



Paradoxical Long-term Memory Augmentation following Temporal Pairing between “Limited” and “Extensive” Motor Sequence Training Experiences

Yaniv Oren¹, Maria Korman^{2*}, and Avi Karni^{1*}

Abstract

■ Consecutive training on two movement sequences often leads to retroactive interference—obstructing memory for the initially trained sequence but not for the second. However, in the context of hippocampal-system dependent memories, a poor learning experience, memory for which would soon decay, can be enhanced if temporally paired with a “strong” memory triggering experience. The synaptic tagging and capture hypothesis explains this paradoxical enhancement by suggesting that only strong experiences generate cellular resources necessary for synaptic remodeling. However, synapses engaged in a “weak” learning experience can capture and utilize plasticity-related resources generated for a subsequent strong learning experience. Here, we tested whether such a “paradoxical” outcome would result in the context of motor (procedural) memory, if two movement sequences are unequally trained, consecutively. We show, in young adults

($n = 100$), that limited practice on a novel sequence of finger-to-thumb opposition movements led to different long-term outcomes, depending on whether and when (5 min, 5 hr) it was followed by extensive training on a different sequence. Five-minute pairing only resulted in overnight gains for the limited-trained sequence that were well-retained a week later; the overnight gains for the extensively trained sequence were compromised. Thus, consecutive training on different motor tasks can result in mnemonic interactions other than interference. We propose that the newly discovered mnemonic interaction provides the first-tier behavioral evidence in support of the possible applicability of notions stemming from the synaptic tagging and capture hypothesis in relation to human motor memory generation, specifically in relation to the practice-dependent consolidation of novel explicitly instructed movement sequences. ■

INTRODUCTION

Temporal pairing between different motor experiences can lead to interference effects. Thus, experiences occurring subsequent to the learning session can disrupt the learner’s ability to express the expected long-term performance gains. For example, young adults training in a session combining, first, the generation of movements in a force field of a given direction and then within a time-window extending up to 4–6 hr, in moving in an orthogonally directed force field, lose the expected long-term gains for movement in the initially trained force field (Brashers-Krug, Shadmehr, & Bizzi, 1996). Such retroactive interference occurs consistently across many motor tasks, including in motor sequence learning. Young adults training in the performance of a finger-to-thumb opposition movement sequence and shortly after receive training on the mirror reversed sequence (using the same hand and same digit movements) fail to express the full delayed consolidation phase benefits of the initial training overnight but show robust delayed gains in the performance

of the second sequence (Friedman & Korman, 2016; Korman et al., 2007; Korman, Raz, Flash, & Karni, 2003; Walker, Brakefield, Morgan, Hobson, & Stickgold, 2002). However, some motor training experiences may be too “weak” to trigger the expression of delayed consolidation phase gains in the performance of a movement sequence, even when no interference is afforded, specifically, when too little training is afforded (e.g., Korman et al., 2003). Here, we describe, for the first time, specific training conditions wherein the temporal pairing between two different motor sequences, trained in succession, results in a paradoxical mnemonic enhancement of the initially trained movement sequence and a suppression of consolidation phase gains for the second, subsequently trained, movement sequence. This “paradoxical” outcome, we show, appears when the two movement sequences have been unequally trained in close temporal pairing.

It was suggested that retroactive interference is critically dependent on an overlap between the brain representations of the two tasks and that such overlap is more likely when the two tasks are novel, competing for (cell-level) general resources for their execution and, specifically, for their consolidation into long-term memory consolidation (Maaravi Hesse, Gal, & Karni, 2016; Krakauer, Ghez,

¹University of Haifa, Israel, ²Ariel University, Israel

*Equal contributions.

& Ghilardi, 2005; Miall, Jenkinson, & Kulkarni, 2004; Tong, Wolpert, & Flanagan, 2002). However, in the domain of hippocampus-dependent learning, the very same logic has been used to explain a quite different mnemonic outcome of consecutive learning experiences, enhancement rather than interference. There is compelling evidence that a weak learning experience can be successfully retained, and its memory enhanced if temporally paired with a “strong” experience (Okuda, Højgaard, Privitera, Bayraktar, & Takeuchi, 2021; Ramirez Butavand et al., 2020; Vishnoi, Raisuddin, & Parvez, 2016; Moncada, Ballarini, & Viola, 2015; Redondo & Morris, 2011; Ballarini, Moncada, Martinez, Alen, & Viola, 2009; Frey & Morris, 1997). The synaptic tagging and capture (STC) hypothesis and the notion of “behavioral tagging” (BT) explain this effect by suggesting that both weak and strong learning experiences engage a large enough overlapping (shared) population of neurons. The mnemonic gain for the weak experience is related to two distinct cellular processes that occur within a short time-window within this common pool of neurons engaged by the two tasks (Okuda et al., 2021; Ramirez Butavand et al., 2020; Vishnoi et al., 2016; Moncada et al., 2015; Redondo & Morris, 2011; Ballarini et al., 2009; Frey & Morris, 1997). The underlying notion is that two dissociable cellular events must occur to achieve long-term synaptic plasticity: one, the generation of local *synaptic tags* that temporarily mark the task/event-specific synapses that have been active in the learning experience; two, the synthesis of *plasticity-related products* (PRPs) proteins and factors that are synthesized de novo and then used in the remodeling of tagged synapses. Some, if not most, PRPs are potentially available to all the synapses of a given neuron but are directed only to tagged synapse (Redondo & Morris, 2011; Frey & Morris, 1998a, 1998b). Importantly, synaptic tagging can proceed even when the learning experience is weak; however, the generation of PRPs is dependent on the intensity of the learning experience (Almaguer-Melian et al., 2012; Redondo & Morris, 2011; Frey & Frey, 2008; Sajikumar & Frey, 2004; Frey & Morris, 1997). Therefore, synaptic tags on synapses selectively engaged by a weak (limited) learning experience can capture PRPs synthesized for a different, strong (sufficient, extensive) learning experience if the two learning experiences occur in temporal proximity to each other and activate an overlapping population of neurons (Nomoto & Inokuchi, 2018; Maaravi Hesseget al., 2016; Nomoto et al., 2016).

Both the STC and the BT stem from experiments addressing hippocampus-dependent LTM, but it is not known whether the same principles apply to the acquisition and consolidation of procedural “how to” knowledge. Both explicit and implicit processes subserve the learning of new motor skills (Krakauer, Hadjiosif, Xu, Wong, & Haith, 2019; Hardwick, Rottschy, Miall, & Eickhoff, 2013; Karni et al., 1998; Karni, 1996). However, a recent review suggests two forms of “how to” knowledge by referring to (a) qualitative changes in motor task performance through

practice, conceptualized as increased motor “acuity,” and (b) to the gaining of knowledge (explicit or implicit) of the necessary discrete actions to be performed to solve the motor task but do not necessarily require changes in the execution of these actions (Krakauer et al., 2019).

Here, we tested whether the predictions of the STC hypothesis would hold true for a motor skill learning task in which motor “acuity” changes have been shown to occur (Korman, Gal, Gabitov, & Karni, 2021; Friedman & Korman, 2016; Rozanov, Keren, & Karni, 2010). The gains attained in training on this task can be consolidated and retained as procedural long-term memory even in amnesiacs, that is, when the hippocampal system is compromised by disease or injury (Korman et al., 2018; Döhring et al., 2017). We hypothesized that the long-term outcome of a limited practice-experience on a novel finger opposition sequence (FOS) movement task may change if closely followed by extensive practice of a different (novel) finger opposition movement sequence. In the FOS task, participants start training only after they have established a clear explicit understanding of the order of the required movement sequence. During and after training, there is an implicit change in the execution quality of the sequence (Friedman & Korman, 2012; Rozanov et al., 2010; Korman et al., 2003). A major hub for motor procedural memory, specifically for finger-to-thumb opposition movements, is motor cortex (Guo et al., 2015; Gabitov, Manor, & Karni, 2014; Xu et al., 2009; Yang, Pan, & Gan, 2009; Matsuzaka, Picard, & Strick, 2007; Karni et al., 1995). If synapses in neurons engaged by two movement sequences are tagged in a sequence specific manner, the two sets of synapses would compete for a limited amount of available PRPs, presumably generated only for the second extensively trained sequence (Okuda et al., 2021; Viola, Ballarini, Martínez, & Moncada, 2014). Thus, in a combined session of a limited-trained sequence and an extensively trained sequence, consolidation gain for the limited-trained sequence may appear but come at a cost of reduced gains for the extensively trained sequence.

An important behavioral marker of motor skill consolidation in humans is the delayed, often overnight, expression of additional gains in the performance of the novel skill, that is, gains that evolve offline, after the termination of the practice session (Gabitov et al., 2014; Friedman & Korman, 2012; Korman et al., 2003, 2007; Walker et al., 2002). Within the framework of the STC hypothesis (applied to mnemonic representations possibly outside the hippocampal system), a possible explanation for the delay (latency) in expressing offline performance gains after extensive practice is that PRP-mediated structural modification of task-specific synapses (tagged during the practice session) is time-demanding (Karni et al., 1998). Thus, delayed gains in performance presumably reflect the engagement of task-specific synapses that were structurally modified or generated de novo during the postsession interval (Xu et al., 2009; Yang et al., 2009; Walker, 2005; Maquet et al., 2003; Karni, 1996). In the current

study, our main prediction was that the overnight expression of delayed gains in the performance of a finger opposition movement sequence for which only limited (LIM) training (16 task iterations) was given will benefit from temporal pairing with extensive (EXT) training (160 task iterations) on a different sequence of opposition movements.

METHODS

Participants

One hundred right-handed, young, healthy, university students (19–35 years old, mean = 24.44 years, $SD = 3.24$ years, 83% female and 17% male participants) enrolled in the study. Group size was estimated using the G*Power 3.1 software program. Analysis of sample size given $\alpha = .05$ and power of .80 with a medium effect size ($f = 0.25$) and a correlation among repeated measures of .5 indicated that a sample size greater than 12 participants per group was required for the assessment of an interaction of Group \times Time-Point and to assess time-point effects within groups. Taking into account that the correlation among repeated measures were expected to be on the order of 0.7 (in fact, the correlations were above 0.8), a group size of 11 was appropriate for examining an even smaller effect size ($f = 0.20$). On the basis of the effects observed in previous studies with neurotypical young adult participants (Korman et al., 2003, 2017), the control groups that did not experience interfering interventions showed overnight (postsession consolidation phase) gains in performance on the order of 20–30% for extensively trained sequence (EXT training in the current study). On the basis of a standard deviation of 15% in each group, and a power of 0.80, a group size of 16 participants per group, randomly assigned, was appropriate. The study was approved by the Human Experimentation Ethics Committee of the Faculty of Education of the University of Haifa (054–16). Ethics approval was received from the Haifa University Institutional Review Board, and all experiments were performed in accordance with the relevant guidelines and regulations. The participants were recruited using the University of Haifa advertisement boards and a dedicated Facebook board. All participants received a detailed explanation about the study and signed the informed consent before the start of the experiment.

All participants were prescreened to exclude neurological, medical, or psychiatric illness that could impair fine motor performance, reported of having at least 6 hr of regular sleep per night, without sleep disorders. Musicians and/or professional typists, shift workers, heavy smokers and alcohol and caffeine consumers (more than four to five cigarettes per day and more than three drinks per day), and individuals who take prescribed medications on a regular basis were excluded. The study task was novel for all participants.

Task

The motor task, FOS task (Karni, 1996), required to repetitively perform a five-element sequence of finger-to-thumb movements, using the left (nondominant) hand “as fast and as accurately as possible” (Figure 1A; Korman et al., 2003, 2007). Sequences A, B, and the New sequence, all composed of the same movements arranged in a different order, were randomly assigned to serve as a first trained sequence, a second trained sequence, and a New (transfer) sequence. Performance of a New sequence is not included in the current analyses. Previous studies confirmed that the three sequences were of similar difficulty (Friedman & Korman, 2016; Balas, Roitenberg, Giladi, & Karni, 2007). The participants performed the FOS task while sitting on a chair with the performing hand resting on the table in front of the participant. The hand was in the direct view of a video camera, to allow capturing of all digit movements. Visual feedback (hand view) was not afforded. Participants were explicitly instructed to rely on the sensory feedback (touch sense) and to avert their gaze from the hand; the experimenter ensured they followed this instruction throughout all training and test sessions.

Procedure

The study included six groups. All participants had their training session around mid-day between 11 a.m. and 1 p.m. Participant’s randomization to different groups was done separately for the main experiments (Figure 1B) and the temporal-interval testing (Figure 1C). Half of the participants in each training condition were randomly assigned to start with Sequence A, and the other half were randomly assigned to start with Sequence B (Figure 1A).

The extent of practice on each sequence, referred from here as “dose,” was either EXT or LIM. The EXT practice included four pretest blocks, 160 repetitions of the sequence (training), and four posttest blocks. Each test block lasted 30 sec and required to perform the sequence as fast and as accurately as possible, with the trainee’s left, nondominant, hand. Auditory cues were used to initiate and terminate each of the test blocks. The training session included 10 sets of 16 sequence repetitions per set, where the initiation of each sequence was cued by a beep every 2.5 sec. To keep a strict limit on the number of sequence iterations in the LIM training, the number of test blocks in the pre and post training was reduced from four to one. Thus, the LIM practice included one pretest block, 16 cued repetitions of the sequence (a single training set), and one posttest block. The structure of the test blocks was similar to the EXT test blocks. Altogether, the number of task iterations for training in the LIM and EXT conditions was chosen to consistently (across most individuals) lead to either no delayed gains or robust delayed gains, respectively. To enable a comparison of the current set of results to previous studies (Friedman & Korman, 2012, 2016; Korman

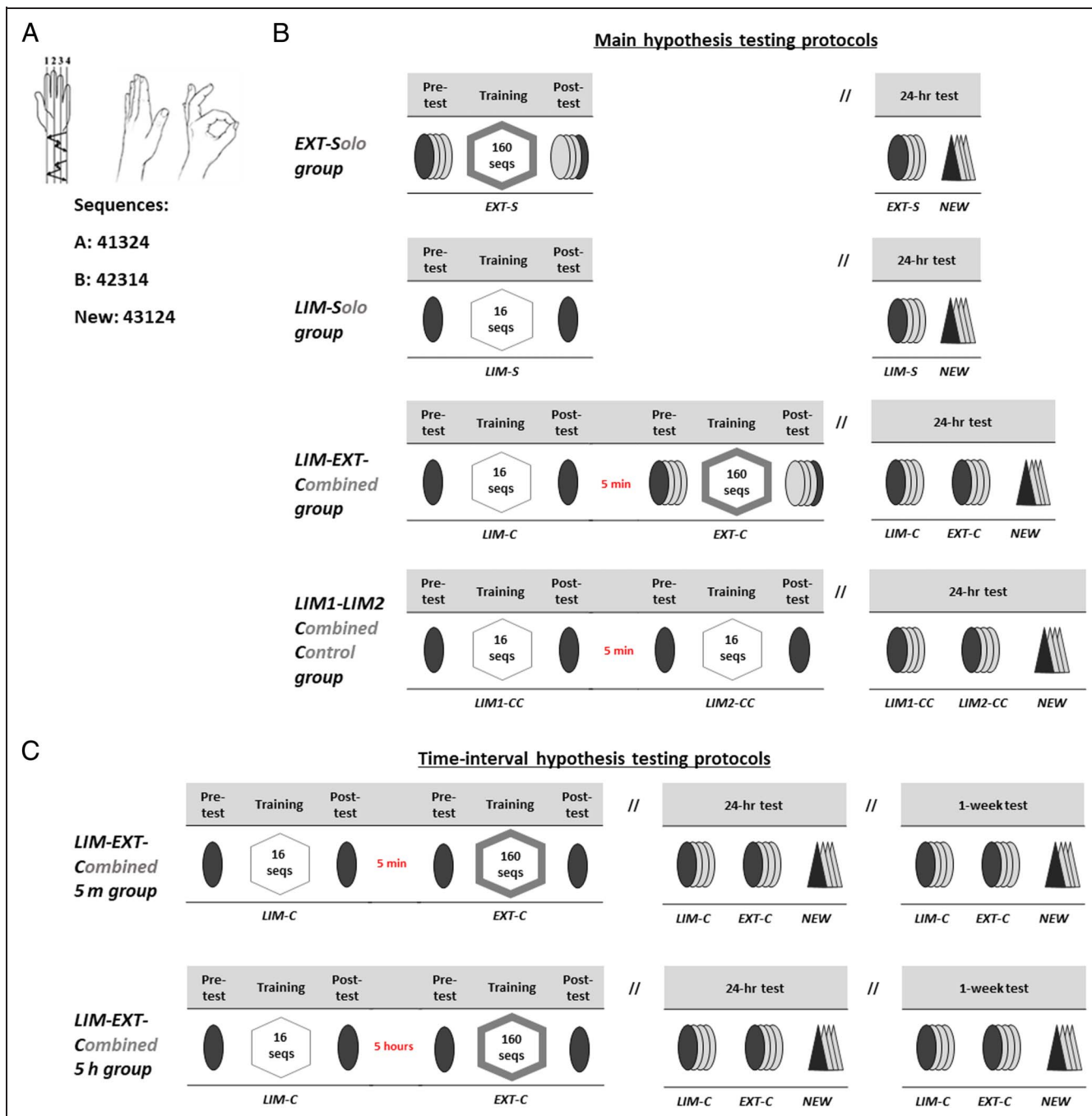


Figure 1. The finger-to-thumb opposition sequence task and training–testing protocol. (A) Finger-to-thumb opposition sequences (A, B, and New). Sequences A and B were matched for number of movements per digit and mirror-reversed in relation to each other (in terms of order). Sequence New had the same component movements as A and B. (B) Main hypothesis testing protocols (four groups, laboratory). (C) Time-interval hypothesis testing protocols (two groups, web-based). Two types of training were afforded—LIM and EXT trainings (16, 160 iterations of the trained movement sequence, respectively). On Day 2, the performance of each of the sequences instructed and trained on Day 1 and of the New sequence were retested. Altogether, the study included six groups of participants: two solo sequence training groups and four combined training groups; participants in each group trained on two sequences, consecutively. In the combined groups, testing and training on LIM-C always preceded training on EXT-C, both on Day 1 and Day 2. In each group, half the participants trained Sequence A as the initially trained (or solo) sequence; all others first trained Sequence B. During training, each sequence was cued; in test blocks, only the first sequence was cued and participants were asked to continuously execute as many sequences as possible in a 30-sec interval. Note that to limit the number of task iterations in LIM training, only one test block was afforded. Black shapes—the test blocks used in the analysis.

et al., 2003, 2007) that have demonstrated (a) the expression of overnight delayed gains and (b) the interference effect in EXT-EXT training, the EXT sequence testing protocol was kept standard. However, because the number of

test blocks in the pre and post training tests was reduced to one in the LIM training condition, all comparisons presented in the current article, between groups and between time-points, were performed on the first blocks of the

performance tests (common to both LIM and the EXT conditions). Note that EXT training in both the solo and the combined conditions also afforded additional experience on the EXT sequence (on average, 15.14 and 17.29 additional sequence iterations per session) compared with the LIM sequence, because of the additional test blocks at pretest and posttest.

The performance was video-recorded. Performance scores of all participants were computed by the same skilled rater offline by observation of the recordings replayed with reduced frame rate. The scoring for the number of correct sequences and the number of sequencing errors produced during 30-sec test intervals (blocks) was performed twice to ensure robustness. As the main rater was informed about the experimental condition, an additional skilled rater unaware of the test conditions rescored randomly picked trials.

In line with the main hypothesis of the study, one group was subjected to the combination of the LIM and the EXT training experiences (LIM-EXT-C group). On Day 1, this group was first afforded the limited training on the first sequence (LIM-C condition), followed by the extensive training on the second sequence (EXT-C condition). A rest period of 5 min separated between the two training conditions. On Day2, 24 hr posttraining, this group was afforded four blocks of the LIM-C condition, four blocks of the EXT-C condition, and four blocks of a New, untrained, sequence (Sequence C, Figure 1A). A direct control group was subjected to the combination of the LIM1-CC and the LIM2-CC training experiences (LIM1-LIM2-CC group). The training protocol was as in the LIM-EXT-C group. As both trained conditions were limited, on Day 1, the second LIM2-CC had a single pre, a single training, and a single post block.

Two additional groups were trained on a single sequence (solo); one group was given LIM training, and the other group was given EXT training (LIM-S and EXT-S groups, respectively); this afforded a baseline time-course for assessing training-induced gains in performance across the 2 days in comparison to the combined training effects. The training protocol of the LIM-S group was similar to the training afforded in the LIM-C condition, and the training protocol of the EXT-S group was similar to the EXT-C condition. On Day2, 24 hr posttraining, the LIM-S group was retested in four blocks of the LIM-S condition and four blocks of a New, untrained, sequence (Figure 1A and B), whereas the EXT-S group was afforded four blocks of the EXT-S condition and four blocks of a New, untrained, sequence (Figure 1A and B). Performance in all test blocks was quantified using two measures: (1) number of correctly completed sequences per performance block and (2) number of incorrectly completed sequences per performance block (absolute errors).

Because of COVID-19-related social restrictions, two experimental groups were run using a web-based, audio–video communication platform (Zoom, <https://>

zoom.us/). The web-based training protocol was adapted in a way to achieve maximum similarity to the face-to-face standard laboratory setting protocol. The web-based protocol included instructions provided by the experimenter. The participant's performance was recorded under the continuous supervision of the experimenter, as in the original laboratory-based experiment. One web-based group was afforded a 5-min interval between the LIM and the EXT training sequences (LIM-EXT-5 m group; a replication of the LIM-EXT-C condition), and the second web-based group was afforded an interval of 5 hr between the two training sequences (LIM-EXT-5 h group; Figure 5A). To test for long-term effects, an additional time point was added relative to the protocol used in the laboratory setting: Participants were retested also a week after the 24 hr post training retest.

Data Analysis

Statistical analysis was carried out in SPSS v.25 (IBM Corporation). Group-wise comparisons were undertaken as repeated measures general linear model (GLM) with time-points as a within-subject factor and practice condition as a between-subjects factor. In cases of sphericity assumption violation, a Greenhouse–Geisser correction was used. Where appropriate (baseline performance or post hoc to explain interactions in rm-GLM), two-tailed *t* tests were used. Multiple comparisons were conducted using Bonferroni corrections. In all analyses, significance level was set to $\alpha = .05$, two-tailed. The outcomes of the two solo trainings conditions were analyzed using one-way ANOVA with 3 time points. The analyses were performed on performance measures derived from a single test block for each time-point (black shapes; Figure 1B and C).

The number of correctly and incorrectly performed sequences obtained from each test block were analyzed separately. The main measure of performance was the number of correct sequences performed in a test block (30 sec), a measure of performance rate (speed), because participants committed very few errors throughout the study and the improvements in the number correct sequences (performance speed) were not at the cost of more errors (absolute accuracy either did not change across the 24-hr window or even improved). Analyses accuracy (number of incorrect sequences performed in a test block) are therefore presented in the appendixes.

To examine the differences in absolute performance (number of correct sequences) at specific time points, independent-samples *t* tests were used to compare between groups. We also compared between the groups using online and offline gains measures. These were individually calculated for each training sequence. Online gains were calculated as the difference between the number of correct sequences completed in the test blocks before and after the training blocks (post – pre); offline gains were calculated as the difference between the

number of correct sequences performed at 24 hr post-training and after the training blocks (24 hr – post).

RESULTS

Extensive Practice Generates Overnight Gains

We first established that of the two solo training protocols (Figure 1B), only the EXT-solo practice protocol (extensive practice on sequence EXT-S) resulted in significant overnight, delayed, gains in sequence performance; that is, we established that there were additional gains in the performance of sequence EXT-S by 24 hr post training, beyond the gains achieved immediately after the practice session (at post). A repeated measures GLM analysis with three time-points (pre, post, and 24 hr post training) showed a significant effect of Time-Point, $F(2, 60) = 64.174, p < .001, \eta^2 = .681$, indicating that sequence performance speed improved in both the EXT-solo and LIM-solo groups with no costs in accuracy. However, the gains attained in group EXT-S were superior, $F(1, 30) = 11.011, p = .002, \eta^2 = .268$, and there was a significant Time-Point \times Practice Dose interaction, $F(2, 60) = 8.859, p < .001, \eta^2 = .228$, reflecting the fact that delayed performance gains, across the overnight postsession interval (consolidation phase), emerged for sequence EXT-S, $t(15) = -3.656, p = .002, d' = 0.85$, but not for sequence LIM-S—a practice dose effect in the solo training conditions (Figure 2A and B; Appendix 1).

In the Combined Limited and Extensive Training, the Limited-trained Sequence Shows Overnight Gains

Next, in the combined training group, LIM-EXT-C (Figure 1B), we showed that, unlike in the LIM-S, there were significant 24-hr posttraining gains expressed for the LIM-C (Figure 2C). A rm-GLM analysis with three time-points (pre, post, and 24 hr) and two practice conditions (LIM-S and LIM-C) showed a significant Time-Point \times Practice-Condition interaction, $F(2, 60) = 4.756, p = .02, \eta^2 = .137$ (Figure 2B and C). Paired *t*-test analyses confirmed that significant overnight delayed gains in speed emerged (only) in the LIM-C sequence, $t(15) = -5.921, p < .001, d' = 0.92$, with unchanged (i.e., no cost in) accuracy (Appendix 2). A direct comparison of the performance of the two sequences (LIM-C and EXT-C), trained in Group LIM-EXT-combined, showed that there was a significant practice dose effect, $F(1, 30) = 4.176, p = .05, \eta^2 = .122$, with overall superior performance for the EXT-C sequence over the course of learning. However, there was also a significant Time-Point \times Practice-Dose interaction, $F(2, 60) = 15.832, p < .001, \eta^2 = .345$ (Figure 2C) because overnight delayed gains were expressed only for the LIM-C sequence. In fact, the overnight performance of the EXT-C sequence was less robust than the immediate

posttraining performance, $t(15) = 3.061; p = .008, d' = 0.67$ (Figure 2C and D; Appendix 3).

A comparison of the performance of EXT-C to that of EXT-S across the three time-points showed a significant effect of Time-Point, $F(2, 60) = 83.237, p < .001, \eta^2 = .735$ (Appendix 4), that is, overall learning in both conditions, but also a significant Time-Point \times Practice-Dose interaction, $F(2, 60) = 11.686, p < .001, \eta^2 = .280$, reflecting the disadvantage in terms of the overnight gains incurred in the EXT-C sequence (vis-à-vis the significant overnight gains that were attained after the EXT-S training (Figure 2D). Thus, the evolving overnight advantage for LIM-C in the combined training protocol co-occurred with a cost to the overnight performance of EXT-C, a pattern of results that one would expect under the assumption that LIM-C consolidation made use of resources generated for EXT-C. However, there was no significant (negative) correlation between the overnight gains expressed for the LIM-C and EXT-C losses ($r = -.032, p = .908$).

Figure 3 shows that the paradoxical effects of the combined practice protocol (LIM-C gains, EXT-C losses) were highly consistent across the performance of individual participants. When practiced by itself, LIM training (LIM-S) led to overnight improvement (vis-à-vis the immediate post-training test) in 9/16 participants; in the combined training session, 15/16 individuals showed overnight gains in LIM-C, $\chi^2(1) = 6.00, p = .014$. Moreover, although 13/16 participants expressed overnight gains in EXT-S, only 2/16 participants improved their performance overnight in EXT-C, $\chi^2(1) = 15.18, p < .001$. Thus, the overall effect of combined training was that the performance of the two sequences, LIM-C and EXT-C, did not differ by 24 hr after the combined practice protocol (Figure 3B). A one-way ANOVA showed significant differences in the total gains attained for the four sequences—LIM-S, EXT-S, LIM-C, and EXT-C, $F(3, 63) = 5.746, p = .002, \eta^2 = .22$; post hoc Tukey comparisons showed that the total gains for EXT-S (7.94 ± 3.53) were significantly larger than those for LIM-S ($3.56 \pm 3.24; p = .001$) but there was no significant difference between the total gains of LIM-C (5.88 ± 2.53) and EXT-C ($5.76 \pm 2.49; p = .999$; Figure 3B; Appendix 5).

According to our work hypothesis, the delayed gains for the LIM-C were made possible because of the pairing with an extensively trained sequence (EXT-C). One testable corollary of this notion was that no delayed gains were to be expected for consecutive training on two limited-trained sequences (Figure 1C). We directly tested this in an additional group of participants (LIM1-LIM2-combined control) that trained in a combined session that had the same structure as the session afforded the LIM-EXT-combined group in the main experiment, but with both sequences afforded limited training (Figure 4A). A comparison of the overnight consolidation interval gains for LIM-C and LIM1-CC in the two conditions, LIM-EXT-combined and LIM1-LIM2-combined control, respectively, showed significantly larger overnight gains in LIM-C

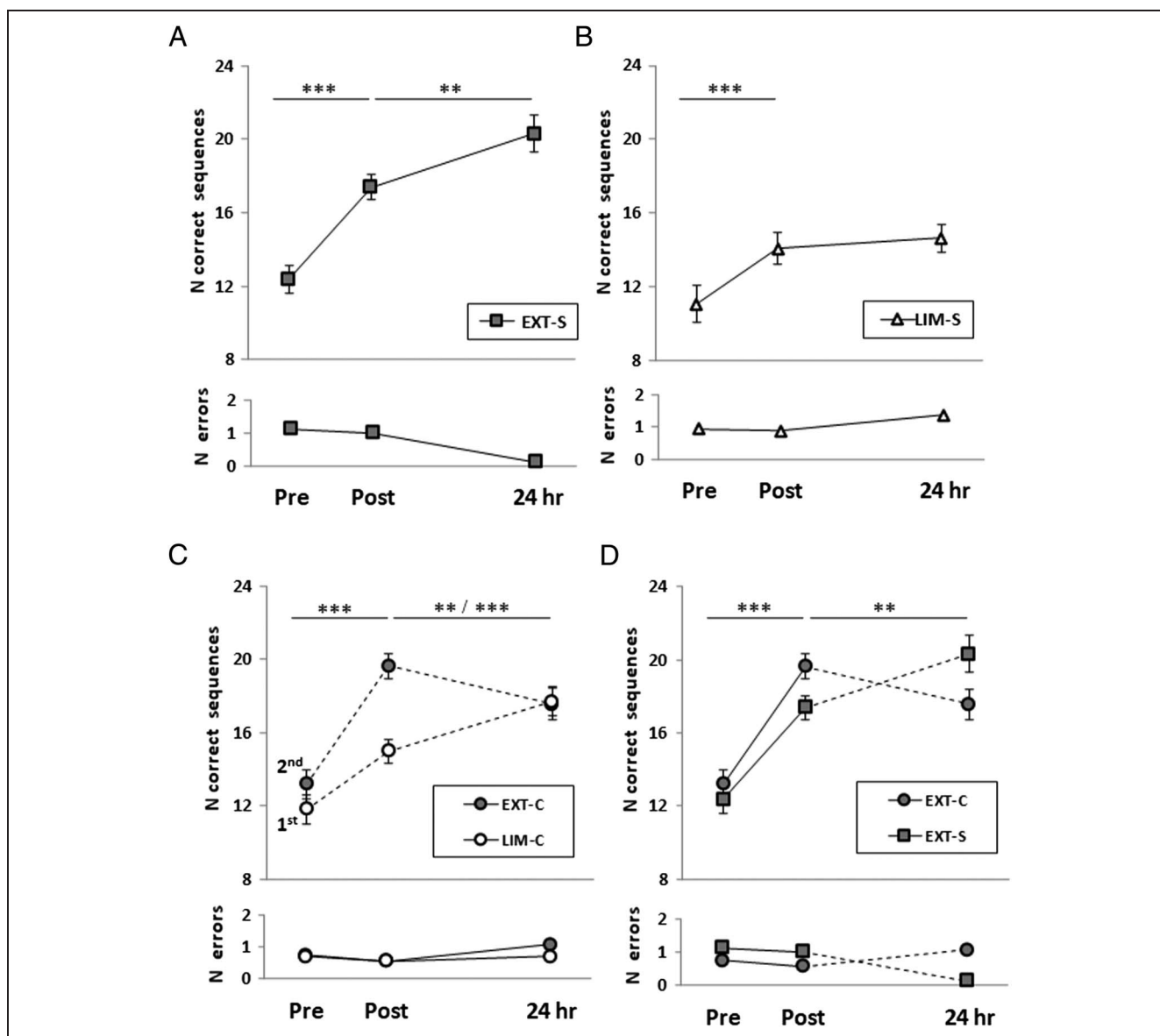


Figure 2. Time-course of changes in performance—number of correct sequences (upper panels) executed in the test and Number of errors (lower panels) committed during the first day's pre and post tests and at the 24 hr post training test. (A) solo extensive training group. (B) Solo limited training group. (C) Within-group direct comparison of performance in the combined experiment training group (group LIM-EXT-combined [laboratory]). There was a significant interaction of Time-Point \times Practice Dose within the session (pre, post) with larger gains expressed for the EXT-C sequence. In addition, there was a significant interaction of Time-Point \times Practice Dose in the postsession interval (post, 24 hr) reflecting additional gains for the LIM-C sequence but a performance decrease for EXT-C. (D) Between-groups comparison of performance for extensively trained sequences in the solo and combined conditions (EXT-S and LIM-EXT-C groups). Both groups showed robust within-session gains (pre, post) for both sequences (EXT-S and EXT-C), but there was a significant interaction of Time-Point \times Practice Dose in the postsession interval (post, 24 hr) reflecting the additional overnight gains for the EXT-S sequence but not for EXT-C. EXT-S but not EXT-C showed in addition an overnight decrease in the number of errors committed (post, 24 hr), that is, an overnight increase in accuracy (Appendix 3). Dashed lines designate significant interactions between practice conditions during the acquisition and consolidation phases. $**p < .01$, $***p < .001$, comparisons across pairs of consecutive timepoints; left to right in correspondence to highest to lower performance at post.

compared with the LIM1-CC, $t(30) = 2.369$; $p = .024$, $d' = 0.84$. Further evidence for the insufficiency of adding a second LIM practice to support (in a combined session) the consolidation gains for a previous LIM-trained sequence is indicated by finding that the gains in LIM1-CC were not significantly different from the gains in LIM-S, that is, when the LIM sequence was trained in a solo session, $t(30) = 1.215$; $p = .234$ (Figure 4B, upper part).

The Time Interval between the Limited and Extensive Training Is Critical

In animal models, synaptic tagging generated in a learning experience is time-limited, on the order of 1–2 hr (Frey & Morris, 1998a, 1998b); a similar time course may be inferred from studies of BT in humans (children; Ramirez Butavand et al., 2020; Ballarini, Martínez, Díaz Perez,

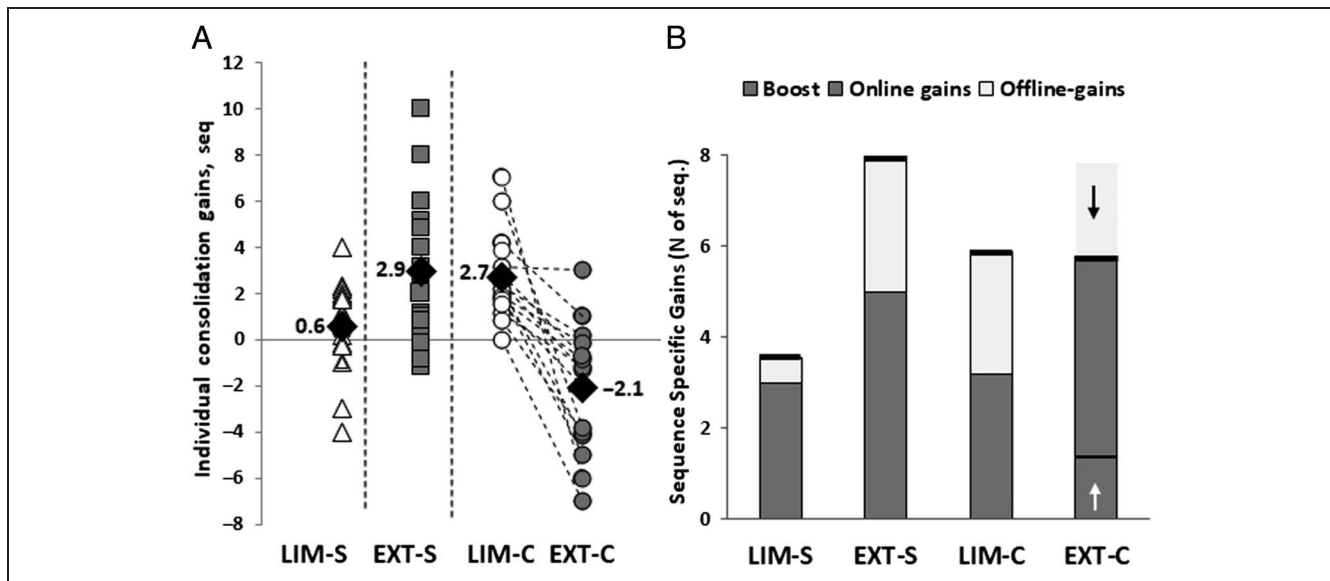


Figure 3. Performance at 24 hr posttraining tests in the solo and combined training conditions. (A) Individual participants' overnight gains in performance (change in the number of correct sequences produced in the 24 hr posttraining test compared with the immediate posttraining test). Black diamond – group mean. (B) Averaged gains during the acquisition (online, dark gray) and the consolidation (offline, light gray) phases. Note that in EXT-C, there was some overnight loss (downward arrow). Upward arrow – baseline shift. By 24 hr post training, there was no significant difference in the performances of LIM-C and EXT-C, 17.69 ± 3.22 ; 17.56 ± 3.41 ; LIM-C and EXT-C, respectively; $t(30) = 0.107$, $p = .916$.

Moncada, & Viola, 2013). To test whether the pairing effects in the combined training protocol occur only within a limited time-window, a second control experiment was used to test the effect of extending the interval (time-window) between training on LIM-C and EXT-C

(Figure 1C). To this end, two additional groups of naive participants were recruited and each group was trained, using a web-based audio–video communication platform, on two FOS sequences in a combined session that had the same structure as the LIM-EXT-combined session

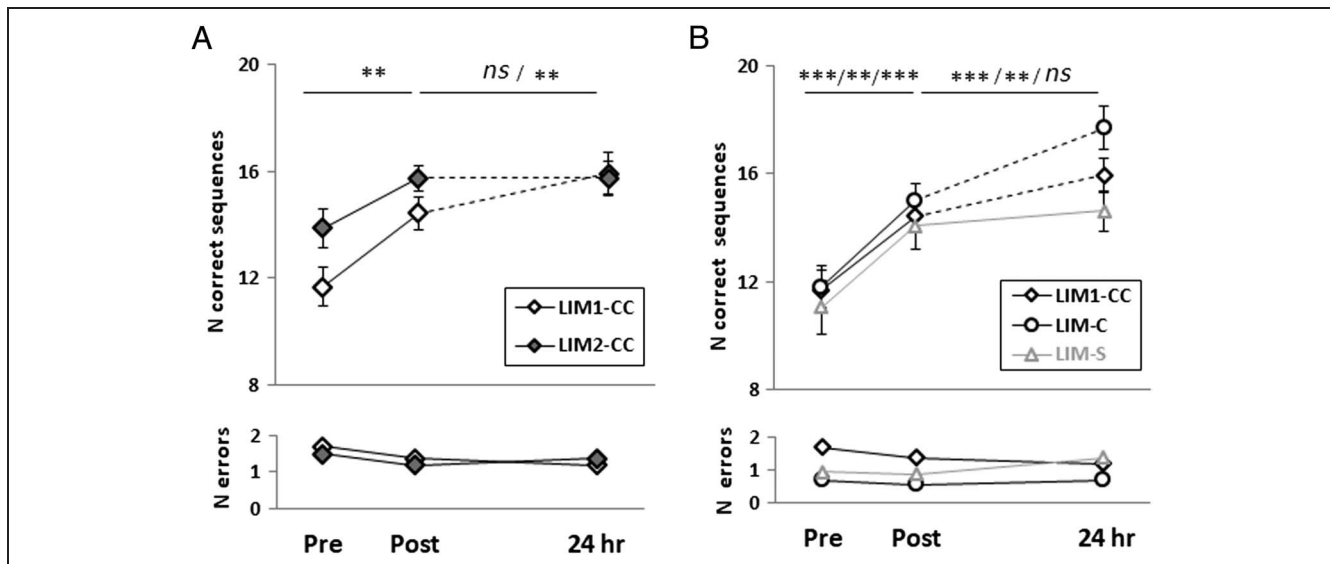


Figure 4. The effects of combining two LIM trained sequences. Number of correct sequences executed in the tests (upper parts) and number of errors (lower parts) during the training session and by 24 hr post training. (A) Combined control training group (LIM1-LIM2-CC). There were additional gains for the LIM1-CC sequence, overnight, but no performance improvement for LIM2-CC. (B) Comparisons between the two initially trained LIM sequences in the two combined training protocols LIM-C versus LIM1-CC (in reference to the LIM-S sequence shown in gray). There was a significant interaction of Time-Point \times Training Condition, for the overnight performance (post, 24 hr) reflecting the more robust gains in LIM-C compared with LIM1-CC, $F(1, 30) = 5.611$, $p = .024$, $\eta^2 = .158$. The overnight consolidation gains expressed in LIM1-CC (1.31 ± 1.45) were significantly smaller than the gains in LIM-C, 2.69 ± 1.81 , $t(30) = 2.369$; $p = .024$, $d' = 0.84$. Dashed lines designate significant interactions between practice conditions during the acquisition and consolidation phases. $**p < .01$, $***p < .001$, left to right in correspondence to highest to lowest performance at 24 hr.

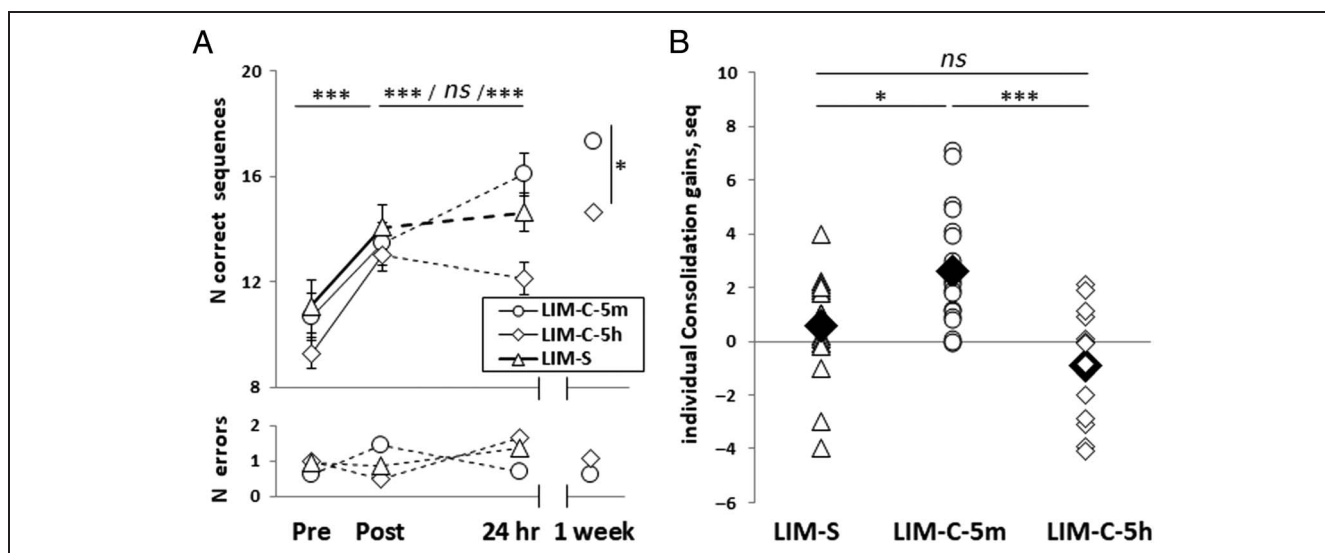


Figure 5. Performance and overnight gains for LIM trained sequences in three study conditions. (A) Number of correct sequences (upper part) and number of errors (lower part) before and after training, at 24 hr post training, and a week later for LIM-C-5 m, and LIM-C-5 h. At 1 week post training, the performance of LIM-C-5 m was superior to that of LIM-C-5 h, $t(34) = 2.495, p = .018, d' = 0.83$. LIM-S was trained in the laboratory, LIM-C-5 m, LIM-C-5 h were obtained in web experiments. (B) Individuals' absolute gains during the 24 hr postsession interval. Black diamond – group mean gains. One-way ANOVA, $F(2, 49) = 13.492, p = .018, \eta^2 = .36$, with pairwise post hoc Tukey comparisons showed that the consolidation gains of LIM-C-5 m (2.61 ± 2.25) were significantly larger than those of LIM-S ($0.56 \pm 2.00; p = .014$) and those of LIM-C-5 h ($-0.89 \pm 1.81; p < .001$); the consolidation gains of LIM-S and LIM-C-5 h did not differ ($p = .104$). * $p < .05$, *** $p < .001$, left to right in correspondence to highest to lower performance at 24 hr.

but with a different interval between the LIM and the EXT training. One group was afforded a 5-min long interval between the LIM and the EXT training events (LIM-EXT-5 m) and constituted a replication of the LIM-EXT-C condition, and the other group was afforded an interval of 5 hr between the two training events (LIM-EXT-5 h; Figures 1C and 5A). To test for long-term effects, an additional time-point was added, and participants were retested 1 week after the 24-hr posttraining assessment.

Changing the mode of training from face-to-face (laboratory) to web-based training (at home) did not affect the time-course of task performance changes in the combined training protocol with a 5-min interval between the LIM and the EXT training (Appendix 6). As can be seen in Figure 5A, significant overnight delayed gains were again generated for the limited-trained sequence, LIM-C-5 m, $t(17) = -4.917, p < .001, d' = 0.77$. However, when the interval afforded between the LIM and the subsequent EXT training was extended to 5 hr, there were significant overnight losses in performance speed for the LIM-C-5 h, $t(17) = 2.082, p = .053, d' = 0.34$, as well as a small increase in the number of errors committed by 24 hr post-training, $t(17) = -3.475, p < .003, d' = 1.07$ (Figure 5A). Moreover, only 4/18 participants improved their overnight performance in LIM-C-5 h (Figure 5B) compared with 15/18 individuals who showed overnight gains in LIM-C-5 m, $\chi^2(1) = 13.49, p < .001$. There were also more overnight gainers in the LIM-C-5 h group compared with overnight gainers in the LIM-S condition in the main experiment, $\chi^2(1) = 4.15, p = .042$.

By 1-week posttraining, although some additional gains were accrued for both sequences, the performance gap (that opened overnight after the two combined training conditions) between the two LIM sequences remained significant; the LIM-C-5 m sequence was performed faster than the LIM-C-5 h sequence, $F(1, 34) = 11.0041, p = .002, \eta^2 = .245$ (Figure 5A).

DISCUSSION

The current study presents evidence for a new interaction (outcome of temporal pairing) between two consecutively learned motor tasks. The interaction occurs when a newly learned movement sequence receives limited practice (LIM), and this is then followed by more extensive practice on a different, equally novel, movement sequence (EXT). In this form, the combined learning-practice session results, paradoxically, in robust benefit to the overnight performance of the LIM sequence – delayed gains unproportional to the amount of actual practice afforded for the sequence. These delayed gains are well-maintained a week later. Two control experiments showed that when the LIM sequence was trained by itself (solo) or when the combined learning-practice session constituted two LIM-trained sequences, there were no significant overnight (delayed) gains for these LIM trained sequences. However, the gains expressed in LIM performance after the LIM-EXT combined session came at a cost; the long-term performance gains that were expected to evolve for the

EXT sequence (as indicated by the robust delayed gains expressed in the EXT solo training condition) were significantly reduced.

Taken together, the current results indicate a hitherto unrecognized long-term outcome of combined training sessions for the acquisition and mastery of how to (procedural) memory, and, specifically, suggest a possible interaction between two novel movement sequences, temporally paired in training, whereby the LIM sequence benefits from a factor or factors made available by the EXT sequence, at a cost to the latter (a “steal” phenomenon). We propose that these results provide the first tier behavioral evidence in support of the STC hypothesis (Maaravi Hesseg et al., 2016; Moncada et al., 2015; Sajikumar, Morris, & Korte, 2014; Fonseca, Nägerl, Morris, & Bonhoeffer, 2004) in relation to human motor (procedural, how to) memory generation. The results are in line with two main assumptions of the STC model: (1) that even a limited experience may suffice to tag the synapses involved in the (repeated) execution of a motor sequence and (2) that a temporal pairing with an extensively trained sequence of movements, performed with the same hand, can lead to the “capture” of critical factors necessary for synaptic remodeling (PRPs) by the (still) tagged synapses of a motor sequence afforded only limited training. Thus, we interpret the current findings as indirect evidence for the notion of a capture of PRPs generated for the EXT sequence by synapses tagged by practice on the LIM sequence. The consistent reduction in the consolidation phase gains expressed for the EXT sequence is in line with the notion of a competition for a limited pool of PRPs suggested by the STC (Maaravi Hesseg et al., 2016; Moncada et al., 2015; Sajikumar et al., 2014; Fonseca et al., 2004).

This STC-inspired interpretation is further supported by the results of the control experiments. The LIM1-LIM2-CC control experiment showed that a combined training session that did not include an extensively trained sequence did not lead to robust gains for LIM trained sequences. Thus, after a combined training session, the enhanced expression of delayed gains for the initially trained LIM sequence was dependent on the affordance of extensive training. The further finding that the LIM enhancement occurred only when the LIM-EXT pairing happened within a short time-window—the interaction effect disappeared when the interval was extended to 5 hr—can also be interpreted in the framework of the STC. The time-course constraint presumably reflects a relatively limited time-window during which LIM-training-related synapses can compete with the EXT-related synapses in capturing PRPs presumably by presenting (still) active tags; cellular level studies have indicated that synaptic tags are short lived (about an hour; Nomoto & Inokuchi, 2018; Nomoto et al., 2016; Moncada & Viola, 2007; Frey & Morris, 1998a, 1998b). The results also show that the pattern of LIM enhancement and relative EXT losses found on Day 2 represents longer-term outcomes of the LIM-EXT combined training protocols; in fact, rather than an attrition,

additional gains were accrued for the LIM sequence by 1 week posttraining. This time-course is again in line with the STC hypothesis in suggesting that the interaction between the two learning experiences has its main effect on mnemonic processes, that is, long-term memory generation. Thus, the current results are compatible with, and extend, the notion that the cell-level synaptic events suggested by the STC can be directly reflected in animal and human behavior, a notion supported by seminal studies of the BT paradigm; when a weak experience is temporally associated with a novel experience, it can turn into a long-term memory (Ramirez Butavand et al., 2020; Ballarini et al., 2013). The BT paradigm, however, was designed, in both animal and human studies, to specifically engage the hippocampal memory system. In addition, the enhancement effect in BT is related to pairing with novel experiences; the novelty of the strong experience, rather than the induction of new learning per se, is considered critical in inducing the generation of PRPs (Nomoto & Inokuchi, 2018; Nomoto et al., 2016; Ballarini et al., 2009, 2013). One problem in moving to the procedural memory domain is that the hippocampus presumably is not required for the generation of skill and long-term maintenance of how to memory for trained movement sequences; the finger-to-thumb opposition sequence learning task, the task used in the current study, can be mastered and consolidated into long-term procedural memory in individuals with hippocampal damage and amnesia (Korman et al., 2018; Döhring et al., 2017). Nevertheless, as the FOS practice starts with a clear explicit cognitive representation of the required component movements and their order, sufficient practice leads to both within-session and between-sessions improvements of motor acuity (Krakauer et al., 2019) and a reduction in the explicit awareness of the sequence (Korman et al., 2003). Indeed, the hippocampus was shown to be only temporally involved in early acquisition-encoding (Albouy, King, Maquet, & Doyon, 2013); hippocampal activity tends to decrease with accumulated practice during the initial phase of motor sequence learning (Albouy et al., 2008). Thus, a different brain hub may provide the grounding for the long-term outcomes of the LIM-EXT interaction described here.

The application of the STC hypothesis to motor learning and memory, specifically FOS learning, must meet two critical requirements to afford a direct competition between the two tasks (FOSs): first, that there is a sizeable population of neurons engaged in the execution of (i.e., shared by) both the two movement sequences during encoding and retrieval, and second, that different (non-overlapping) subsets of synapses in these neurons are (differentially) activated in the training of the two movement sequences, given that a major hub of motor skill acquisition and consolidation is primary motor cortex (Guo et al., 2015; Gabbitov et al., 2014; Xu et al., 2009; Yang et al., 2009; Matsuzaka et al., 2007; Karni et al., 1998), that M1 units co-represent component movements irrespective of the

sequence (though differentially, after extensive practice) providing a common ground for interactions between sequences (Ninokura, Mushiake, & Tanji, 2004; Mushiake, Inase, & Tanji, 1991), and that the generation of procedural memory is possible even when the hippocampal system is dysfunctional (Korman et al., 2018; Döhring et al., 2017). One possible candidate structure that can support a STC-like interaction in motor learning is primary motor cortex and its striatum connections (Gabitov, Manor, & Karni, 2015). However, it may be the case that the common neuronal population relevant for the newly discovered paradoxical mnemonic interaction between LIM and EXT training may extend to other areas such as premotor and parietal cortices (Zimnik & Churchland, 2021; Berlot, Popp, & Diedrichsen, 2020).

Taken together, the current results show that an apparently paradoxical long-term memory augmentation can follow the temporal pairing between “limited” and “extensive” motor sequence training experiences—LIM-wins, EXT-loses. Thus, consecutive training on different motor tasks can result in mnemonic interactions other than interference. The long-term behavioral consequences of a LIM-EXT combined training experience may be explained by a set of synaptic plasticity events and mechanisms similar to those suggested by the STC hypothesis, but presumably occurring outside the hippocampal system. Animal model studies are needed to replicate the behavioral results and then to directly address the underlying neural mechanisms as well as to establish and test competing cell or systems-level explanations for the behavioral phenomenon. Altogether, we propose that the STC hypothesis can be generalized to apply to both long-term memory systems, declarative and procedural; the current findings support the possibility that a similar repertoire of neuronal level mechanisms of plasticity subserves both hippocampus-dependent and hippocampus-independent learning and memory consolidation.

APPENDIX 1. COMPARISONS BETWEEN TWO SOLO PROTOCOLS LIM-S AND EXT-S

Number of Correct Sequences (Performance Speed)

An rm-GLM analysis comparing performance for the groups training on LIM-S and training on EXT-S across three time-points (pre, post, and 24 hr post training) showed a significant effect of Time-Point, $F(2, 60) = 64.174, p < .001, \eta^2 = .681$, indicating that performance improved over the 24-hr period in both groups (training conditions). The Practice Dose was also significant, $F(1, 30) = 11.011, p = .002, \eta^2 = .268$, with overall superior performance for the EXT-S condition. Moreover, there was a significant Time-Point \times Practice Dose interaction, $F(2, 60) = 8.859, p < .001, \eta^2 = .228$, reflecting the different rates of improvement in the groups across the overnight, post-24 hr, posttraining interval. Paired t tests

showed that EXT-S resulted in significant gains not only across the practice session, within-session gains, $t(15) = -6.594, p < .001, d' = 1.75$, but also in the 24-hr posttraining interval, between-sessions gains, $t(15) = -3.656, p = .002, d' = 0.85$. In contrast, although training with the LIM-S resulted in within-session gains, $t(15) = -5.132, p < .001, d' = 0.81$, there were no significant gains in the 24-hr posttraining interval, $t(15) = -1.126, p = .278$. Note that all significant results remained significant ($p_{\text{crit}} = .025$) also after Bonferroni correction for multiple comparisons.

Number of Errors

The gains in performance speed were not at the cost of accuracy (no speed-accuracy trade-off). An rm-GLM analysis with three time-points (pre, post, and 24 hr) showed no significant effect of Time-Point in the number of errors committed, $F(2, 60) = 0.892, p = .415, \eta^2 = .029$, and no significant effect of Practice Dose, $F(1, 30) = 1.216, p = .279, \eta^2 = .039$. However there was a significant Time-Point \times Practice Dose interaction effect, $F(2, 60) = 7.181, p = .002, \eta^2 = .193$. The interaction reflected that participants made fewer errors at 24 hr post training, but only in the EXT-S condition. Paired t -test analysis for EXT-S sequence showed a significant decrease in the number of errors in the 24-hr posttraining interval, $t(15) = 2.907, p = .011, d' = 0.90$. In contrast, in the LIM-S sequence, the number of error tended to increase overnight, $t(15) = -2.236, p = .041, d' = 0.49$.

APPENDIX 2. A COMPARISON OF ACCURACY IN THE SOLO AND COMBINED PROTOCOLS IN REFERENCE TO THE LIM SEQUENCE (LIM-S, LIM-C)

Number of Errors

An rm-GLM analysis with three time-points (pre, post, and 24 hr) showed no significant effect of Time-Point, $F(2, 60) = 1.537, p = .223, \eta^2 = .049$; no significant effect of Practice Condition, $F(1, 30) = 2.353, p = .136, \eta^2 = .073$; and no significant interaction of Time-Point \times Practice Condition, $F(2, 60) = 0.837, p = .438, \eta^2 = .027$.

APPENDIX 3. COMPARISONS BETWEEN PERFORMANCE ON LIM AND EXT SEQUENCES IN THE COMBINED PROTOCOL (LIM-C AND EXT-C)

Number of Correct Sequences

An rm-GLM analysis with three time-points (pre, post, and 24 hr post training) showed a significant effect of Time-Point across the two sequences, $F(2, 60) = 88.657, p < .001, \eta^2 = .747$, indicating that performance improved over the 24-hr period in both groups. The Practice Dose was also significant, $F(1, 30) = 4.176, p = .05, \eta^2 = .122$,

reflecting overall superior performance for the EXT-C condition over the course of learning. Moreover, there was a significant Time-Point \times Practice Dose interaction, $F(2, 60) = 15.832, p < .001, \eta^2 = .345$, reflecting the different rates of improvement in the groups.

Paired t -test analyses conducted for LIM-C and EXT-C, $t(15) = -8.091, p < .001, d' = 2.17$, showed that both sequences resulted in significant within-session gains. However, LIM-C also showed significant gains evolved overnight, in the 24-hr posttraining interval (between-sessions gains; see Data S2), whereas EXT-C showed an overnight decrease in performance, $t(15) = 3.061; p = .008, d' = 0.67$, that is, no overnight gains.

An rm-GLM analysis with two time-points (post and 24 hr post training) showed no significant effect of Time-Point, $F(1, 30) = 0.592, p = .448, \eta^2 = .019$. There was a significant Practice Dose effect, $F(1, 30) = 5.260, p = .029, \eta^2 = .149$, reflecting overall superior performance for the EXT-C condition over the course of learning. Moreover, the Time-Point \times Practice Dose interaction was also significant, $F(1, 30) = 34.191, p < .001, \eta^2 = .533$, reflecting performance improvement in the LIM-C group compared with a performance decrease in the EXT-C group.

Number of Errors

An rm-GLM analysis with three time-points (pre, post, and 24 hr) showed no significant effect of Time-Point in the number of errors committed, $F(2, 60) = 1.474, p = .237, \eta^2 = .047$; no significant effect of Practice Dose, $F(1, 30) = 0.439, p = .513, \eta^2 = .014$; and no significant Time-Point \times Practice Dose interaction effect, $F(2, 60) = 0.609, p = .547, \eta^2 = .020$ (Figure 2C). Paired t -test analyses conducted for the LIM-C sequence showed (like EXT-C) no change in accuracy across the practice session, $t(15) = 0.415, p = .684$.

However, unlike EXT-C, which showed a trend for an increase in the number of errors committed overnight, this was followed by no change in accuracy overnight, in the 24-hr posttraining interval, $t(15) = -0.620, p = .544$.

An rm-GLM analysis with two time-points (pre and post) showed no significant effect of Time-Point in the number of errors committed, $F(1, 30) = 0.614, p = .440, \eta^2 = .020$; no significant effect of Practice Dose, $F(1, 30) = 0.016, p = .900, \eta^2 = .001$; and no significant Time-Point \times Practice Dose interaction effect, $F(1, 30) = 0.025, p = .877, \eta^2 = .001$. An rm-GLM analysis with two time-points (post and 24 hr post training) showed a significant effect of Time-Point in the number of errors committed, $F(1, 30) = 4.310, p = .047, \eta^2 = .126$, indicating that the number of errors committed over the 24-hr period increased in both groups. There was no significant effect of Practice Dose, $F(1, 30) = 0.722, p = .402, \eta^2 = .023$, and no significant Time-Point \times Practice Dose interaction effect, $F(1, 30) = 1.552, p = .223, \eta^2 = .049$.

APPENDIX 4. COMPARISONS BETWEEN THE EXTENSIVE TRAINED SEQUENCES IN THE SOLO AND COMBINED PROTOCOLS (EXT-S, EXT-C)

Number of Correct Sequences

An rm-GLM analysis with three time-points (Pre1, Post4, and 24 hr post training) showed that there was no significant Practice Condition effect, $F(1, 30) = 0.012, p = .914, \eta^2 = .000$. There was a significant effect of Time-Point, $F(2, 60) = 83.237, p < .001, \eta^2 = .735$, indicating that performance improved over the 24-hr posttraining interval in both conditions. Moreover, there was a significant interaction of Time-Point \times Practice Condition, $F(2, 60) = 11.686, p < .001, \eta^2 = .280$, reflecting the larger overnight gains in the EXT-S sequence. Paired t test analyses conducted for the EXT-C sequence showed within-session improvements, but unlike the EXT-S, this was followed by a between-session decrease in performance, that is, no overnight delayed gains in EXT sequence performance (Figure 2D). Note that all significant results remained significant ($p_{\text{crit}} = .025$) also after Bonferroni correction for multiple comparisons.

An rm-GLM analysis with two time-points (pre and post) showed a significant effect of Time-Point, $F(1, 30) = 108.285, p < .001, \eta^2 = .783$, indicating that performance improved over the 24-hr period in both groups. The Practice Dose was not significant, $F(1, 30) = 3.086, p = .089, \eta^2 = .093$. Moreover, there was no significant Time-Point \times Practice Dose interaction, $F(1, 30) = 1.710, p = .201, \eta^2 = .054$ (Figure 2D). An rm-GLM analysis with two time-points (post and 24 hr post training) showed no significant effect of Time-Point, $F(1, 30) = 0.696, p = .411, \eta^2 = .023$, and no significant effect of Practice Condition, $F(1, 30) = 0.059, p < .809, \eta^2 = .002$. However, there was a significant Time-Point \times Practice Dose interaction, $F(1, 30) = 22.738, p < .001, \eta^2 = .431$, reflecting an overnight performance improvement in the EXT-S group compared with a performance decrease in the EXT-C group (Figure 2D).

Number of Errors

An rm-GLM analysis with three time-points (pre, post, and 24 hr) showed no significant effect of Time-Point on the number of errors committed, $F(2, 60) = 1.299, p = .280, \eta^2 = .041$, and no significant effect of Practice Condition, $F(1, 30) = 0.032, p = .860, \eta^2 = .001$. However, there was a significant interaction of Time-Point \times Practice Condition, $F(2, 60) = 6.608, p = .003, \eta^2 = .181$. The interaction indicated that (only) in the EXT-S condition participants made fewer errors at 24 hr post training. Paired t -test analyses conducted for EXT-C sequence showed no change in accuracy across the practice session, $t(15) = 0.716, p = .485$. However, there was a trend for an increase in the number of errors committed overnight, $t(15) = -2.236, p = .041, d' = 0.67$.

An rm-GLM analysis with two time-points (pre and post) showed no significant effect of Time-Point in the number of errors committed, $F(1, 30) = 0.449, p = .508, \eta^2 = .015$;

no significant effect of Practice Dose, $F(1, 30) = 1.776, p = .193, \eta^2 = .056$; and no significant Time-Point \times Practice Dose interaction effect, $F(1, 30) = 0.018, p = .894, \eta^2 = .001$ (Figure 2D). An rm-GLM analysis with two time-points (post and 24 hr post training) showed no significant effect of Time-Point in the number of errors committed, $F(1, 30) = 1.000, p = .325, \eta^2 = .032$, and no significant effect of Practice Dose, $F(1, 30) = 1.008, p = .323, \eta^2 = .033$. However, there was a significant Time-Point \times Practice Dose interaction effect, $F(1, 30) = 13.444, p = .001, \eta^2 = .309$. The interaction indicated that only in the EXT-S condition, participants made fewer errors at 24 hr post training.

APPENDIX 5. COMPARISONS BETWEEN THREE LIM SEQUENCES IN SOLO, COMBINED AND COMBINED CONTROL PROTOCOLS (LIM-S, LIM-C AND LIM1-CC)

Number of Correct Sequences

One-way ANOVA showed a significant difference in the overnight consolidation gains for the LIM sequence in the three conditions—LIM-solo, LIM-combined, and LIM1-combined control, $F(2, 45) = 5.939, p = .005, \eta^2 = .21$. Post hoc Tukey comparisons showed that the overnight consolidation gains of LIM-C (2.69 ± 1.81) were significantly larger than those of LIM-S ($0.56 \pm 2.00; p = .004$). The gains of LIM1-CC (1.31 ± 1.45) were not significantly different from the gains of LIM-C ($p = .082$) and those of LIM-S ($p = .460$).

Number of Errors

An rm-GLM analysis with two time-points (post and 24 hr) and three training conditions (LIM-S, LIM-C, and LIM1-CC) showed a significant effect of Training Condition, $F(2, 45) = 3.416, p = .042, \eta^2 = .132$, reflecting overall superior performance (less errors made) for the LIM-C condition over the course of learning. There was no significant effect of Time-Point in the number of errors committed, $F(2, 90) = 0.504, p = .606, \eta^2 = .011$, and no significant Time-Point \times Training Condition interaction effect, $F(4, 90) = 1.206, p = .314, \eta^2 = .051$.

APPENDIX 6. A COMPARISON BETWEEN LIMITED AND EXTENSIVE TRAINED SEQUENCES IN THE TWO COMBINED-TRAINING PROTOCOL VERSIONS (LABORATORY AND ZOOM BASED)

Number of Correct Sequences

First performance on the LIM sequence in the two experiments (laboratory and Zoom) was compared. An rm-GLM analysis with three time-points (pre, post, and 24 hr post training) showed a significant effect of Time-Point in both protocol versions, $F(2, 64) = 108.662, p < .001, \eta^2 = .773$.

The Experiment Condition (protocol version) was not significant, $F(1, 32) = 1.935, p = .174, \eta^2 = .057$. Moreover, there was no significant Time-Point \times Experiment Condition interaction, $F(2, 64) = 0.234, p = .754, \eta^2 = .007$.

For the EXT sequence, an rm-GLM analysis with three time-points (pre, post, and 24 hr post training) showed a significant effect of Time-Point, $F(2, 64) = 91.166, p < .001, \eta^2 = .740$. The Experiment Condition was not significant, $F(1, 32) = 1.996, p = .167, \eta^2 = .059$. Moreover, there was no significant Time-Point \times Experiment Condition interaction, $F(2, 64) = 1.978, p = .147, \eta^2 = .058$.

Number of Errors

No significant differences were found either in relation to the LIM or the EXT sequences.

Acknowledgments

We thank Dr. Rinatia Maarvi-Hesseg, Yaseen Awad-Igbaria, and Nebal Egbarieh for their assistance in data collection and data management.

Corresponding author: Maria Korman, Occupational Therapy Department, Faculty of Health Sciences, Lab 30.3.18, Kampus Milkun, Ariel University, Ariel, Ramat HaGolan St 65, Ariel, 40700, Israel, or via e-mail: maria.korman@ariel.ac.il.

Data Availability Statement

Additional data related to this article may be requested from the authors.

Author Contributions

Yaniv Oren: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Visualization; Writing—Original draft; Writing—Review & editing. Maria Korman: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Supervision; Validation; Visualization; Writing—Original draft; Writing—Review & editing. Avi Karni: Conceptualization; Data curation; Formal analysis; Investigation; Methodology; Supervision; Validation; Visualization; Writing—Original draft; Writing—Review & editing.

Diversity in Citation Practices

Retrospective analysis of the citations in every article published in this journal from 2010 to 2021 reveals a persistent pattern of gender imbalance: Although the proportions of authorship teams (categorized by estimated gender identification of first author/last author) publishing in the *Journal of Cognitive Neuroscience (JoCN)* during this period were M(an)/M = .407, W(oman)/M = .32, M/W = .115, and W/W = .159, the comparable proportions for the articles that these authorship teams cited were M/M = .549,

W/M = .257, M/W = .109, and W/W = .085 (Postle and Fulvio, *JoCN*, 34:1, pp. 1–3). Consequently, *JoCN* encourages all authors to consider gender balance explicitly when selecting which articles to cite and gives them the opportunity to report their article's gender citation balance.

REFERENCES

- Albouy, G., King, B. R., Maquet, P., & Doyon, J. (2013). Hippocampus and striatum: Dynamics and interaction during acquisition and sleep-related motor sequence memory consolidation. *Hippocampus*, 23, 985–1004. <https://doi.org/10.1002/hipo.22183>, PubMed: 23929594
- Albouy, G., Sterpenich, V., Balteau, E., Vandewalle, G., Desseilles, M., Dang-Vu, T., et al. (2008). Both the hippocampus and striatum are involved in consolidation of motor sequence memory. *Neuron*, 58, 261–272. <https://doi.org/10.1016/j.neuron.2008.02.008>, PubMed: 18439410
- Almaguer-Melian, W., Bergado-Rosado, J., Pavón-Fuentes, N., Alberti-Amador, E., Merceron-Martínez, D., & Frey, J. U. (2012). Novelty exposure overcomes foot shock-induced spatial-memory impairment by processes of synaptic-tagging in rats. *Proceedings of the National Academy of Sciences, U.S.A.*, 109, 953–958. <https://doi.org/10.1073/pnas.1114198109>, PubMed: 22215603
- Balas, M., Roitenberg, N., Giladi, N., & Karni, A. (2007). When practice does not make perfect: Well-practiced handwriting interferes with the consolidation phase gains in learning a movement sequence. *Experimental Brain Research*, 178, 499–508. <https://doi.org/10.1007/s00221-006-0757-3>, PubMed: 17091292
- Ballarini, F., Martínez, M. C., Díaz Perez, M., Moncada, D., & Viola, H. (2013). Memory in elementary school children is improved by an unrelated novel experience. *PLoS One*, 8, e66875. <https://doi.org/10.1371/journal.pone.0066875>, PubMed: 23840541
- Ballarini, F., Moncada, D., Martinez, M. C., Alen, N., & Viola, H. (2009). Behavioral tagging is a general mechanism of long-term memory formation. *Proceedings of the National Academy of Sciences, U.S.A.*, 106, 14599–14604. <https://doi.org/10.1073/pnas.0907078106>, PubMed: 19706547
- Berlot, E., Popp, N. J., & Diedrichsen, J. (2020). A critical re-evaluation of fMRI signatures of motor sequence learning. *eLife*, 9, e55241. <https://doi.org/10.7554/eLife.55241>, PubMed: 32401193
- Brashers-Krug, T., Shadmehr, R., & Bizzi, E. (1996). Consolidation in human motor memory. *Nature*, 382, 252–255. <https://doi.org/10.1038/382252a0>, PubMed: 8717039
- Döhring, J., Stoldt, A., Witt, K., Schönfeld, R., Deuschl, G., Born, J., et al. (2017). Motor skill learning and offline-changes in TGA patients with acute hippocampal CA1 lesions. *Cortex*, 89, 156–168. <https://doi.org/10.1016/j.cortex.2016.10.009>, PubMed: 27890324
- Fonseca, R., Nägerl, U. V., Morris, R. G. M., & Bonhoeffer, T. (2004). Competing for memory: Hippocampal LTP under regimes of reduced protein synthesis. *Neuron*, 44, 1011–1020. <https://doi.org/10.1016/j.neuron.2004.10.033>, PubMed: 15603743
- Frey, S., & Frey, J. U. (2008). ‘Synaptic tagging’ and ‘cross-tagging’ and related associative reinforcement processes of functional plasticity as the cellular basis for memory formation. In W. Sossin, J. Lacaille, V. Castellucci, & S. Belleville (Eds.), *Progress in brain research* (Vol. 169, pp. 117–143). Elsevier. [https://doi.org/10.1016/S0079-6123\(07\)00007-6](https://doi.org/10.1016/S0079-6123(07)00007-6)
- Frey, U., & Morris, R. G. (1997). Synaptic tagging and long-term potentiation. *Nature*, 385, 533–536. <https://doi.org/10.1038/385533a0>, PubMed: 9020359
- Frey, U., & Morris, R. G. (1998a). Synaptic tagging: Implications for late maintenance of hippocampal long-term potentiation. *Trends in Neurosciences*, 21, 181–188. [https://doi.org/10.1016/S0166-2236\(97\)01189-2](https://doi.org/10.1016/S0166-2236(97)01189-2), PubMed: 9610879
- Frey, U., & Morris, R. G. (1998b). Weak before strong: Dissociating synaptic tagging and plasticity-factor accounts of late-LTP. *Neuropharmacology*, 37, 545–552. [https://doi.org/10.1016/S0028-3908\(98\)00040-9](https://doi.org/10.1016/S0028-3908(98)00040-9), PubMed: 9704995
- Friedman, J., & Korman, M. (2012). Kinematic strategies underlying improvement in the acquisition of a sequential finger task with self-generated vs. cued repetition training. *PLoS One*, 7, e52063. <https://doi.org/10.1371/journal.pone.0052063>, PubMed: 23272210
- Friedman, J., & Korman, M. (2016). Offline optimization of the relative timing of movements in a sequence is blocked by retroactive behavioral interference. *Frontiers in Human Neuroscience*, 10, 623. <https://doi.org/10.3389/fnhum.2016.00623>, PubMed: 28066205
- Gabitov, E., Manor, D., & Karni, A. (2014). Done that: Short-term repetition related modulations of motor cortex activity as a stable signature for overnight motor memory consolidation. *Journal of Cognitive Neuroscience*, 26, 2716–2734. https://doi.org/10.1162/jocn_a_00675, PubMed: 24893741
- Gabitov, E., Manor, D., & Karni, A. (2015). Patterns of modulation in the activity and connectivity of motor cortex during the repeated generation of movement sequences. *Journal of Cognitive Neuroscience*, 27, 736–751. https://doi.org/10.1162/jocn_a_00751, PubMed: 25390206
- Guo, J.-Z., Graves, A. R., Guo, W. W., Zheng, J., Lee, A., Rodríguez-González, J., et al. (2015). Cortex commands the performance of skilled movement. *eLife*, 4, e10774. <https://doi.org/10.7554/eLife.10774>, PubMed: 26633811
- Hardwick, R. M., Rottschy, C., Miall, R. C., & Eickhoff, S. B. (2013). A quantitative meta-analysis and review of motor learning in the human brain. *Neuroimage*, 67, 283–297. <https://doi.org/10.1016/j.neuroimage.2012.11.020>, PubMed: 23194819
- Karni, A. (1996). The acquisition of perceptual and motor skills: A memory system in the adult human cortex. *Cognitive Brain Research*, 5, 39–48. [https://doi.org/10.1016/S0926-6410\(96\)00039-0](https://doi.org/10.1016/S0926-6410(96)00039-0), PubMed: 9049069
- Karni, A., Meyer, G., Jezard, P., Adams, M. M., Turner, R., & Ungerleider, L. G. (1995). Functional MRI evidence for adult motor cortex plasticity during motor skill learning. *Nature*, 377, 155–158. <https://doi.org/10.1038/377155a0>, PubMed: 7675082
- Karni, A., Meyer, G., Rey-Hipolito, C., Jezard, P., Adams, M. M., Turner, R., et al. (1998). The acquisition of skilled motor performance: Fast and slow experience-driven changes in primary motor cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, 95, 861–868. <https://doi.org/10.1073/pnas.95.3.861>, PubMed: 9448252
- Korman, M., Doyon, J., Doljansky, J., Carrier, J., Dagan, Y., & Karni, A. (2007). Daytime sleep condenses the time course of motor memory consolidation. *Nature Neuroscience*, 10, 1206–1213. <https://doi.org/10.1038/nn1959>, PubMed: 17694051
- Korman, M., Gal, C., Gabitov, E., & Karni, A. (2021). Better later: Evening practice is advantageous for motor skill consolidation in the elderly. *Learning and Memory*, 28, 72–75. <https://doi.org/10.1101/lm.052522.120>, PubMed: 33593924
- Korman, M., Herling, Z., Levy, I., Egbarieh, N., Engel-Yeger, B., & Karni, A. (2017). Background matters: Minor vibratory

- stimulation during motor skill acquisition selectively reduces off-line memory consolidation. *Neurobiology of Learning and Memory*, *140*, 27–32. <https://doi.org/10.1016/j.nlm.2017.02.002>, PubMed: 28189551
- Korman, M., Raz, N., Flash, T., & Karni, A. (2003). Multiple shifts in the representation of a motor sequence during the acquisition of skilled performance. *Proceedings of the National Academy of Sciences, U.S.A.*, *100*, 12492–12497. <https://doi.org/10.1073/pnas.2035019100>, PubMed: 14530407
- Korman, M., Shakkai, S., Cisamariu, K., Gal, C., Maaravi-Hesseg, R., Levy, I., et al. (2018). Atypical within-session motor procedural learning after traumatic brain injury but well-preserved between-session procedural memory consolidation. *Frontiers in Human Neuroscience*, *12*, 10. <https://doi.org/10.3389/fnhum.2018.00010>, PubMed: 29441005
- Krakauer, J. W., Ghez, C., & Ghilardi, M. F. (2005). Adaptation to visuomotor transformations: Consolidation, interference, and forgetting. *Journal of Neuroscience*, *25*, 473–478. <https://doi.org/10.1523/JNEUROSCI.4218-04.2005>, PubMed: 15647491
- Krakauer, J. W., Hadjiosif, A. M., Xu, J., Wong, A. L., & Haith, A. M. (2019). Motor learning. *Comprehensive Physiology*, *9*, 613–663. <https://doi.org/10.1002/cphy.c170043>, PubMed: 30873583
- Maaravi Hesseg, R., Gal, C., & Karni, A. (2016). Not quite there: Skill consolidation in training by doing or observing. *Learning and Memory*, *23*, 189–194. <https://doi.org/10.1101/lm.041228.115>, PubMed: 27084926
- Maquet, P., Laureys, S., Perrin, F., Ruby, P., Melchior, G., Boly, M., et al. (2003). Festina lente: Evidences for fast and slow learning processes and a role for sleep in human motor skill learning. *Learning and Memory*, *10*, 237–239. <https://doi.org/10.1101/lm.64303>, PubMed: 12888539
- Matsuzaka, Y., Picard, N., & Strick, P. L. (2007). Skill representation in the primary motor cortex after long-term practice. *Journal of Neurophysiology*, *97*, 1819–1832. <https://doi.org/10.1152/jn.00784.2006>, PubMed: 17182912
- Miail, R. C., Jenkinson, N., & Kulkarni, K. (2004). Adaptation to rotated visual feedback: A re-examination of motor interference. *Experimental Brain Research*, *154*, 201–210. <https://doi.org/10.1007/s00221-003-1630-2>, PubMed: 14608451
- Moncada, D., Ballarini, F., & Viola, H. (2015). Behavioral tagging: A translation of the synaptic tagging and capture hypothesis. *Neural Plasticity*, *2015*, 650780. <https://doi.org/10.1155/2015/650780>, PubMed: 26380117
- Moncada, D., & Viola, H. (2007). Induction of long-term memory by exposure to novelty requires protein synthesis: Evidence for a behavioral tagging. *Journal of Neuroscience*, *27*, 7476–7481. <https://doi.org/10.1523/JNEUROSCI.1083-07.2007>, PubMed: 17626208
- Mushiake, H., Inase, M., & Tanji, J. (1991). Neuronal activity in the primate premotor, supplementary, and precentral motor cortex during visually guided and internally determined sequential movements. *Journal of Neurophysiology*, *66*, 705–718. <https://doi.org/10.1152/jn.1991.66.3.705>, PubMed: 1753282
- Ninokura, Y., Mushiake, H., & Tanji, J. (2004). Integration of temporal order and object information in the monkey lateral prefrontal cortex. *Journal of Neurophysiology*, *91*, 555–560. <https://doi.org/10.1152/jn.00694.2003>, PubMed: 12968014
- Nomoto, M., & Inokuchi, K. (2018). Behavioral, cellular, and synaptic tagging frameworks. *Neurobiology of Learning and Memory*, *153*, 13–20. <https://doi.org/10.1016/j.nlm.2018.03.010>, PubMed: 29535042
- Nomoto, M., Ohkawa, N., Nishizono, H., Yokose, J., Suzuki, A., Matsuo, M., et al. (2016). Cellular tagging as a neural network mechanism for behavioural tagging. *Nature Communications*, *7*, 12319. <https://doi.org/10.1038/ncomms12319>, PubMed: 27477539
- Okuda, K., Højgaard, K., Privitera, L., Bayraktar, G., & Takeuchi, T. (2021). Initial memory consolidation and the synaptic tagging and capture hypothesis. *European Journal of Neuroscience*, *54*, 6826–6849. <https://doi.org/10.1111/ejn.14902>, PubMed: 32649022
- Ramirez Butavand, D., Hirsch, I., Tomaiuolo, M., Moncada, D., Viola, H., & Ballarini, F. (2020). Novelty improves the formation and persistence of memory in a naturalistic school scenario. *Frontiers in Psychology*, *11*, 48. <https://doi.org/10.3389/fpsyg.2020.00048>, PubMed: 32063874
- Redondo, R. L., & Morris, R. G. M. (2011). Making memories last: The synaptic tagging and capture hypothesis. *Nature Reviews Neuroscience*, *12*, 17–30. <https://doi.org/10.1038/nrn2963>, PubMed: 21170072
- Rožanov, S., Keren, O., & Karni, A. (2010). The specificity of memory for a highly trained finger movement sequence: Change the ending, change all. *Brain Research*, *1331*, 80–87. <https://doi.org/10.1016/j.brainres.2010.03.019>, PubMed: 20298683
- Sajikumar, S., & Frey, J. U. (2004). Late-associativity, synaptic tagging, and the role of dopamine during LTP and LTD. *Neurobiology of Learning and Memory*, *82*, 12–25. <https://doi.org/10.1016/j.nlm.2004.03.003>, PubMed: 15183167
- Sajikumar, S., Morris, R. G. M., & Korte, M. (2014). Competition between recently potentiated synaptic inputs reveals a winner-take-all phase of synaptic tagging and capture. *Proceedings of the National Academy of Sciences, U.S.A.*, *111*, 12217–12221. <https://doi.org/10.1073/pnas.1403643111>, PubMed: 25092326
- Tong, C., Wolpert, D. M., & Flanagan, J. R. (2002). Kinematics and dynamics are not represented independently in motor working memory: Evidence from an interference study. *Journal of Neuroscience*, *22*, 1108–1113. <https://doi.org/10.1523/JNEUROSCI.22-03-01108.2002>, PubMed: 11826139
- Viola, H., Ballarini, F., Martínez, M. C., & Moncada, D. (2014). The tagging and capture hypothesis from synapse to memory. In Z. U. Khan & E. Muly (Eds.), *Progress in molecular biology and translational science* (Vol. 122, pp. 391–423). Elsevier. <https://doi.org/10.1016/B978-0-12-420170-5.00013-1>, PubMed: 24484708
- Vishnoi, S., Raisuddin, S., & Parvez, S. (2016). Behavioral tagging: A novel model for studying long-term memory. *Neuroscience & Biobehavioral Reviews*, *68*, 361–369. <https://doi.org/10.1016/j.neubiorev.2016.05.017>, PubMed: 27216211
- Walker, M. P. (2005). A refined model of sleep and the time course of memory formation. *Behavioral and Brain Sciences*, *28*, 51–64. <https://doi.org/10.1017/S0140525X05000026>, PubMed: 16047457
- Walker, M. P., Brakefield, T., Morgan, A., Hobson, J. A., & Stickgold, R. (2002). Practice with sleep makes perfect: Sleep-dependent motor skill learning. *Neuron*, *35*, 205–211. [https://doi.org/10.1016/S0896-6273\(02\)00746-8](https://doi.org/10.1016/S0896-6273(02)00746-8), PubMed: 12123620
- Xu, T., Yu, X., Perlik, A. J., Tobin, W. F., Zweig, J. A., Tennant, K., et al. (2009). Rapid formation and selective stabilization of synapses for enduring motor memories. *Nature*, *462*, 915–919. <https://doi.org/10.1038/nature08389>, PubMed: 19946267
- Yang, G., Pan, F., & Gan, W.-B. (2009). Stably maintained dendritic spines are associated with lifelong memories. *Nature*, *462*, 920–924. <https://doi.org/10.1038/nature08577>, PubMed: 19946265
- Zimnik, A. J., & Churchland, M. M. (2021). Independent generation of sequence elements by motor cortex. *Nature Neuroscience*, *24*, 412–424. <https://doi.org/10.1038/s41593-021-00798-5>, PubMed: 33619403