganizations, observed
195 for 15 monkeys) and the *concatenation* of two chunks into a single one (49.6% of the
196 reorganizations, observed for 17 monkeys). An illustration of these reorganizations is provided
197 in Figure 2: from Block 1 to Block 3, the first six positions of the sequence were parsed into
198 two equivalent chunks of three positions, indicated by the significant increase in RTs between
199 Position 3 and 4. Starting Block 4, the chunks are recombined with Position 4 being slowly
200 integrated to the first chunk as its difference with Position 3 disappears and a significant
201 increase in RTs is now appearing between Position 4 and 5. The concatenation of two chunks
202 into a single one is also illustrated in Figure 2: while Chunks 2 and 3 were separated by a
203 significant increase in RT between Position 6 and Position 7 until Block 8, this difference
204 disappeared in Block 9, and the two chunks were then grouped into one. Table 1 provides a
205 detailed description of the chunks appearing in Figure 2. For each block of trials, it counts the
206 number of chunks, their respective size, and each occurrence of a concatenation or a
207 recombination. At the group level, Table 2 provides the total number of concatenations and
208 recombinations obtained for each block and for all monkeys². A repeated measure ANOVA
209 with Block and Mechanisms (concatenation vs. recombination) did not reveal any significant
210 effect (all $p_s > .05$).

211 < insert Table 1 & 2 here >

212 3.4. Additional analyses

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.

213 Our data show that a chunking pattern emerges and evolves over time for each monkey
214 (i.e., Figures S1 to S17). The averaged data on Block 1 indicates that a first chunking pattern
215 rapidly emerges after a few dozen of trials. However, we checked that monkeys produced the
216 same ascending pattern of RTs during their first 10 trials on the repeated sequence than the one
217 we computed from the random baseline trials (see Figure S18). This ensures that our sequences
218 were correctly designed and that chunking mechanisms are rapidly efficient.

219 We also compared the mean response times computed on every trial and averaged over
220 blocks in the random and repeated sequence conditions. The result indicates that chunking
221 mechanisms qualitatively and quantitatively influenced the monkey's behavior by significantly
222 increasing their global efficacy in this visuo-motor task (see Figure S19).

223 **4. Discussion**

224 Three main findings were obtained in the present study. First, we found that non-human
225 primates spontaneously segmented long sequences into short chunks. Second, with extended
226 practice, chunks became longer and fewer. Third, based on these observations, we found this
227 decrease in the number of chunks and this increase in chunks' size was due to two types of
228 reorganizations: the recombination of several preexisting chunks and the concatenation of two
229 distinct chunks into a single one.

230 Our first finding is consistent with previous studies on sequence learning in humans as
231 participants were found to initially segment the sequence into small chunks (e.g., Nissen &
232 Bullemer, 1987; Verwey, 1996; 2001; 2003, Verwey et al., 2002). Chunks typically contained
233 2 to 4 items, sometimes 5 (Sakai et al., 2003), which is similar to what we observed in the first
234 block of the experiment. With extended practice, the mean chunk size could reach up to 8
235 successive positions. However, these large chunks are very rare in our data and in the literature,
236 although Kennerly, Sakai and Rushworth (2004) found in humans a mean chunk size of 7.83
237 items for sequences of 12 items. This difference is likely due to motor constraints that vary

238 across experimental paradigms and may facilitate the development of longer chunks. This first
239 result indicates that chunking is a fundamental mechanism of sequence learning in both human
240 and non-human primates, and that the cognitive system spontaneously forms strong
241 associations between repeated co-occurring events (e.g., Perruchet & Vinter, 1998; 2002; Rey
242 et al., 2019). However, the results also show that there is an initial limit to the number of
243 associations that can be formed successively.

244 Our second finding is in line with previous research and indicates that extended practice
245 allows the baboons to exceed these limits, as we found that chunks can be reorganized into
246 longer segments. This feature is robust and we observed reorganizations in 17 monkeys out of
247 18. Ramkumar et al. (2016) suggested that to limit the cost of computation, learning new
248 sequences of movements starts with many short chunks. With practice, the execution of short
249 chunks becomes more efficient, which reduces the computation's complexity. This increase in
250 efficiency for short chunks would promote more complex computations leading to the
251 development of longer chunks. However, the precise neural mechanisms supporting these
252 changes over time are still unclear, as are the conditions that promote switching from one
253 pattern to another.

254 Our third main finding is related to the evolution of chunks and the two reorganization
255 mechanisms that allow chunks to become longer: the concatenation and recombination of
256 chunks. Abrahamse et al. (2013) define the concatenation mechanism as the process by which
257 two successive chunks are performed more fluidly and the temporal gap between them
258 decreases. This description is consistent with some of our findings, but concatenations were not
259 the only form of reorganization we observed. To our knowledge, there is no other study
260 reporting the second type of reorganization we found, i.e., recombinations. Indeed,
261 recombinations differs from concatenations by corresponding to the emergence of a new
262 segmentation pattern across chunks. As shown in Table 2, the occurrence of a concatenation or

263 a recombination does not seem related to a specific stage of learning as it seems to be evenly
264 distributed throughout the experiment.

265 The absence of differences in the distribution of recombinations and concatenations
266 throughout learning can be surprising considering that, in earlier stages of learning, chunks are
267 less stable, and one would expect that recombinations would occur more frequently at the
268 beginning of the task. Conversely, as the sequence is learned, chunks stability should increase,
269 and one would expect more concatenations to occur. However, as indicated by the linear
270 decrease in mean RTs (see Figure S19), it is possible that learning of the sequence was not over
271 even after 1,000 trials, and that baboons did not yet reached their optimal performance.
272 Therefore, an increase of the number of concatenations over the number of recombinations
273 could occur during later stages of learning that were not observed here.

274 The present findings therefore indicate that baboons initially organize their pattern of
275 responses by segmenting the sequence of nine touches into small chunks (i.e., from 2 to 4
276 elements). With extended practice, the chunking pattern either remain the same or is
277 reorganized be either concatenating or recombining chunks. We identify three parameters that
278 are likely determining the dynamical evolution of a chunking pattern. The first parameter
279 corresponds to the *internal stability* of chunks. Consider, for example, Chunk 3 in Figure 2
280 from Block 2 to 4. The mean RTs for Position 7, 8 and 9 of Chunk 3 are extremely far apart,
281 indicating that this chunking pattern is highly stable. To operationalize this notion of stability,
282 one could compute the standard deviation among these three RTs and find a large value (e.g.,
283 the standard deviation in Block 2 for Chunk 3 is 68.8). Alternatively, Chunk 1 in Block 3
284 certainly has a weaker internal stability due to the small differences between the five RTs
285 composing this chunk (the standard deviation is here equal to 19.7). Internal stability of chunks
286 therefore likely determines the survival of a chunking pattern or its reorganization.

287 The second parameter corresponds to the *stability of chunk boundaries* that can be
288 estimated by computing the difference between the last RT of one chunk and the first RT of the
289 next chunk. Consider, for example, the boundary between Chunk 2 and 3 in Block 2 of Figure
290 2, the difference between the last RT of Chunk 2 (i.e., 378.5 ms) and the first RT of Chunk 3
291 (i.e., 501.1 ms) is equal to 122.6 ms. Here the difference is huge, and this boundary does not
292 change until Block 8.

293 The third parameter is related to the *increased efficiency* in the realization of chunks
294 during extended practice (an idea proposed by Ramkumar et al., 2016). Indeed, the reason why
295 the boundary between Chunk 2 and 3 decreases from Block 1 to 7 is likely due the increase in
296 efficiency in the realization of Chunk 3 (leading to faster RTs) and consequently, to the smaller
297 difference between the last RT of Chunk 2 and the first RT of Chunk 3.

298 The present data therefore suggest that the evolution of chunks in sequence learning
299 probably depends on these three parameters: 1) the internal stability of chunks; 2) the stability
300 of chunk boundaries; and 3) the increased efficiency in the realization of chunks during
301 extended practice. Concatenations will occur when two chunks have a strong internal stability
302 and when the realization of the second chunk will become more efficient, leading to a lower
303 difference between the last RT of the first chunk and the first RT of the second chunk and
304 therefore, a lower stability of chunks boundary. Recombinations are likely to be observed when
305 the internal stability of chunks is low and when there is a low stability of chunks boundary.
306 Finally, the chunking pattern is likely to remain the same when the chunk stability and the
307 chunk boundaries are high.

308 Note that the fact that some chunks are being recombined throughout the task also
309 challenges the idea that there is always a hierarchical organization in sequence learning (Restle,
310 1970; Yokoi & Diedrichsen, 2019) and that the reorganizations only operate by concatenating

311 preexisting chunks. Indeed, chunks appear to be more flexible than previously thought and
312 linear associations between adjacent elements does not seem to be the only mechanism at work.

313 Interestingly, current models of chunking only partially account for the reorganizations
314 we found. Most models, like the Competitive Chunking model (Servan-Schreiber & Anderson,
315 1990) or the PARSER model (Perruchet & Vinter, 1998) assume that repeated sequences lead
316 to the formation of chunks that are stored in memory. They also assume that two co-occurring
317 chunks can lead to a new and larger chunk by concatenating two previously smaller chunks.
318 Recombination, though, is not a mechanism implemented in any model of chunking. The
319 present data, therefore, represent a strong challenge for current models of statistical and
320 associative learning.

321 To summarize, these data indicate that when non-verbal non-human animals are
322 repeatedly exposed to a long sequence of 9 elements, associations are formed initially between
323 a limited number of co-occurring elements (i.e., from 2 to 4 elements). With extended practice,
324 these patterns of associations can become longer. The reorganization mechanisms supporting
325 these evolutions are not yet accounted for by any model of associative learning but
326 understanding the dynamics of these mechanisms represents a strong challenge for the future
327 generation of computational models of associative and statistical learning.

328

329

330 Acknowledgements

331 This work was supported by the BLRI Labex (ANR-11-LABX-0036), Institut
332 Convergence ILCB (ANR-16-CONV-0002), the CHUNKED ANR project (#ANR-17-CE28-
333 0013-02) and IDEXLYON Fellowship of the University of Lyon as part of the Programme
334 Investissements d’Avenir (ANR-16-IDEX-0005). The funders had no role in study design, data
335 collection and analysis, decision to publish, or preparation of the manuscript. We are grateful
336 to Laura Ordonez-Magro for her helpful comments on a previous version of the manuscript.

337 Open Practices Statements

338 Data from the experiment are available on Open Science Framework at
339 https://osf.io/xcw95/?view_only=73f8cc59abe343948a6c6ea090d6dbe2.

340 Conflict of Interest

341 The authors declare that they have no conflict of interest.

342 Animal rights

343 This research adhered to the applicable French rules for ethical treatment of research
344 animals and received ethical approval from the French Ministry of Education (approval
345 APAFIS#2717-2015111708173794 10 v3).

346

347

348 **References**

- 349 Abrahamse, E., Ruitenberg, M., De Kleine, E., & Verwey, W. B. (2013). Control of automated
350 behavior: insights from the discrete sequence production task. *Frontiers in Human*
351 *Neuroscience*, 7, 82.
- 352 Acuna, D. E., Wymbs, N. F., Reynolds, C. A., Picard, N., Turner, R. S., Strick, P. L., ... &
353 Kording, K. P. (2014). Multifaceted aspects of chunking enable robust
354 algorithms. *Journal of neurophysiology*, 112(8), 1849-1856.
- 355 Allen, P. A., & Coyne, A. C. (1988). Age differences in primary organization or processing
356 variability? Part I: An examination of age and primary organization. *Experimental*
357 *Aging Research*, 14(3), 143-149.
- 358 Aslin, R. N. (2017). Statistical learning: a powerful mechanism that operates by mere
359 exposure. *Wiley Interdisciplinary Reviews: Cognitive Science*, 8(1-2), e1373.
- 360 Bera, K., Shukla, A., & Bapi, R. S. (2021). Motor chunking in internally guided
361 sequencing. *Brain Sciences*, 11(3), 292.
- 362 Bo, J., Borza, V., & Seidler, R. D. (2009). Age-related declines in visuospatial working memory
363 correlate with deficits in explicit motor sequence learning. *Journal of*
364 *Neurophysiology*, 102(5), 2744-2754.
- 365 Bottary, R., Sonni, A., Wright, D., & Spencer, R. M. (2016). Insufficient chunk concatenation
366 may underlie changes in sleep-dependent consolidation of motor sequence learning in
367 older adults. *Learning & Memory*, 23(9), 455-459.
- 368 Chase, W. G., & Simon, H. A. (1973). Perception in chess. *Cognitive Psychology*, 4(1), 55-81.
- 369 Cohen, A., Ivry, R. I., & Keele, S. W. (1990). Attention and structure in sequence
370 learning. *Journal of Experimental Psychology: Learning, Memory, and*
371 *Cognition*, 16(1), 17.

- 372 Conway, C. M., & Christiansen, M. H. (2001). Sequential learning in non-human
373 primates. *Trends in Cognitive Sciences*, 5(12), 539-546.
- 374 Cowan, N. (1988). Evolving conceptions of memory storage, selective attention, and their
375 mutual constraints within the human information-processing system. *Psychological*
376 *Bulletin*, 104(2), 163.
- 377 Cowan, N. (2017). The many faces of working memory and short-term storage. *Psychonomic*
378 *Bulletin & Review*, 24(4), 1158-1170.
- 379 Diedrichsen, J., & Kornysheva, K. (2015). Motor skill learning between selection and
380 execution. *Trends in Cognitive Sciences*, 19(4), 227-233.
- 381 Fagot, J., & Bonté, E. (2010). Automated testing of cognitive performance in monkeys: Use of
382 a battery of computerized test systems by a troop of semi-free-ranging baboons (*Papio*
383 *papio*). *Behavior Research Methods*, 42(2), 507-516.
- 384 Gobet, F., Lane, P. C., Croker, S., Cheng, P. C., Jones, G., Oliver, I., & Pine, J. M. (2001).
385 Chunking mechanisms in human learning. *Trends in Cognitive Sciences*, 5(6), 236-243.
- 386 Gobet, F., Lloyd-Kelly, M., & Lane, P. C. (2016). What's in a name? The multiple meanings of
387 "Chunk" and "Chunking". *Frontiers in Psychology*, 7, 102.
- 388 Grafton, S. T., Hazeltine, E., & Ivry, R. B. (2002). Motor sequence learning with the
389 nondominant left hand. *Experimental Brain Research*, 146(3), 369-378.
- 390 Johnson, N. F. (1970). The role of chunking and organization in the process of recall.
391 In *Psychology of Learning and Motivation* (Vol. 4, pp. 171-247). Academic Press.
- 392 Kennerley, S. W., Sakai, K., & Rushworth, M. F. S. (2004). Organization of action sequences
393 and the role of the pre-SMA. *Journal of Neurophysiology*, 91(2), 978-993.
- 394 Lazartigues, L., Mathy, F., & Lavigne, F. (2021). Statistical learning of unbalanced exclusive-
395 or temporal sequences in humans. *Plos one*, 16(2), e0246826.

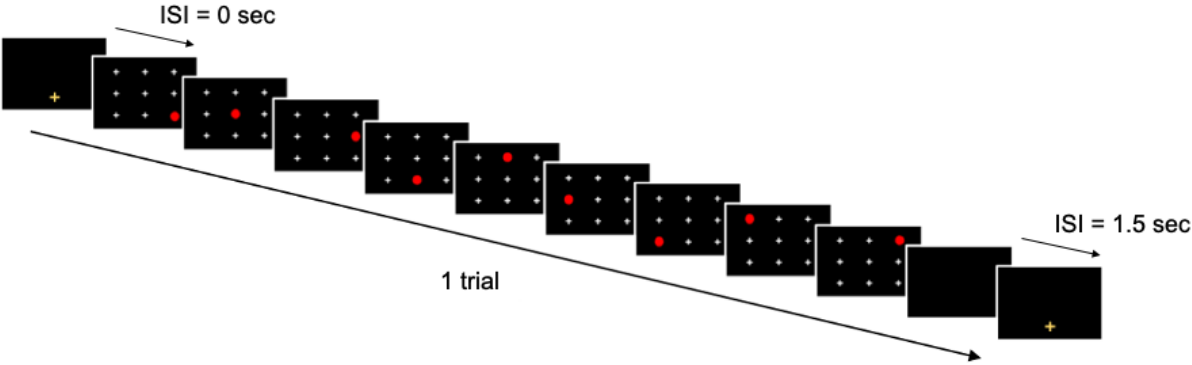
- 396 Mathy, F., & Feldman, J. (2012). What's magic about magic numbers? Chunking and data
397 compression in short-term memory. *Cognition*, *122*(3), 346-362.
- 398 Miller, G. A. (1956). The magical number seven, plus or minus two: Some limits on our
399 capacity for processing information. *Psychological Review*, *63*(2), 81.
- 400 Minier, L., Fagot, J., & Rey, A. (2016). The temporal dynamics of regularity extraction in non-
401 human primates. *Cognitive Science*, *40*(4), 1019-1030.
- 402 Nemeth, D., Janacek, K., Londe, Z., Ullman, M. T., Howard, D. V., & Howard, J. H. (2010).
403 Sleep has no critical role in implicit motor sequence learning in young and old
404 adults. *Experimental Brain Research*, *201*(2), 351-358.
- 405 Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from
406 performance measures. *Cognitive Psychology*, *19*(1), 1-32.
- 407 Perruchet, P., & Vinter, A. (1998). PARSER: A model for word segmentation. *Journal of*
408 *Memory and Language*, *39*(2), 246-263.
- 409 Perruchet, P., & Vinter, A. (2002). The self-organizing consciousness. *Behavioral and Brain*
410 *Sciences*, *25*(3), 297-330.
- 411 Restle, F. (1970). Theory of serial pattern learning: Structural trees. *Psychological*
412 *Review*, *77*(6), 481.
- 413 Rey, A., Minier, L., Malassis, R., Bogaerts, L., & Fagot, J. (2019). Regularity extraction across
414 species: Associative learning mechanisms shared by human and non-human
415 primates. *Topics in Cognitive Science*, *11*(3), 573-586.
- 416 Ramkumar, P., Acuna, D. E., Berniker, M., Grafton, S. T., Turner, R. S., & Kording, K. P.
417 (2016). Chunking as the result of an efficiency computation trade-off. *Nature*
418 *Communications*, *7*(1), 1-11.
- 419 Sakai, K., Kitaguchi, K., & Hikosaka, O. (2003). Chunking during human visuomotor sequence
420 learning. *Experimental Brain Research*, *152*(2), 229-242.

- 421 Scarf, D., Smith, C. D., Jaswal, V. K., Magnuson, J. S., & Terrace, H. (2018). Chunky Monkey?
422 The spontaneous temporal chunking of simultaneous chains by Humans (Homo
423 Sapiens) and rhesus monkeys (*Macaca mulatta*). In A. Ibáñez-Contreras & B.
424 Hernandez-Godinez (Eds), *Studies of Rhesus Monkeys and their Behaviors* (p. 39-58).
425 Nova Science Publishers.
- 426 Servan-Schreiber, E., & Anderson, J. R. (1990). Learning artificial grammars with competitive
427 chunking. *Journal of Experimental Psychology: Learning, Memory, and*
428 *Cognition*, 16(4), 592.
- 429 Song, S., & Cohen, L. (2014). Impact of conscious intent on chunking during motor
430 learning. *Learning & Memory*, 21(9), 449-451.
- 431 Terrace H.S. (2002) The Comparative Psychology of Chunking. In S.B. Fountain, M.D.
432 Bunsey, J.H. Danks, & M.K. McBeath (eds), *Animal Cognition and Sequential Behavior*
433 (p. 23-55). Springer, Boston, MA.
- 434 Verwey, W. B. (1996). Buffer loading and chunking in sequential keypressing. *Journal of*
435 *Experimental Psychology: Human Perception and Performance*, 22(3), 544.
- 436 Verwey, W. B. (2001). Concatenating familiar movement sequences: The versatile cognitive
437 processor. *Acta Psychologica*, 106(1-2), 69-95.
- 438 Verwey, W. B., & Eikelboom, T. (2003). Evidence for lasting sequence segmentation in the
439 discrete sequence-production task. *Journal of Motor Behavior*, 35(2), 171-181.
- 440 Verwey, W. B., Lammens, R., & van Honk, J. (2002). On the role of the SMA in the discrete
441 sequence production task: a TMS study. *Neuropsychologia*, 40(8), 1268-1276.
- 442 Willingham, D. B. (1998). A neuropsychological theory of motor skill learning. *Psychological*
443 *Review*, 105(3), 558.
- 444 Willingham, D. B. (1999). Implicit motor sequence learning is not purely perceptual. *Memory*
445 *& Cognition*, 27(3), 561-572.

- 446 Wright, D. L., Rhee, J. H., & Vaculin, A. (2010). Offline improvement during motor sequence
447 learning is not restricted to developing motor chunks. *Journal of motor behavior*, 42(5),
448 317-324.
- 449 Wymbs, N. F., Bassett, D. S., Mucha, P. J., Porter, M. A., & Grafton, S. T. (2012). Differential
450 recruitment of the sensorimotor putamen and frontoparietal cortex during motor
451 chunking in humans. *Neuron*, 74(5), 936-946.
- 452 Yokoi, A., & Diedrichsen, J. (2019). Neural organization of hierarchical motor sequence
453 representations in the human neocortex. *Neuron*, 103(6), 1178-1190.
- 454
- 455

456 **Figure 1**

457 *Experimental display and stimuli presentation*

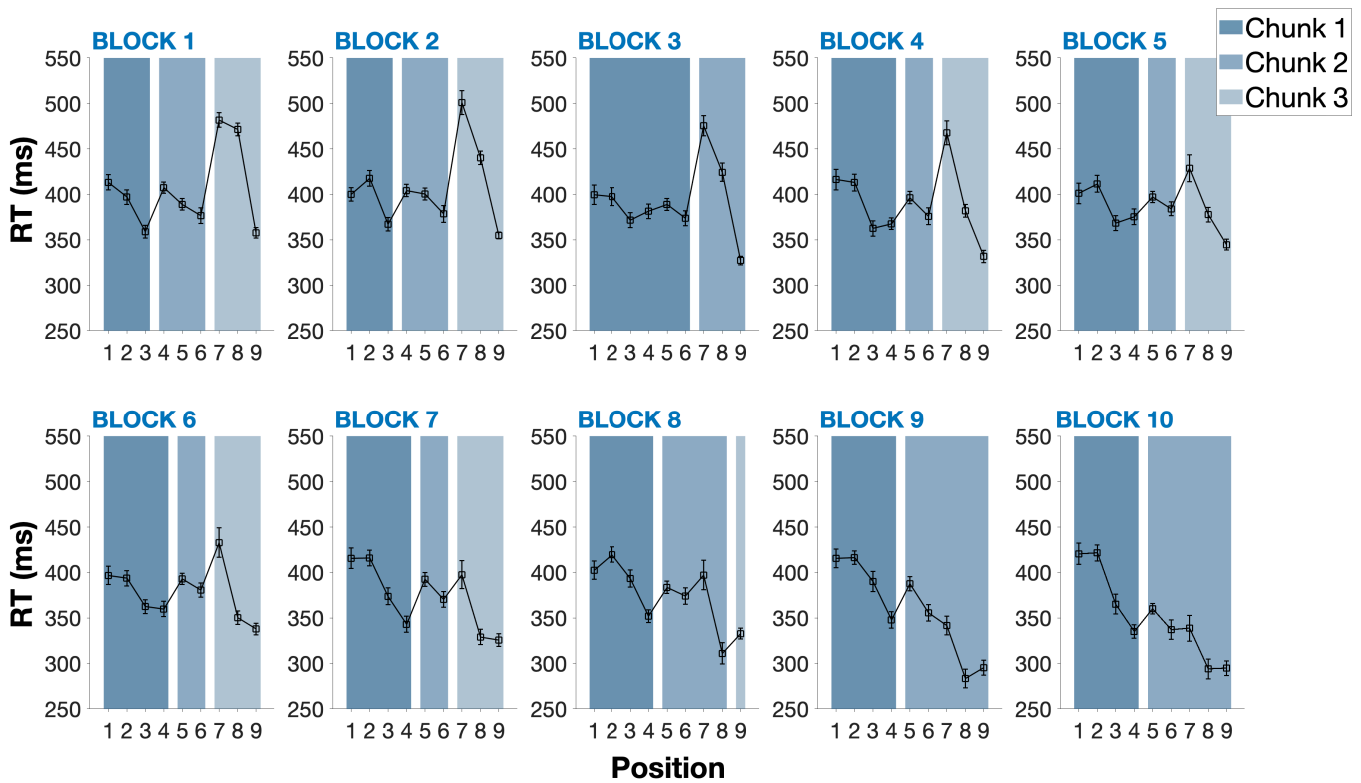


458

459

460 **Figure 2**

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



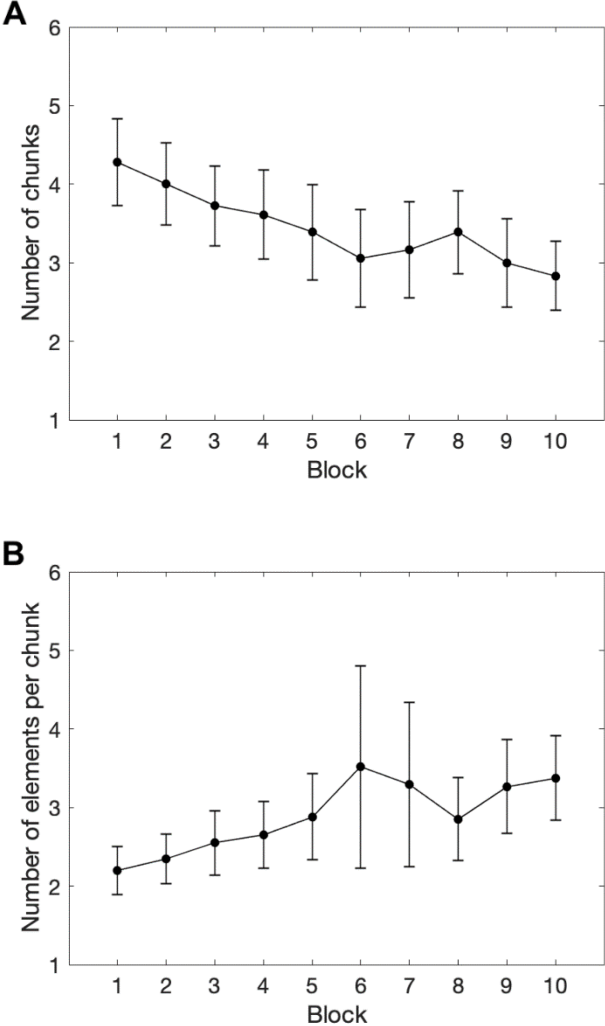
462

463 *Note.* Mean RT per position across the 10 blocks of trials for one baboon (Atmosphere) showing the evolution of
 464 the chunking pattern. This individual initially parses the sequence into three chunks of three positions in the first
 465 three blocks of trials. From Block 7 to 8, the 2nd and 3rd chunks (having a size of 2 and 3, respectively) are
 466 recombined into two chunks of size 4 and 1, respectively. From Block 8 to 9, the 2nd and 3rd chunks are
 467 concatenated into one single chunk of size 5. Error bars represent 95% confidence intervals.

468

469 **Figure 3**

470 *Evolution of chunks across blocks*



471
472 *Note.* A. Mean number of chunks per block. B. Mean chunk size (i.e. number of elements per chunk) per block
473 (error bars represent 95%confidence intervals).
474

475 **Table 1**476 *Summary table of the reorganizations observed throughout the task for one baboon*477 *(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

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

479 any) in the sequence.

480

481 **Table 2**482 *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

483

484

485

Appendix A486 *Mean response times over the entire group of baboons for each of the 72 possible transitions*487 *calculated from the 1000 random trials.*

1 st position in Transition	2 nd 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	-

488 *Note.* All transitions are in milliseconds (ms) and correspond to the time elapsed between the489 disappearance of the red circle from the 1st position of the Transition and the monkey's touch on the 2nd position

490 of the Transition.

491

492

Appendix B493 *Selected sequences and corresponding mean transition times (based on the baseline*494 *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

495