Chunk Concatenation Evolves with Practice and Sleep-Related Enhancement Consolidation in a Complex Arm Movement Sequence

by
Klaus Blischke[1], Andreas Malangré[1]

This paper addresses the notion of chunk concatenation being associated with sleep-related enhancement consolidation of motor sequence memory, thereby essentially contributing to improvements in sequence execution speed. To this end, element movement times of a multi-joint arm movement sequence incorporated in a recent study by Malangré et al. (2014) were reanalyzed. As sequence elements differed with respect to movement distance, element movement times had to be purged from differences solely due to varying trajectory lengths. This was done by dividing each element movement time per subject and trial block by the respective "reference movement time" collected from subjects who had extensively practiced each sequence element in isolation. Any differences in these "relative element movement times" were supposed to reflect element-specific "production costs" imposed solely by the sequence context. Across all subjects non-idiosyncratic, lasting sequence segmentation was shown, and four possible concatenation points (i.e. transition points between successive chunks) within the original arm movement sequence were identified. Based on theoretical suppositions derived from previous work with the discrete sequence production task and the dual processor model (Abrahamse et al., 2013), significantly larger improvements in transition speed occurring at these four concatenation points as compared to the five fastest transition positions within the sequence (associated with mere element execution) were assumed to indicate increased chunk concatenation. As a result, chunk concatenation was shown to proceed during acquisition with physical practice, and, most importantly, to significantly progress some more during retention following a night of sleep, but not during a waking interval.

Key words: motor sequence learning, memory consolidation, offline learning, sleep, gross motor skill.

Introduction

In the motor domain, numerous studies have demonstrated by now that sleep after initial training of motor skills, in the absence of any further physical practice, can produce significant improvements in performance at later retesting. Although in most of these studies fine motor skills like sequential-finger-tapping were used for a criterion task (e.g. Fischer et al., 2005; Walker, 2005; Doyon et al., 2009), recently these findings have also successfully been extended to gross motor tasks involving multi-joint limb movements of the upper as well as of the lower extremities (Al-Sharman and Siengsukon, 2013; Genzel et al., 2012; Kempler and Richmond, 2012; Malangré et al., 2014; Morita et al., 2012).

This phenomenon, usually referred to as "offline learning" or "enhancement consolidation", has been associated with an "active system consolidation" process (Born and Wilhelm, 2012). Here it is assumed that newly encoded skill representations are being actively and repeatedly reprocessed during sleep,

[1] - *Institute of Sport Science, Saarland University, Saarbruecken, Germany.*

resulting in a qualitative reorganization of the respective memory representations (Albouy et al., 2013; Tamminen et al., 2010). Among those qualitative changes in motor sequence memory representation, a prominent role could be ascribed to chunking-processes.

"Chunking" here means parsing long movement sequences into shorter units of integrated motor elements, which can be prepared by a single motor program and executed as a unified action (Verwey, 1996). Since fast and accurate sequence execution requires rapid *in-advance specification* of requisite elements to be organized in immediate succession, motor chunking can essentially facilitate motor sequence performance and effectively reduce the memory load during sequence execution (Bo and Seidler, 2009; Verwey et al., 2015). Preparatory processes associated with the instantiation of a specific sequence chunk typically are reflected in prolonged duration or reaction time of the first as compared to the following elements of that chunk (Verwey, 2001). Theoretically, those differences in element production times are attributed to different *processing phases* within a chunk: chunk selection and preparation are thought to involve higher cognitive processes carried out by the central (or cognitive) processor, while execution of elements comprised within a chunk are triggered by a motor processor from a motor buffer (Abrahamse et al., 2013).

It is assumed that the formation of the respective hierarchical memory structures is driven by two separable processes: one is involved in parsing long strings of contiguous elements into shorter action sets or "chunks" (*segmentation*). The other one involves the transition between adjacent motor chunks (*concatenation*) and is thought to rely on associations between successive motor chunks or on cognitive selection processes. Efficient concatenation is supposed to be essential for the fluid transitions between motor chunks, thus allowing for distinct chunks within a sequence to be executed smoothly and in rapid succession. As a consequence, concatenation points are supposed to appear less and less prominent in the course of sequence skill development. As motor chunks are consolidated, the central processor will still select, load into the motor buffer and initiate each next motor chunk, but will be increasingly relieved

from more detailed processing requirements. With regard to sequences of *aiming* movements such requirements particularly involve specification of movement parameters (i.e. speed, size, absolute force, muscle group, etc.) to each element of the next-following subsequence. After sufficient practice, in addition to the type and number of submovements and movement order, these parameters are also thought to be integrated in the motor chunk representation, rendering a parameter specification stage thereby unnecessary (Verwey et al., 2015). As a result, loading and initiation of such a chunk will proceed faster now.

Motor chunk concatenation usually is associated with extended practice (Sakai et al., 2003), but has also been shown following nocturnal sleep in a sequential finger-tapping task (Kuriyama et al., 2004). Thus, it seems likely that chunk concatenation, in itself a qualitative change in sequence representation, also contributes to quantitative improvements in motor sequence performance (i.e. increased performance speed) typically observed after initial learning at retest following a period of sleep. This notion was tested by reanalyzing behavioral data from a recent study (Malangré et al., 2014), which has proven sleep-related enhancement consolidation in a complex arm movement sequence. Participants initially practiced this task either in the morning (the *M*orning-*E*vening-*M*orning or MEM-group) or in the evening (the *E*vening-*M*orning-*E*vening or EME-group), and were retested 12 hrs (Retest1) and 24 hrs later (Retest2). Thus, subjects were provided with a regular night of sleep either during the first or during the second retention interval. No day-time naps were allowed. *Total Execution Time* per sequence (which is inversely proportional to performance speed) decreased significantly in both experimental groups during acquisition and significantly decreased again in either group during retention *following sleep*, but not following the respective wake intervals (Figure 1). Considering changes in the error rate (i.e. the number of erroneous sequences), there was no indication of any speed-accuracy trade off.

The present paper addresses the following questions: *(a)* Did *sequence segmentation* occur in the course of the study by Malangré et al. (2014) at all? And if so, *(b)* did *chunk concatenation* contribute to the sleep-related performance enhancements reported in that study?

To this end, behavioral data from that study were analyzed at the sequence *element* level.

Material and Methods

Subjects

24 subjects (25.5 ± 3.9 years; 6 females, 1 left-handed; 18 males) participated in the original study by Malangré et al. (2014), which was conducted in accordance with the ethical standards of the Declaration of Helsinki (1964). All participants gave their written informed consent. Subjects were randomly assigned to two experimental groups, the *E*vening-*M*orning-*E*vening-group (EME), and the *M*orning-*E*vening-*M*orning-group (MEM), with N = 12 each.

Task and procedure in the original study

For a criterion task, subjects originally carried out a 15-element movement sequence with their non-dominant arm. Subjects were seated comfortably in a height-adjustable chair with their upper trunk against the backrest in front of a table-mounted electronic pegboard and a vertical computer screen. With their hand visible all the time, participants could freely move their shoulder, elbow and wrist. On each trial, following a start signal they had to successively fit a small peg into the respective target-holes (depth: 22.22 mm; diameter: 12.7 mm) on the pegboard (Figure 2). The spatial pattern they had to follow was void of any apparent regularity. Transport movements differed in direction and extent (range: 3.83 to 33.75 cm). Precision requirements for sequence elements amounted to an index of difficulty (cf. Fitts, 1954) of 4.95 (± 1.01) on average (Table 1).

Targets were visualized prior to each transport movement by changing the color of the target symbol on the screen from grey to red. Correct execution of a sequence element was indicated by a color change of the respective target stimulus from red to green, while the next target symbol was illuminated red. When a reaching error occurred, the symbol representing the target that had been missed turned green as well, while the next target was illuminated red. Thus, explicit error control always required participants to compare the peg's present position on the pegboard to the target position indicated on the screen. Once a sequence element was terminated, the next movement had to be started immediately, until the sequence was completed.

Start and termination of transport movements were registered by breaking/closing a magnetic contact when the peg was lifted from/inserted into the respective target hole. The respective dwell times (i.e. peg stationary in the hole) proved to be negligible and completely unrelated to the significant treatment effects reported above.

Practice as well as retention tests were organized in successive blocks of 10 trials, separated by a 30 second resting period. Performance measures were the number of Erroneous Sequences per trial block (i.e. the error rate), and *T*otal *E*xecution *T*ime (TET) per sequence, with Total Execution Time averaged for each subject across correct sequences in a trial block. When a reaching error occurred, subjects were to continue the respective sequence trial without correcting for that error. However, suchlike erroneous sequences were excluded from calculation of TET. Initial training amounted to 100 trials (10 blocks of 10 trials), while retests (Retest1 and Retest2) comprised 30 trials (3 blocks of 10 trials) each. Start and Post-Training measures were calculated from the first and the last 30 acquisition trials respectively.

Providing a data base for analyzing performance at the sequence element level

In the original study, Total Execution Time per sequence turned out to be the central outcome variable, with the error rate remaining low throughout the whole experiment and largely unrelated to the treatment. Thus, in the present study, only Total Execution Time data were analyzed, solely reflecting the impact of sleep on performance speed. As a consequence, in order to investigate chunking processes our data reanalysis had to address sequence element duration, which was chiefly expressed by element movement time. However, since reaching movements originally differed in amplitude, elementary movement times within a sequence would have naturally differed due to the different transport distances that had to be covered. Therefore, firstly, discrepancies in element movement times resulting from elements having different trajectory lengths had to be eliminated from the original movement time data.

This was done by relating participants' original sequence element movement times to respective *reference movement times*. The latter were derived from three laboratory staff members

who previously had acquired high expertise in handling the pegboard. These "experts" executed each single sequence element all by itself 20 times in succession as fast as possible with their non-dominant arm. Since each element was carried out *in isolation*, any impact of sequence context on element duration was reliably eliminated. For each element, the ten fastest trials from each expert were averaged. These mean element movement times were highly inter-correlated among experts ($\bar{r}$ = 0.862; p < .001). They were averaged once more across experts, resulting in element-specific reference movement times, which in turn were significantly correlated with the elements' index of difficulty (*r* = 0.843; *p* < .001).

Furthermore, for each participant of the original study and each (correct) sequence execution trial, movement time of each single sequence element was determined. These element-specific movement times were then averaged across all correct sequences per subject and trial block. The resulting *mean element movement time* was then divided by the *reference movement time* of that specific element described in the previous paragraph. This resulted in a *r*elative *E*lement *M*ovement *T*ime (rEMT) of that sequence element for each subject and trial block. Thus, any *differences* in rEMT values still present within a subject's sequence trial are supposed to reflect element-specific "production costs" imposed solely by an element's *position within the sequence*.

Assessing chunk concatenation

Chunk concatenation was assumed whenever rEMTs of putative concatenation points (i.e. *long* rEMTs) were shown to decrease *relatively more* than rEMTs of elements classified as integral parts of a chunk (i.e. *short* rEMTs), thereby representing the mere execution phase. Thus, for each group and time point interval (Acquisition, Retention 1, Retention 2), it was tested if rEMTs of putative concatenation points decreased significantly more than rEMTs of elements identified as integral parts of chunks.

Particularly long and short rEMTs characterizing either putative concatenation points or integral parts of chunks were determined by a comprehensive comparison of all element-specific rEMT values, as explained in detail in the results section below (see the paragraph on sequence segmentation).

Results

For inferential statistics, two- and three-way ANOVAs were run. With respect to repeated-measures factors, in case of violation of the sphericity assumption *df*-correction according to Greenhouse-Geisser was applied. A significance level of *p* ≤ .05 was used for all inferential statistics. Calculations were conducted with SPSS-PC, version 15.0.

Error rate

In the original study, over all error rate was small right from the beginning (i.e. 2.8 ± 2.01 Erroneous Sequences per trial block on average), and did not fluctuate significantly across trial blocks during acquisition (*p* = .079). There was no Group x Block interaction (*p* = .273), nor did groups differ in performance (*p* = .374). During retention, again the error rate did not change across time points (*p* = .239), nor did groups differ (*p* = .904). Thus, the number of Erroneous Sequences remained largely unaffected by treatment conditions during retention (i.e. participants either sleeping or staying awake), and did not increase as execution speed increased. There were also no indications for reaching errors to systematically occur at preferred locations within the criterion sequence. Thus, the error rate did not provide any indication of sequence segmentation.

Sequence segmentation

According to a 2[Group] x 15[Element] x 16[Block] ANOVA (with rep. measures on the last two factors), rEMTs differed and decreased significantly across blocks ($F_{[Element]}$ (5.629, 123.836) = 47.379, *p* < .001, η^2_p = .683; $F_{[Block]}$ (2.972, 65.391) = 59.508, *p* < .001, η^2_p = .730). The rEMT-patterns in both groups were not different, since neither the "Group" main effect (*p* = .362), nor the "Element x Group" (*p* = .374), nor the triple-interaction (*p* = .708) reached statistical significance.

Presence of sequence segmentation then was demonstrated by a comprehensive comparison of all element-specific rEMT values. The respective mean rEMTs for each sequence element (collapsed across all subjects and all 16 trial blocks) with their respective 95%-confidence intervals are listed in Table 1. Two mutually exclusive sets of 5 particularly *long* and 6 particularly *short* rEMTs were identified (Duncan Post-Hoc-Test). The "long" rEMT set comprised sequence elements 4,

6, 7, 11 and 13. The "short" rEMT set included elements 1, 3, 5, 8, 9 and 15. rEMTs of elements 2, 10, 12 and 14 were considered "intermediate". Over all rEMTs were not related to the elements' index of difficulty any more ($r = -.440$; $p = .101$).

Non-idiosyncratic segmentation patterns

To obtain a measure for segmentation similarity across members of an experimental group, we computed the average rEMT pattern in our 15-element arm movement sequence at each time point (i.e. Start, Post-Training, Retest1, and Retest2) for each group. These group-average patterns were used as reference, and at each time point correlations (Pearson's r) were computed for the rEMT-pattern of each subject with the respective group average (cf. Verwey and Eikelboom, 2003). A 2[Group] x 4[Time Point] ANOVA on those Fisher Z-transformed correlations with repeated measures on the factor "Time Point" showed that groups did not differ ($p_{[Group]} = .701$; $p_{[Group \times Time-Point]} = .530$), but revealed a significant "Time Point" main effect (F (3, 66) = 10.114, $p < .001$, $\eta^2_P = .315$). Following respective pairwise comparisons, similarity of individual rEMT patterns with the respective group-average in our task decreased somewhat with practice (Start − Post-Training: F (1, 22) = 25.131, $p < .001$, $\eta^2_P = .533$), but then remained stable across retention ($p \geq .469$). As indicated by the fairly high correlation coefficients (Pearson's r of individual rEMT patterns with group-average patterns, averaged across all 24 subjects: Start: .873; Post-Training: .737; Retest1: .757; Retest2: .775), different participants had altogether very similar segmentation patterns at each time point throughout the whole experiment.

Functional classification of rEMT-clusters

Drawing on the notion of different processing phases, the rEMT clusters identified above were functionally characterized as follows: *(a)* Particularly *"long"* rEMTs immediately followed by at least one element characterized by a "short" or an "intermediate" rEMT most likely indicate typical chunk *concatenation points* (i.e. elements 4, 7, 11 and 13). *(b)* Particularly *"short"* rEMTs immediately following assumed concatenation points are thought to reflect the mere *execution* process (i.e. elements 3, 5, 8, 9 and 15). *(c)* *Intermediate* rEMT values (elements 2, 10, 12 and 14) could possibly be caused by individual differences, but could also designate elements not

fully integrated into a motor chunk yet because of sequence knowledge still being somewhat weak.

Special cases: element 6 cannot be considered a concatenation point, because it was nearly always followed by another "long" rEMT (i.e. element 7). Element 1 represents sequence initiation, usually associated with a comparatively long preparation time. In the original study, however, participants were free to initiate the criterion sequence within a certain time limit following the start-signal. This preparation time preceding actual sequence initiation at that time had not been recorded and was not included in the TET measure. Therefore, in the present data-reanalysis sequence initiation was not taken into account.

Chunk concatenation

Based on the two distinct sets of four "long" rEMTs considered putative concatenation points (i.e. elements 4, 7, 11 and 13), and five "short" rEMTs considered integral parts of chunks (i.e. elements 3, 5, 8, 9 and 15), the factor "Processing Phase" (levels: "concatenation" = long rEMTs, "execution" = short rEMTs) was subjected to further analysis. Three 2[Group] x 2[Time Point] x 2[Processing Phase] ANOVAs (rep. measures on the last two factors) were conducted separately for acquisition and each retention interval (Acquisition: Start − Post-Training; Retention 1: Post-Training − Retest1; Retention 2: Retest1 − Retest2).

Globally, long rEMTs decreased significantly more than short ones during the first two intervals, but not during the third any more (Acquisition: $F_{[Time Point \times Processing Phase]}$ (1, 22) = 30.173, $p < .001$, $\eta^2_P = .578$; Retention 1: $F_{[Time Point \times Processing Phase]}$ (1, 22) = 7.417, $p = .012$, $\eta^2_P = .254$; Retention 2: $F_{[Time Point \times Processing Phase]}$ (1, 22) = 2.504, $p = .128$). While groups did not differ in this respect during acquisition (triple interaction: F (1, 22) = .211, $p = .650$), the respective *triple interactions* turned out *significant* for each of the two subsequent *retention intervals* (Retention 1: F (1, 22) = 7.900, $p = .010$, $\eta^2_P = .264$; Retention 2: F (1, 22) = 5.931, $p = .023$, $\eta^2_P = .212$). That is, across each of the retention intervals the two experimental groups exhibited significantly different patterns of change in rEMTs (Figure 3).

According to our expectation of sleep boosting chunk concatenation, this should have been the case indeed if long rEMTs had decreased more than short ones in subjects provided with sleep

during the respective retention interval than in those subjects that had to stay awake.

At this point, for final clarification detailed follow-up analyses were required. With respect to the repeated measures factors "Time Point" and "Processing Phase" (levels: long rEMTs, short rEMTs), 2[Time Point] x 2[Processing Phase] repeated measures-ANOVAs were calculated separately for each group and each retention interval (Retention 1: Post-Training − Retest1; Retention 2: Retest1 − Retest2), respectively. For both groups in both retention intervals long and short rEMTs significantly differed ($p < .001$). During *Retention 1*, there was no significant change of rEMTs in the *MEM* group any more ($F_{\text{[Time Point]}}$ (1, 11) = 4.581, $p = .056$, $\eta^2_p = .294$; interaction: $p = .961$), while in the *EME* group (with sleep provided during this first interval) the main effect "Time Point" as well as the interaction

term turned out significant ($F_{\text{[Time Point]}}$ (1, 11) = 37.903, $p < .001$, $\eta^2_p = .775$; $F_{\text{[Time Point x Processing Phase]}}$ (1, 11) = 19.321, $p = .001$, $\eta^2_p = .637$).

During *Retention 2*, however, the main effect "Time Point" and the interaction term turned out significant this time in the *MEM* group, i.e. the group provided with sleep during this second retention interval ($F_{\text{[Time Point]}}$ (1, 11) = 30.566, $p < .001$, $\eta^2_p = .735$; $F_{\text{[Time Point x Processing Phase]}}$ (1, 11) = 5.327, $p = .041$, $\eta^2_p = .326$), while in the *EME* group neither rEMT changes nor interactions were significant any more ($p \geq .405$).

Thus, comparison of changes over time between the two sets of rEMTs revealed that in each experimental group the *long* rEMTs *decreased significantly more* than short ones during *practice*, and then once more during *retention* when tested *following sleep*, but not when tested following wake.

Table 1

*Relative Element Movement Time (rEMT) and Index of Difficulty (ID).*

Element	ID	rEMT		
		short	intermediate	long
1	5.74	1.79 (1.70 − 1.89)		
2	4.59		1.97 (1.84 − 2.11)	
3	5.21	1.81 (1.70 − 1.92)		
4	2.58			2.24 (2.05 − 2.42)
5	5.01	1.85 (1.76 − 1.95)		
6	5.18			2.35 (2.24 − 2.46)
7	5.11			2.71 (2.55 − 2.87)
8	5.60	1.70 (1.62 − 1.78)		
9	5.86	1.84 (1.76 − 1.91)		
10	5.46		2.09 (1.98 − 2.21)	
11	2.58			2.38 (2.15 − 2.61)
12	5.01		2.07 (1.95 − 2.20)	
13	5.46			2.28 (2.15 − 2.41)
14	5.11		1.98 (1.87 − 2.08)	
15	5.74	1.80 (1.71 − 1.90)		

Annotations: Reported are for each sequence element (a) Fitts' index of difficulty
(ID = log$_2$ [2A/W]; A = amplitude, and W = target width),
and (b) rEMT mean values (dimensionless) for each sequence element,
calculated across all 24 subjects and all 16 trial blocks,
and the respective 95% confidence intervals (in parentheses).
Note that upper confidence interval limits of short rEMTs
and lower confidence interval limits of long rEMTs do not overlap.
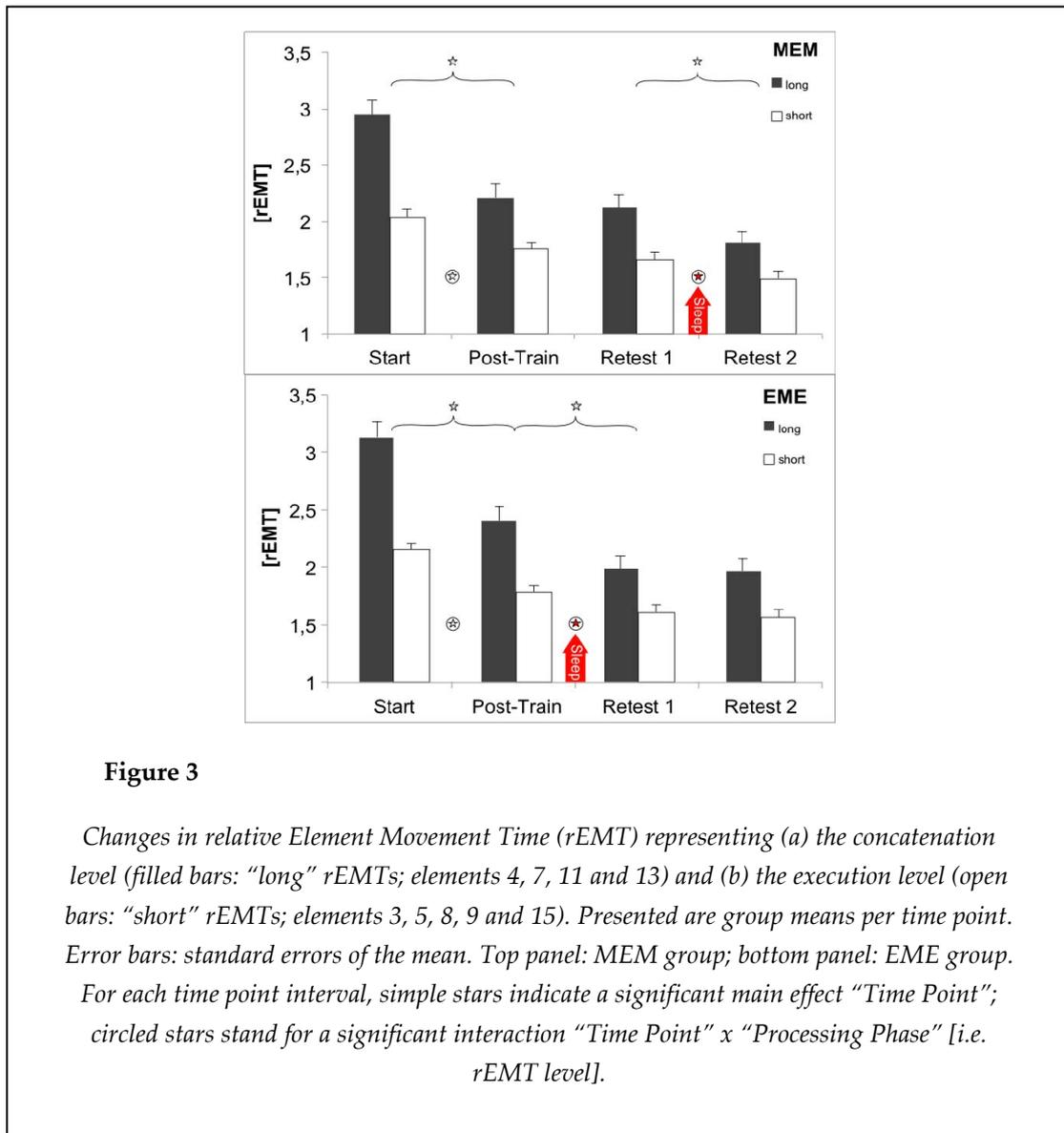*See also the results section for classification of rEMT values.*
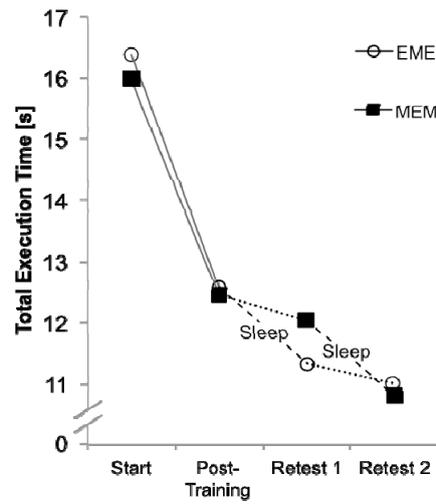
Figure 1

Mean total sequence execution time (TET [s]; correct sequences only) per group (MEM, EME) and time point (Start: block 1, 2 & 3; Post-Training: block 8, 9 & 10; Retest 1 (after 12 hrs): block 11, 12 & 13; Retest 2 (after 24 hrs): block 14, 15 & 16). Acquisition: solid lines. Retention: dotted lines = wake periods; dashed lines = sleep-filled periods (data from Malangré et al., 2014).
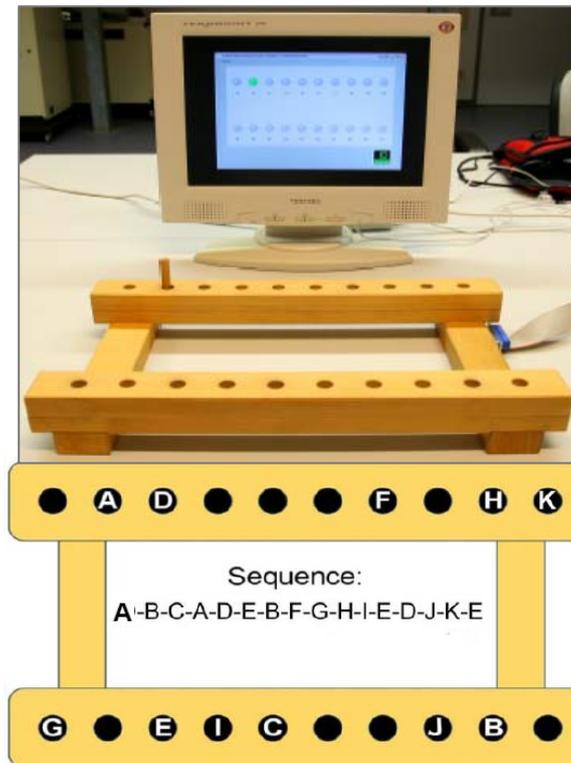
Figure 2

Experimental apparatus (upper panel), and spatial locations to be reached for one after the other, defining the fifteen-element arm movement sequence (lower panel).

Figure 3

Changes in relative Element Movement Time (rEMT) representing (a) the concatenation level (filled bars: "long" rEMTs; elements 4, 7, 11 and 13) and (b) the execution level (open bars: "short" rEMTs; elements 3, 5, 8, 9 and 15). Presented are group means per time point. Error bars: standard errors of the mean. Top panel: MEM group; bottom panel: EME group. For each time point interval, simple stars indicate a significant main effect "Time Point"; circled stars stand for a significant interaction "Time Point" x "Processing Phase" [i.e. rEMT level].

Discussion

In order to better understand what kind of representational changes take place in the course of an "active system consolidation" process in motor skills during sleep, performance data from a recent study by Malangré et al. (2014) were reanalyzed. In that study, a sequence of 15 precise reaching movements had to be executed as fast as possible with as few errors as possible. For the purpose of data reanalysis, participants' original element-specific movement times were subjected to a transformation procedure in order to eliminate differences in movement duration resulting purely from dissimilar transport distances which had to be covered for different sequence elements. The resulting dimensionless "relative Element Movement Times" (rEMTs) eventually allowed participants' performance to be addressed at the sequence element level. As it was then evidenced by the present data reanalysis, sleep-related offline improvements in performance speed found in the original study were closely associated with increased concatenation of adjacent sets of sequence elements (i.e. chunks).

Sequence segmentation is found to be largely non-idiosyncratic

Presence of *sequence segmentation* then was demonstrated by comparison of element-specific rEMT values. Three different sets of rEMTs

were identified: five were long, six were short and four were intermediate. Drawing on the concept of different processing phases in sequence production (i.e. initiation, concatenation and execution) proposed in the context of the dual processor model (Abrahamse et al., 2013; Verwey et al., 2015), those elements with particularly *long* rEMTs *immediately preceding* short or intermediate ones were understood as possible chunk *concatenation* points, indexing the transition from one motor chunk to the next (the concatenation phase). The resulting segmentation pattern turned out *fairly consistent* across time points and appeared also to be largely *non-idiosyncratic*.

This requires closer consideration for it has been shown previously participants sometimes tend to develop *individually different* segmentation patterns when rapidly executing key-pressing sequences with little or no apparent regularities (Kuriyama et al., 2004; Sakai et al., 2003). More specifically, Verwey and Eikelboom (2003) showed that the amount of idiosyncratic sequence segmentation differed for different sequences and was virtually identical for all subjects only with one particular sequence. Thus, in order to check if our results were possibly obscured by initially averaging rEMTs across subjects and trial blocks, we also applied *k-means clustering* (MacQueen, 1967) for motor chunk identification (cf. Song and Cohen, 2014). This data-driven procedure does not require any a priori assumptions on chunk length or on the ratio of long and short rEMTs. It is sensitive to idiosyncratic chunking behavior (which means the concatenation points are found to occur at different sequential locations for different people) and it allows tracing chunk development across time points. As reported in Appendix 1, our above finding of sequence segmentation turning out *non-idiosyncratic* and fairly stable across retention was fully confirmed.

Sequence segmentation being non-idiosyncratic in our study is explained most parsimoniously by specific task constraints which impose segmentation at the same locations within the sequence across participants. In finger-tapping tasks, sequence segmentation was imposed artificially by *(a)* introducing regularities in response order, *(b)* temporarily inserting a pause between two successive stimuli, *(c)* using differently colored key-specific stimuli to distinguish subsequences (Abrahamse et al., 2013). None of these conditions immediately applies to our arm movement sequence. However, reliably locating the next target on the pegboard always proved to be difficult, when there was a whole string of holes extending to each side of the new target and long transport distances had to be covered. To execute such elements correctly, early on in practice required elaborate visual search and implementation of auxiliary strategies (e.g. counting out holes). This in turn could have had an effect equivalent to imposing an artificial pause at this point of the sequence. The same holds true where the "dynamical flow" of sequence production was interrupted on account of the spatial arrangement of two successive target holes.

Thus, all in all our reanalysis revealed a (none-idiosyncratic) segmentation pattern of five sub-sets (including the one immediately following sequence initiation) comprising two to four elements each. This result nicely reconciles the notion that longer, fixed movement sequences (> 4 elements) generally are subdivided into a succession of multiple motor chunks (Abrahamse et al., 2013; Bo and Seidler, 2009). The number and size of sub-sets found in our 15-element arm movement sequence are well in accordance with capacity limitations of working memory (Cowan, 2010) as well as the notion of the motor buffer being limited to about 3 to 5 elements, thereby defining the average length of a single motor chunk (Bo and Seidler, 2009; Verwey and Eikelboom, 2003).

Chunk concatenation evolves with sleep and physical practice

In the discrete sequence production task, sequence segmentation was shown to be maintained even after considerable practice (Verwey and Eikelboom, 2003). Lasting segmentation might imply that the length of motor chunks is structurally limited, although there may be individual differences as to how many elements are comprised into one chunk. This considered, we did not expect the absolute *number* of concatenation points initially identified in our arm movement sequence to significantly decrease across time points on account of smaller chunks being totally concatenated into longer ones. As it turned out, participants' segmentation patterns were indeed well preserved throughout

the whole experiment, and the number of concatenation points on average did not change.

We did, however, expect relative Element Movement Times at chunk concatenation points to *decrease significantly more* during a *sleep*-filled retention period than relative Element Movement Times reflecting merely rapid execution *within* a chunk. This notion of a selective sleep-related offline improvement had already been proposed by Kuriyama et al. (2004) more than ten years ago. These authors found sleep-related performance enhancements (increased execution speed and error reduction) in four finger-tapping tasks of different length (five-elements or nine-elements) and complexity (unimanual or bimanual). They also assessed each participant's slowest ("problem point") and fastest ("easy point") transition from one key press to the next in each sequence. At the group level, they found significant overnight improvements in transition speed occurring at the slowest, but not at the fastest transition position. This selective improvement was specific to the overnight learning process and was neither found during acquisition, nor during a wake retention interval in a control group.

Other than Kuriyama et al. (2004), we did not compare just two transition points between sequence elements. Instead we compared two whole sets of elements which were either classified as putative concatenation points (long rEMTs) or identified as integral parts of a chunk (short rEMTs) before. During *retention*, in each of our experimental groups those *long* rEMTs decreased *significantly more* than the short ones *following* sleep, while there were no significant changes in rEMTs following wake. According to the "Cognitive framework for Sequential Motor Behavior" (Verwey et al., 2015), this selective improvement may be conceived as the central processor becoming more efficient and thereby faster at selecting, retrieving and loading each chunk into the motor buffer for subsequent execution during sleep, but not during wake.

Interestingly, we observed the same selective improvement also during *acquisition*. Here, both sets of rEMTs significantly decreased, but the *long* ones again decreased *significantly more* than the short ones. While Kuriyama et al. (2004) also found transition points to improve at a different rate across acquisition, they notwithstanding interpreted significant improvements of *both* transitions (fastest *and* slowest) in evidence of chunk concatenation being completely *absent* during initial training. In the light of our present findings, this position ought to be reconsidered. We would rather argue that progressive chunk concatenation always comes about as an integral part of sequence learning, either "online" associated with overt physical activity and processing of error information or "offline" associated with a sleep-dependent rearrangement of the respective movement representation (i.e. enhancement consolidation).

These findings also bear implications relevant to sports training. This would specifically pertain to complex motor skills requiring fast and precise execution of a series of sub-movements in Euclidian space, like, e.g., performing a karate kata or isolation dance routine. Memory representations of suchlike skills are almost certainly subject to chunking processes. Efficient chunk concatenation will make reproduction of such a skill easier and more reliable specifically under free recall conditions in a tournament. Thus, whenever motor tasks of this kind have to be learned or relearned efficiently, we would recommend practice to be distributed over two or more sessions separated by night or daytime sleep respectively in order to facilitate chunk formation and concatenation. Distributed practice schedules of this kind might be especially beneficial in situations, where the amount of practice per session is limited of necessity, as is often the case in rehab-training or acquisition of sport skills requiring extraordinary high physical or mental effort.

Conclusion

In this paper, performance data of a recent study by Malangré et al. (2014) showing sleep-related offline learning in a complex arm movement sequence were reanalyzed at the sequence element level. This way motor chunk concatenation was shown to evolve in the course of sleep-related enhancement consolidation, thereby making up a relevant portion of overnight improvements in sequence execution speed reported previously. Chunk concatenation also proceeded with physical practice, but was not observed at retests following a waking retention interval. To our knowledge for the first time, such qualitative changes in sequence memory

representation have been shown in a task involving unrestrained multi-joint movements of an entire limb. These findings corroborate and complement earlier findings by Kuriyama et al. (2004), and successfully extend the respective theoretical suppositions based on previous work with the Discrete Sequence Production-task and derived from the Dual Processor Model and the "Cognitive framework for Sequential Motor Behavior" to gross motor abilities closer to so many every-day and vocational activities. Occurrence of non-idiosyncratic sequence segmentation was plausibly attributed to constraints imposed by features specific to our criterion task. This might encourage researchers to increasingly consider task-specific constraints imposed by tool use and environment, and thereby to arrive at conclusions of improved ecological validity.

Acknowledgements

Appendix 1

K-means clustering. First, for each participant element-specific rEMT values were averaged across trial blocks per time-point (Start, Post-Training, Retest1 and Retest2). Then, the clustering procedure (matlab 'kmeans', cluster number = 2, empty action = drop, outliers beyond 2 SD of the mean rEMT are set to 2SD of the mean) was applied on each individual's rEMT-data set per time point. With a cluster number = 2, rEMTs of the 15 sequence elements self-sorted as either long (i.e. slow element) or short (i.e. fast element).

Sequence segmentation. After clustering, motor chunks were easily identified as one or more short rEMTs (i.e. fast elements) following sequence initiation or one of the long rEMTs. With the first (sequence initiation) and the last element excluded, 261 such possible individual concatenation points could be identified in all 24 subjects across the four time points (Start, Post-Training, Retest1 and Retest2). Of these a total of 204 (78.2%) were allotted to the same four sequence elements, namely the elements 4 (16.8%), 7 (35%), 11 (9.6%) and 13 (16.8%). The rest (21.8%) was distributed over eight other sequence elements (≤ 5.4% per element). Including the initial subset of elements following sequence initiation, subjects on the average subdivided the 15-element movement sequence into 3.61 ± .96 chunks (SE: ± .11).

Possible differences in the number of chunks between groups and across time points (TP) were analyzed by means of a 2[Group] x 4[Time Point] ANOVA (rep. measures on "Time Point"). This ANOVA altogether failed to reach significance ($F_{[Time Point]}$ (3, 66) = 2.050, p = .115; $F_{[Group]}$ (1, 22) = .223, p = .642; no significant interaction: F (3, 66) = 1.162, p = .331). Once established during acquisition, the resulting pattern of slow and fast elements appeared to be fairly stable across both retention intervals, as indicated by highly significant correlations (p < .01) of segmentation patterns between the respective time points (mean correlation coefficients calculated from Fisher's Z-transformed correlations at the subject level; Pearson's r: .782 (Post-Training – Retest1), and .682 (Retest1 – Retest2)).

Total chunk concatenation. According to Song and Cohen (2014), as smaller chunks are *completely* (!) concatenated into longer chunks, the ratio of the *maximum* number of "fast" elements comprised in one

chunk to the *total* number of "fast" elements within *all* motor chunks should increase and eventually approach one (fully concatenated). This ratio was calculated for each participant at each time point. On these ratios, a 2[Group] x 4[Time Point] ANOVA (rep. measures on "Time Point") was run. Neither the main effects ($F_{[Time\ Point]}$ (3, 66) = 1.825, p = .151; $F_{[Group]}$ (1, 22) = .379, p = .544), nor the interaction turned out significant (F (3, 66) = .4402, p = .725). Thus, at the group level no indication of *total* chunk concatenation across time points was found.

References

Abrahamse EL, Ruitenberg MFL, de Kleine E, Verwey WB. Control of automated behaviour: insights from the discrete sequence production task. *Front Hum Neurosci*, 2013; 7: Article 82

Al-Sharman A, Siengsukon CF. Sleep enhances learning of a functional motor task in young adults. *Phys Ther*, 2013; 93: 1625-1635

Albouy G, King BR, Maquet P, Doyon J. Hippocampus and striatum: Dynamics and interaction during acquisition and sleep-related motor sequence memory consolidation. *Hippocampus*, 2013; 23: 985-1004

Blischke K, Erlacher D, Kresin H, Brueckner S, Malangré A. Benefits of sleep in motor learning – prospects and limitations. *Journal of Human Kinetics*, 2008; 20: 23-35

Bo J, Seidler RD. Visuospatial working memory capacity predicts the organization of acquired explicit motor sequences. *J Neurophysiol*, 2009; 101: 3116-3125

Born J, Wilhelm I. System consolidation of memory during sleep. *Psychol Res*, 2012; 76: 192-203

Cowan N. The magical mystery four: How is working memory capacity limited, and why? *Curr Dir Psychol Sci*, 2010; 19: 51-57

Doyon J, Korman M, Morin A, Dosti V, Tahar AH, Benali H, Karni A, Ungerleider LG, Carrier J. Contribution of night and day sleep vs. simple passage of time to the consolidation of motor sequence and visuomotor adaptation learning. *Exp Brain Res*, 2009; 195: 15-26

Fischer S, Nitschke MF, Melchert UH, Erdmann C, Born J. Motor memory consolidation in sleep shapes more effective neural representations. *J Neurosci*, 2005; 25: 1248-1255

Fitts PM. The information capacity of the human motor system in controlling the amplitude of movement. *J Exp Psychol*, 1954; 47: 381-391

Genzel L, Quack A, Jäger E, Konrad B, Steiger A, Dresler M. Complex motor sequence skills profit from sleep. *Neuropsychobiology*, 2012; 66: 237–243

Hoedlmoser K, Birklbauer J, Schabus M, Eibenberger P, Rigler S, Mueller E. The impact of diurnal sleep on the consolidation of a complex gross motor adaptation task. *J Sleep Res*, 2015; 24: 100-109

Kempler L, Richmond JL. Effect of sleep on gross motor memory. *Memory*, 2012; 20: 907-914

Kuriyama K, Stickgold R, Walker MP. Sleep-dependent learning and motor-skill complexity. *Learn Memory*, 2004; 11: 705-713

MacQueen J. Some methods for classification and analysis of multivariate observations. In Lecam L, Neyman J (eds.). *Proc 5th Berkely Symp Math Stat Prob 1965/66*. Berkely, 1: 281-297; 1967

Malangré A, Leinen P, Blischke K. Sleep-related offline learning in a complex arm movement sequence. *Journal of Human Kinetics*, 2014; 40: 7-20

Morita Y, Ogawa K, Uchida S. The effect of a daytime 2-hour nap on complex motor skill learning. *Sleep and Biological Rhythms*, 2012; 10: 302–309

Sakai K, Kitaguchi K, Hikosaka O. Chunking during human visuomotor sequence learning. *Exp Brain Res*, 2003; 152: 229-242

Song S, Cohen L. Impact of conscious intent on chunking during motor learning. *Learn Mem*, 2014; 21: 449-451

Tamminen J, Payne JD, Stickgold R, Wamsley EJ, Gaskell MG. Sleep spindle activity is associated with the integration of new memories and existing knowledge. *J Neurosci*, 2010; 30: 14356-14360

Verwey WB. Buffer loading and chunking in sequential key-pressing. *J Exp Psychol Human*, 1996; 22: 544-562

Verwey WB. Concatenating familiar movement sequences: The versatile cognitive processor. *Acta Psychol*, 2001; 106(1-2): 69-95

Verwey WB, Eikelboom T. Evidence for lasting sequence segmentation in the discrete sequence-production task. *J Mot Behav*, 2003; 35: 171-181

Verwey WB, Shea CH, Wright DL. A cognitive framework for explaining serial processing and sequence execution strategies. *Psychon Bull Rev*, 2015; 22: 54-77

Walker MP. A refined model of sleep and the time course of memory formation. *Behav Brain Sci*, 2005; 28: 51-104

Corresponding author:

Klaus Blischke
Saarland University, Institute of Sport Science
Campus B8-2
D-66123 Saarbruecken, Germany.
Phone: +49-681-302-4172
e-mail: k.blischke@mx.uni-saarland.de