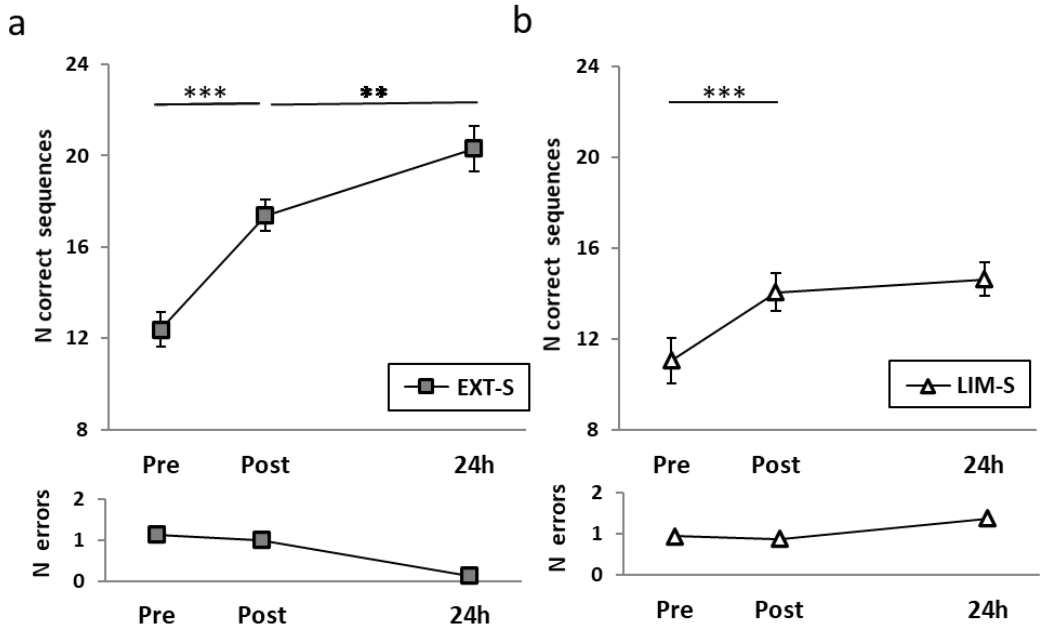
Supplementary Information

Supplementary Information 1. A comparison between two Solo protocols in reference to limited and extensive trainings: LIM-S and EXT-S

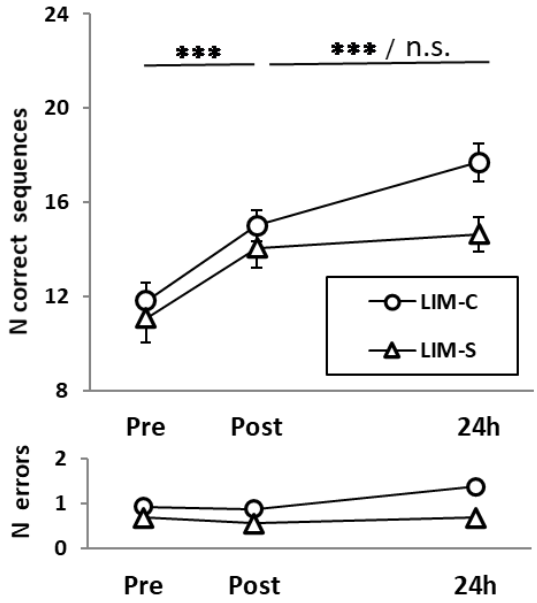


Number of correct sequences (performance speed). An rm-GLM analysis with 3 time-points (Pre, Post and 24h post training) showed a significant effect of time-point **(**$F_{(2,60)}$= 64.174, p <0.001, η^2=0.681), indicating that performance improved over the 24h period in both groups (training conditions). The practice dose was also significant ($F_{(1,30)}$=11.011, p=0.002, η^2=0.268), with overall superior performance for the EXT-S condition. Moreover, there was a significant time-point*practice dose interaction *(*$F_{(2,60)}$=8.859, p<0.001, η^2=0.228), reflecting the different rates of improvement in the groups across the overnight, Post-24h, post-training interval (Figure 2a,b). 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<0.001, d'=1.75) but also in the 24h post-training interval (between-sessions gains) (t(15)=-3.656, p=0.002, d'=0.85) (Figure 2a). In contrast, although training with the LIM-S resulted in within-session gains (t(15)=-5.132, p<0.001, d'=0.81), there were no significant gains in the 24h post-training interval (t(15)=-1.126, p=0.278) (Figure 2b). Note that all significant results remained significant (p_{crit}=0.025) also after Bonferroni correction for multiple comparisons.

Number of Errors. Participants committed overall only a very few errors in the performance of the sequences in all conditions and groups tested, with high accuracy already in the initial (Pre) tests. Moreover, the gains in performance speed were not at the cost of accuracy (no speed-accuracy tradeoff). An rm-GLM analysis with 3 time-points (Pre, Post and 24h) showed no significant effect of time-points in the number of errors committed **(**$F_{(2,60)}$=0.892, p=0.415, η^2=0.029) and no

significant effect of practice dose ($F_{(1,30)}$=1.216, p=0.279, η^2=0.039). However there was a significant time-point*practice dose interaction effect ($F_{(2,60)}$= 7.181, p=0.002, η^2=0.193). The interaction reflected that participants made fewer errors at 24h post training, but only in the EXT-S condition (Figure 2a,b). Paired t-tests analysis for EXT-S sequence showed a significant decrease in the number of errors in the 24h post-training interval (t(15)=2.907, p=0.011, d'=0.90) (Figure 2a). In contrast, in the LIM-S sequence, the number of error tended to increase overnight (t(15)=-2.236, p=0.041, d'=0.49) (Figure 2b).

Supplementary Information 2. A comparison between Solo and Combined protocols in reference to the limited sequence: LIM-S and LIM-C

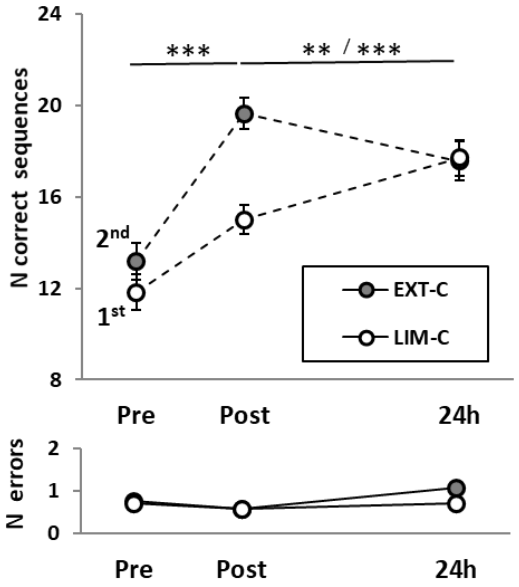


Number of correct sequences. An rm-GLM analysis with 3 time-points (Pre, Post, and 24h post training) showed that there was no significant practice condition effect ($F_{(1,30)}$=2.317, p=0.138, η^2=0.072). There was a significant effect of time-point **(**$F_{(2,60)}$= 66.265, p <0.001, η^2=0.688), indicating that performance improved over the 24h period in both groups. Moreover, there was a significant interaction of time-point*practice condition *(*$F_{(2,60)}$=4.756, p=0.02, η^2=0.137), reflecting the different rates of improvement in the groups (Figure 2b,c). Unlike LIM-S (see Data S1) paired t-test analyses conducted for LIM-C sequence showed significant increase not only in the number of correct sequences performed across the practice session (within-session gains) (t(15)=-6.693, p<0.001, d'=1.11) but also (unlike LIM-S) significant gains evolved overnight (t(15)=-5.921, p<0.001, d'=0.92) (Figure 2c).

Number of errors. An rm-GLM analysis with 3 time-points (Pre, Post and 24h) showed no significant effect of time-points **(**$F_{(2,60)}$=1.537, p=0.223, η^2=0.049), no significant effect of practice

condition ($F_{(1,30)}$=2.353, p=0.136, η^2=0.073), and no significant interaction of time-point*practice condition ($F_{(2,60)}$= 0.837, p=0.438, η^2=0.027) (Figure 2b,c).

Supplementary Information 3. A comparison between the limited and extensive sequences in the Combined protocol: LIM-C and EXT-C



Number of correct sequences. An rm-GLM analysis with 3 time-points 3 time-points (Pre, Post, and 24h post training) showed a significant effect of time-point across the 2 sequences **(**$F_{(2,60)}$= 88.657, p <0.001, η^2=0.747), indicating that performance improved over the 24h period in both groups. The practice dose was also significant ($F_{(1,30)}$=4.176, p=0.05, η^2=0.122), reflecting overall superior performance for the EXT-C condition over the course of learning. Moreover, there was a significant time-point*practice dose interaction *(*$F_{(2,60)}$=15.832, p<0.001, η^2=0.345), reflecting the different rates of improvement in the groups (Figure 2c). Paired t-tests analyses conducted for LIM-C (See Data S2) and EXT-C (t(15)=-8.091, p<0.001, d'=2.17) showed that both sequences resulted in significant within-session gains. However, LIM-C also showed significant gains evolved overnight, in the 24h post-training interval (between-sessions gains, See Data S2), whereas EXT-C showed an overnight decrease in performance (t(15)=3.061; p=0.008, d'=0.67), i.e., no overnight gains (Figure 2c).

An rm-GLM analysis with 2 time-points (Pre, Post) showed a significant effect of time-point **(**$F_{(1,30)}$= 107.735, p <0.001, η^2=0.782), indicating that performance improved over the 24h period in both groups. The practice dose was also significant ($F_{(1,30)}$=10.603, p=0.003, η^2=0.261), reflecting overall superior performance for the EXT-C condition over the course of learning. Moreover, there was a significant time-point*practice dose interaction *(*$F_{(1,30)}$=12.283, p=0.001, η^2=0.291) (Figure 2c), reflecting the different rates of improvement in the groups (Figure 2c). An rm-GLM analysis

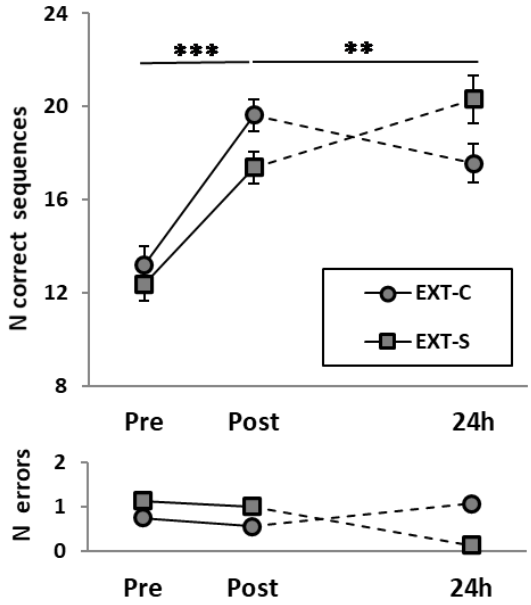
with 2 time-points (Post, 24h post training) showed no significant effect of time-point (**$F_{(1,30)}$**= 0.592, p=0.448, η^2=0.019). There was a significant practice dose effect ($F_{(1,30)}$=5.260, p=0.029, η^2=0.149), reflecting overall superior performance for the EXT-C condition over the course of learning. Moreover, the time-point*practice dose interaction was also significant ($F_{(1,30)}$=34.191, p<0.001, η^2=0.533), reflecting performance improvement in the LIM-C group compared to a performance decrease in the EXT-C group (Figure 2c).

Number of Errors. An rm-GLM analysis with 3 time-points (Pre, Post and 24h) showed no significant effect of time-points in the number of errors committed (**$F_{(2,60)}$**=1.474, p=0.237, η^2=0.047), no significant effect of practice dose ($F_{(1,30)}$=0.439, p=0.513, η^2=0.014), and no significant time-point*practice dose interaction effect ($F_{(2,60)}$= 0.609, p=0.547, η^2=0.020) (Figure 2c). Paired t-test analyses conducted for the LIM-C sequence showed (like EXT-C, see Data S3) no change in accuracy across the practice session (t(15)=0.415, p=0.684).

However, Unlike EXT-C which showed a trend for an increase in the number of errors committed overnight (see Data S3), this was followed by no change in accuracy overnight, in the 24h post-training interval (t(15)=-0.620, p=0.544), (Figure 2c).

An rm-GLM analysis with 2 time-points (Pre, Post) showed no significant effect of time-point in the number of errors committed (**$F_{(1,30)}$**= 0.614, p=0.440, η^2=0.020), no significant effect of practice dose ($F_{(1,30)}$=0.016, p=0.900, η^2=0.001), and no significant time-point*practice dose interaction effect ($F_{(1,30)}$=0.025, p=0.877, η^2=0.001) (Figure 2c). An rm-GLM analysis with 2 time-points (Post, 24h post training) showed a significant effect of time-point in the number of errors committed (**$F_{(1,30)}$**= 4.310, p=0.047, η^2=0.126), indicating that the number of errors committed over the 24h period increased in both groups. There was no significant effect of practice dose ($F_{(1,30)}$=0.722, p=0.402, η^2=0.023), and no significant time-point*practice dose interaction effect ($F_{(1,30)}$=1.552, p=0.223, η^2=0.049) (Figure 2c).

Supplementary Information 4. A comparison between the Extensive trained sequences in the Solo and Combined protocols: EXT-S and EXT-C



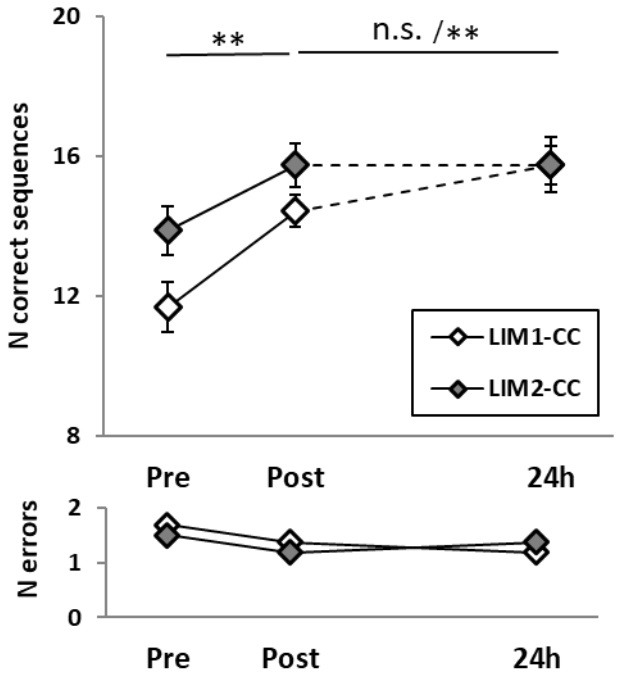
Number of correct sequences. An rm-GLM analysis with 3 time-points (Pre1, Post4, and 24h post training) showed that there was no significant practice condition effect ($F_{(1,30)}$=0.012, p=0.914, η^2=0.000). There was a significant effect of time-point **(**$F_{(2,60)}$= 83.237, p <0.001,η^2 =0.735), indicating that performance improved over the 24h post-training interval in both conditions. Moreover, there was a significant interaction of time-point*practice condition *(*$F_{(2,60)}$=11.686, p<0.001, η^2 =0.280) (Figure 2d), reflecting the larger overnight gains in the EXT-S sequence. Paired t-test analyses conducted for the EXT-C sequence (See Data S3) showed significant within- but unlike the EXT-S (see Data S1) this was followed by an overnight decrease in performance (see Data S3), i.e., no overnight gains (Figure 2d). Note that all significant results remained significant (p_{crit}=0.025) also after Bonferroni correction for multiple comparisons.

An rm-GLM analysis with 2 time-points (Pre, Post) showed a significant effect of time-point **(**$F_{(1,30)}$= 108.285, p <0.001, η^2=0.783), indicating that performance improved over the 24h period in both groups. The practice dose was not significant ($F_{(1,30)}$=3.086, p=0.089, η^2=0.093). Moreover, there was no significant time-point*practice dose interaction *(*$F_{(1,30)}$=1.710, p=0.201, η^2=0.054) (Figure 2d). An rm-GLM analysis with 2 time-points (Post, 24h post training) showed no significant effect of time-point **(**$F_{(1,30)}$= 0.696, p=0.411, η^2=0.023) and no significant effect of practice condition ($F_{(1,30)}$=0.059, p<0.809, η^2=0.002). However, there was a significant time-point*practice dose interaction *(*$F_{(1,30)}$=22.738, p<0.001, η^2=0.431), reflecting an overnight performance improvement in the EXT-S group compared to a performance decrease in the EXT-C group (Figure 2d).

Number of errors. An rm-GLM analysis with 3 time-points (Pre, Post and 24h) showed no significant effect of time-points on the number of errors committed (*F*$_{(2,60)}$=1.299, *p*=0.280, η^2=0.041) and no significant effect of practice condition (*F*$_{(1,30)}$=0.032, *p*=0.860, η^2=0.001) . However, there was a significant interaction of time-point*practice condition (*F*$_{(2,60)}$= 6.608, *p*=0.003, η^2=0.181) (Figure 2d). The interaction indicated that (only) in the EXT-S condition participants made fewer errors at 24h post training (see Data S1). Paired t-test analyses conducted for EXT-C sequence showed no change in accuracy across the practice session (t(15)=0.716, *p*=0.485). However, there was a trend for an increase in the number of errors committed overnight (t(15)=-2.236, *p*=0.041, d'=0.67) (Figure 2d).

An rm-GLM analysis with 2 time-points (Pre, Post) showed no significant effect of time-point in the number of errors committed (*F*$_{(1,30)}$= 0.449, *p*=0.508, η^2=0.015), no significant effect of practice dose (*F*$_{(1,30)}$=1.776, *p*=0.193, η^2=0.056), and no significant time-point*practice dose interaction effect (*F*$_{(1,30)}$=0.018, *p*=0.894, η^2=0.001) (Figure 2d). An rm-GLM analysis with 2 time-points (Post, 24h post training) showed no significant effect of time-point in the number of errors committed (*F*$_{(1,30)}$= 1.000, *p*=0.325, η^2=0.032), and no significant effect of practice dose (*F*$_{(1,30)}$=1.008, *p*=0.323, η^2=0.033). However there was a significant time-point*practice dose interaction effect (*F*$_{(1,30)}$=13.444, *p*=0.001, η^2=0.309). The interaction indicated that only in the EXT-S condition, participants made fewer errors at 24h post training (Figure 2d).

Supplementary Information 5. A comparison between the two Limited sequences in the Combined Control condition: LIM1-CC and LIM2-CC



Number of correct sequences. An rm-GLM analysis with 3 time-points (Pre, Post, and 24h post training) showed no significant training condition effect ($F_{(1,30)}$=1.963, p=0.171, η^2=0.061). There was a significant effect of time-point **(**$F_{(2,60)}$= 44.488, p <0.001,η^2 =0.597), indicating that performance improved over the 24h period in both groups. Moreover, there was a significant time-point* training condition interaction *(*$F_{(2,60)}$=5.544, p=0.010, η^2=0.156), reflecting the different rates of improvement in the groups (Figure 4a). Paired t-tests analyses conducted for LIM1-CC sequence showed significant within-session gains (t(15)=-4.282, p=0.001, d'=1.14) and also a significant increase in the amount of correct sequences performed overnight (between-sessions gains) (t(15)=-3.626; p=0.002, d'=0.65) (Figure 4a). In contrast, although training with the LIM2-CC resulted in within-session gains (t(15)=-4.392, p=0.001, d'=0.71), there were no significant gains in the 24h post-training interval (t(15)=0.000, p=1.000) (Figure 4a).

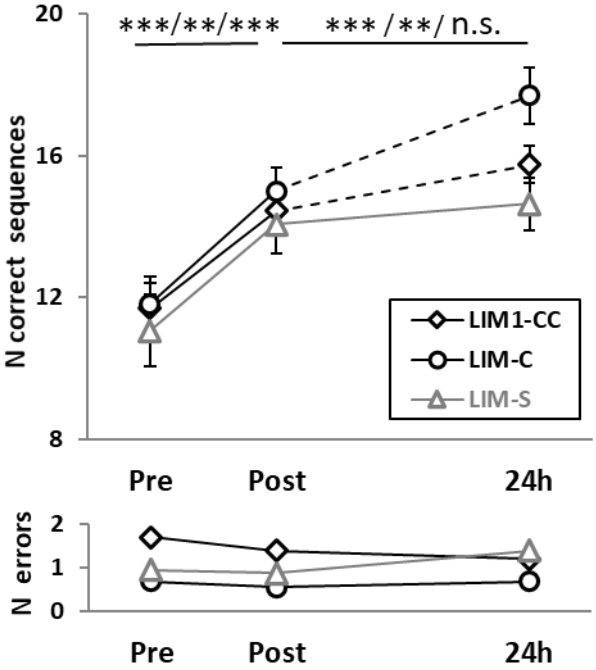
An rm-GLM analysis with 2 time-points (Pre, Post) showed a significant effect of time-point **(**$F_{(1,30)}$= 35.963, p <0.001, η^2=0.545), indicating that performance improved over the 24h period in both groups. The practice dose was also significant ($F_{(1,30)}$=4.700, p=0.038, η^2=0.135), reflecting overall superior performance for the LIM2-CC condition over the course of learning. However, there was no significant time-point*practice dose interaction effect ($F_{(1,30)}$= 1.287, p=0.266, η^2=0.041) (Figure 4a). An rm-GLM analysis with 2 time-points (Post, 24h post training) showed no significant training condition effect ($F_{(1,30)}$=0.624, p=0.436, η^2=0.020). There was a significant effect of time-point **(**$F_{(1,30)}$= 6.517, p =0.016, η^2=0.178), indicating that performance improved over the 24h period in both groups. Moreover, the time-point*training condition interaction was also

significant *(*$F_{(1,30)}$=6.517, *p*=0.016, η^2=0.178), reflecting performance improvement in the LIM1-CC group compared to a performance decrease in the LIM2-CC group (Figure 4a).

Number of Errors. The number of errors committed in the two Combined conditions did not differ (t(30)=0.377, *p*=0.709) (Figure 4a). An rm-GLM analysis with 3 time-points (Pre, Post and 24h) showed no significant effect of time-points in the number of errors committed **(***$F_{(2,60)}$=1.081, *p*=0.346, η^2=0.035**)**, no significant effect of practice dose ($F_{(1,30)}$=0.033, *p*=0.856, η^2=0.001), and no significant time-point*practice dose interaction effect ($F_{(2,60)}$= 0.389, *p*=0.679, η^2=0.013) (Figure 4a). Paired t-tests analyses conducted for LIM1-CC sequence showed no decrease in accuracy across the practice session (t(15)=0.689, p=0.502) and in the number of errors overnight (t(15)= 0.565, p=0.580) (Figure 4a). Similarly, paired t-tests analyses conducted for LIM2-CC sequence showed no decrease in accuracy across the practice session (t(15)=1.431, p=0.173) and in the number of errors overnight (t(15)=-0.643, p=0.530) (Figure 4a).

An rm-GLM analysis with 2 time-points (Pre, Post) showed no significant effect of time-point in the number of errors committed **(***$F_{(1,30)}$= 1.540, *p*=0.224, η^2=0.049), no significant effect of practice dose ($F_{(1,30)}$=0.231, *p*=0.635, η^2=0.008), and no significant time-point*practice dose interaction effect *(*$F_{(1,30)}$=0.000, *p*=1.000, η^2=0.000) (Figure 4a). An rm-GLM analysis with 2 time-points (Post, 24h post training) showed no significant effect of time-point in the number of errors committed **(***$F_{(1,30)}$= 0.000, *p*=1.000, η^2=0.000), no significant effect of practice dose ($F_{(1,30)}$= 0.000, *p*=1.000, η^2=0.000), and no significant time-point*practice dose interaction effect *(*$F_{(1,30)}$=0.720, *p*=0.403, η^2=0.023) (Figure 4a).

Supplementary Information 6. A comparison between three Limited training sequences in Solo, Combined and Combined Control protocols: LIM-S, LIM-C and LIM1-CC



Number of correct sequences. An rm-GLM analysis with 2 time-points (Post, 24h post training) and 2 training conditions (LIM-C, LIM1-CC) showed no significant effect of training condition ($F_{(1,30)}$=2.206, p=0.148, η^2=0.069). There was a significant effect of time-point **(**$F_{(1,30)}$= 47.481, p <0.001,η^2 =0.613), indicating that performance improved over the 24h period in all of the groups. Moreover, there was a significant time-point* training condition interaction *(*$F_{(1,30)}$=5.611, p=0.024, η^2=0.158), reflecting the different rates of improvement in the groups. LIM-S results as reference showed in grey (Figure 4b). Paired t-tests analyses conducted for LIM-S (see Data S1), LIM-C (see Data S2), and LIM1-CC (see Data S5) showed that all sequences resulted in significant within-session gains, while LIM-C and LIM1-CC showed also significant gains evolved overnight, in the 24h post-training interval (between-sessions gains), (Figure 4b).

One-Way Anova showed a significant difference in the overnight consolidation gains for the LIM sequence in the 3 conditions - LIM-solo, LIM-combined, and LIM1-combined control ($F_{(2,45)}$= 5.939, p=0.005, η2=0.21); 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±2.00) (p=0.004). The gains of LIM1-CC (1.31±1.45) were not significantly different from the gains of LIM-C (p=0.082) and those of LIM-S (p=0.460) (Figure 4b).

Number of Errors. An rm-GLM analysis with 2 time-points (Post, 24h) and 2 training conditions (LIM-C, LIM1-CC) showed a significant effect of training condition ($F_{(2,45)}$=3.416, p=0.042, η^2=0.132), reflecting overall superior performance (less errors made) for the LIM-C condition over the course of learning. There was no significant effect of time-points in the number of errors

committed (*F*$_{(2,90)}$=0.504, *p*=0.606, η^2=0.011) , and no significant time-point* training condition interaction effect (*F*$_{(4,90)}$= 1.206, *p*=0.314, η^2=0.051) (Figure 4b).

Supplementary Information 7. A comparison between Limited and Extensive trained sequences in the two Combined protocol versions, Laboratory and Zoom based: LIM-C, EXT-C

Supplementary Information 7.1 We first compared performance on the limited-trained sequence in the two experiments (lab, Zoom). *Number of correct sequences*. a. Limited-trained conditions: An rm-GLM analysis with 3 time-points (Pre, Post, and 24h post training) showed a significant effect of time-point (*F*$_{(2,64)}$= 108.662, *p* <0.001, η^2=0.773). The experiment condition was not significant (*F*$_{(1,32)}$=1.935, *p*=0.174, η^2=0.057). Moreover, there was no significant time-point*experiment condition interaction (*F*$_{(2,64)}$=0.234, *p*=0.754, η^2=0.007). b. Extensively-trained conditions: An rm-GLM analysis with 3 time-points 3 time-points (Pre, Post, and 24h post training) showed a significant effect of time-point (*F*$_{(2,64)}$= 91.166, *p* <0.001, η^2=0.740). The experiment condition was not significant (*F*$_{(1,32)}$=1.996, *p*=0.167, η^2=0.059). Moreover, there was no significant time-point*experiment condition interaction (*F*$_{(2,64)}$=1.978, *p*=0.147, η^2=0.058).

Supplementary Information 7.2 In the zoom experiments, we added a 4th time-point and re-tested the participants for long-term retention one week after the 24h post-training test (1-week). *Number of correct sequences*. An rm-GLM analysis with the 4 time-points of the experiment (Pre, Post, 24h post training, and 1-week) showed a significant effect of time-point in both sequences (*F*$_{(3,102)}$= 82.553, *p* <0.001, η^2=0.708), indicating that performance improved over the course of the experiment. Although the effect of the practice dose was not significant (*F*$_{(1,34)}$=1.293, *p*=0.264, η^2=0.037) the time-point*practice dose interaction was significant (*F*$_{(3,102)}$=15.418, *p*<0.001, η^2=0.312) reflecting the different rates of improvement in the two sequences.

Post-hoc pair-wise comparisons showed that the time-point*practice dose interaction was due to the overnight, 24h post-training, interval. While for LIM-C-5m there was an increase in performance from the Post to the 24h test, p<0.001, the EXT-C-5m showed a significant decrease over the interval, p=0.003. Additionally, while for LIM-C-5m there was an increase in performance from the 24h test to the 1-week later test, p=0.054, the EXT-C-5m showed no significant additional gains over the 1-week interval, p=0.224.

Paired t-test analyses showed that for LIM-C-5m sequence there were significant gains across the practice session (t(17)=-6.320, *p*<0.001, d'=0.77) as well as overnight gains (t(17)=-4.917, *p*<0.001, d'=0.77), and additional gains 1-week later (t(17)=-2.945, *p*=0.009, d'=0.38). Paired t-test analysis conducted for EXT-C-5m sequence showed significant increase in the number of correct sequences

performed across the practice session (within gains) (t(17)=-8.312, p<0.001, d'=1.36), significant decrease overnight (t(17)= 4.025, p=0.001, d'=0.37), and a trend for an increase 1-week later (t(17)=-2.258, p=0.037, d'=0.26). Note that all significant results remained significant (p_{crit}=0.017) after Bonferroni correction for multiple comparisons, except for the improvement over the 1-week interval for EXT-C-5m.

Number of Errors. An rm-GLM analysis with 4 time-points (Pre, Post, 24h post training, and a week after the 24h post training) showed no significant effect of time-points **(**$F_{(3,102)}$= 2.579, p =0.058, η^2=0.070), no significant effect of practice dose ($F_{(1,34)}$=0.081, p=0.778, η^2=0.002), and no significant time-point*practice dose interaction *(*$F_{(3,102)}$=2.272, p=0.085, η^2=0.063), in the number of errors committed.

Supplementary Information 7.3 Next, we performed a comparison between the Extensive and Limited sequences of the Zoom Combined 5h interval protocol: LIM-C-5h and EXT-C-5h.

Number of correct sequences. An rm-GLM analysis with 4 time-points (Pre, Post, 24h post training, and a week after the 24h post training) showed a significant effect of time-point across the 2 sequences **(**$F_{(3,102)}$= 66.716, p <0.001, η^2=0.662), indicating that performance improved over the 24h period in both groups. There was a significant time-point*practice dose interaction *(*$F_{(3,102)}$=3.891, p=0.017, η^2=0.103), reflecting the different rates of improvement in the groups. The practice dose effect was not significant ($F_{(1,34)}$=3.868, p=0.057, η^2=0.102).

Post-hoc pair-wise comparisons showed that the time-point*practice dose interaction was due to the practice session (within-session gains) interval. Both LIM-C-5h (p<0.001) and EXT-C-5h (p<0.001) showed a significant increase from the Pre to the Post test interval. However, the within-session gains of LIM-C-5h were smaller than those of EXT-C-5h (3.72 vs. 4.78, respectively). Additionally, while for LIM-C-5h there was an increase in performance from the 24h test to the 1-week later test, p<0.001, the EXT-C-5h showed no significant change (additional gains) over the 1-week interval, p=1.000.

Paired t-tests analyses conducted for LIM-C-5h sequence showed significant increase in the number of correct sequences performed across the practice session (within gains) (t(17)=-11.581, p<0.001, d'=1.45), significant decrease overnight (between gains) (t(17)= 2.082, p=0.053, d'=0.34) , and significant increase a week after 24h post training (t(17)= -5.284, p<0.001, d'=0.85).

Paired t-tests analyses conducted for EXT-C-5h sequence showed significant increase in the number of correct sequences performed across the practice session (within gains) (t(17)=-9.397, p<0.001, d'=1.70), and significant decrease overnight (between gains) (t(17)= 2.184, p=0.043, d'=0.37). A week after 24h post training there was no loss of performance (t(17)=-0.878, p=0.392). Note that

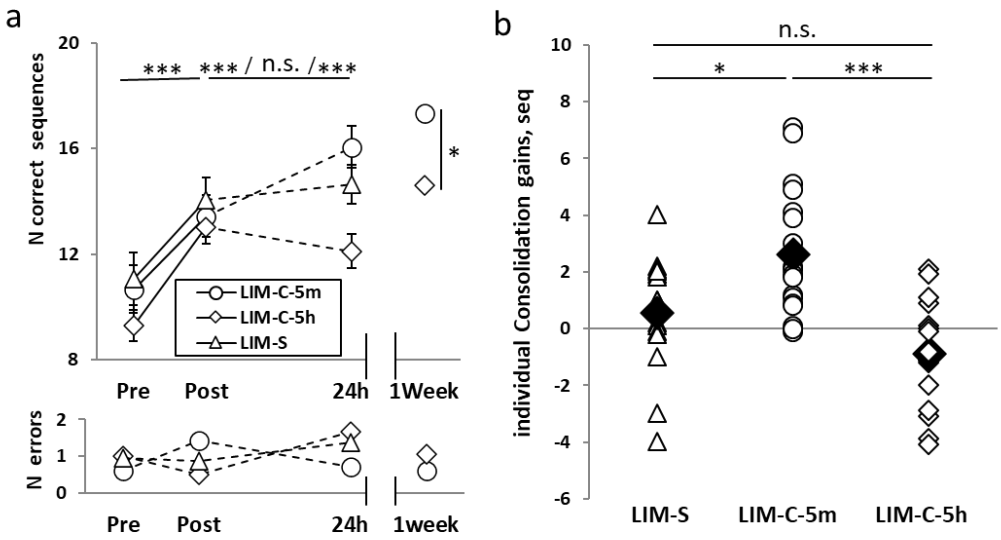
not all significant results remained significant (p_{crit}=0.017) after Bonferroni correction for multiple comparisons.

Number of Errors. An rm-GLM analysis with 4 time-points (Pre, Post, 24h post training, and a week after the 24h post training) showed no significant effect of time-points **(**$F_{(3,102)}$= 0.977, p =0.406, η^2=0.028), and no significant practice dose ($F_{(1,34)}$=0.003, p=0.955, η^2=0.000). There was a significant time-point*practice dose interaction in the number of errors committed *(*$F_{(3,102)}$=3.124, p=0.029, η^2=0.084).

Post-hoc pair-wise comparisons showed that the time-point*practice dose interaction was due to the overnight, 24h post-training, interval. While for LIM-C-5h there was an increase in the number of errors committed from the Post to the 24h test, p=0.017, the EXT-C-5h showed no significant change, p=1.000. Both LIM-C-5h and EXT-C-5h showed no significant change (additional gains) over the 1-week interval, p=0.563; p=1.000, respectively.

Paired t-tests analyses conducted for LIM-C-5h sequence showed no significant difference in the number of errors committed across the practice session (within gains) (t(17)=1.844, p=0.083), significant increase overnight (between gains) (t(17)= -3.475, p=0.003, d'=1.07), and no significant difference in the number of errors committed 1-week later (t(17)= 1.775, p=0.094). Paired t-tests analyses conducted for EXT-C-5h sequence showed no significant differences in the number of errors committed across the practice session (within gains) (t(17)=-0.458, p=0.653), overnight (between gains) (t(17)= 0.946, p=0.357), and 1-week later (t(17)=-0.809, p=0.430).

Supplementary Information 7.4. Finally, we performed a comparison between Limited sequences in three protocols, two Combined Zoom protocols and a Solo laboratory protocol: LIM-C-5m, LIM-C-5h and LIM-S



Number of correct sequences. An rm-GLM analysis with 3 time-points 3 time-points (Pre, Post, and 24h post training) showed no significant effect of training condition ($F_{(2,49)}$=2.372, p=0.104, η^2=0.088). There was a significant effect of time-point **(**$F_{(2,98)}$= 86.413, p <0.001,η^2 =0.638), indicating that performance improved over the 24h period in all of the groups. Moreover, there was a significant time-point* training condition interaction *(*$F_{(4,98)}$=5.828, p=0.001, η^2=0.192), reflecting the different rates of improvement in the groups (Figure 5a). Paired t-tests analyses conducted for LIM-S (see Data S1), LIM-C-5m, and LIM-C-5h (see Data S8) showed that all sequences resulted in significant within-session gains, while only LIM-C-5m showed also significant gains evolved overnight, in the 24h post-training interval (between-sessions gains), (Figure 5a). An rm-GLM analysis with 2 time-points (Pre, Post) showed a significant effect of time-point **(**$F_{(1,49)}$= 147.475, p <0.001, η^2=0.751. There was no significant effect of training condition ($F_{(2,49)}$=0.887, p=0.418, η^2=0.035). Moreover, there was no significant time-point* training condition interaction effect ($F_{(2,49)}$= 1.237, p=0.299, η^2=0.048) (Figure 5a). An rm-GLM analysis with 2 time-points (Post, 24h post training) showed that there was no significant effect of training condition ($F_{(2,49)}$= 2.875, p=0.066, η^2=0.105); there was, however, a significant time-point effect ($F_{(1,49)}$= 7.294, p<0.009, η^2=0.130) and a significant time-point* training condition interaction effect ($F_{(2,49)}$= 13.492, p<0.001, η^2=0.355), reflecting the robust overnight gains in the LIM-C-5m compared to the other two conditions (LIM-S; LIM-C-5h) (Figure 5a).

An rm-GLM analysis with 2 groups (LIM-C-5m, LIM-C-5h), and 2 time-points (24h post training, 1 week post training) showed a significant effect of time-point **(**$F_{(1,34)}$= 34.629, p <0.001, η^2=0.505), and a significant effect of training condition ($F_{(1,34)}$=11.0041, p=0.002, η^2=0.245). However, there was no significant time-point* training condition interaction effect ($F_{(1,34)}$= 3.625, p=0.065, η^2=0.096). Independent-samples t-tests showed that at 1-week post training the performance of LIM-C-5m was superior to that of LIM-C-5h (t(34)=2.495, p=0.018, d'=0.83) (Figure 5a).

Only few of the individual participants showed overnight gains in the LIM-C-5h group compared to more than half in the LIM-S group and the majority in the LIM-C-5m group (Figure 5b). While 9/16 participants showed overnight gains (better performance at 24h post-training compared to performance in the immediate post-training test) in LIM-S, 15/18 individuals showed overnight gains in LIM-C-5m (χ^2(1)= 2.99, p=0.08). In contrast, while 9/16 participants expressed overnight gains in LIM-S, only 4/18 participants improved their performance overnight in LIM-C-5h (χ^2(1)= 4.15, p=0.042). LIM-C-5m vs. LIM-C-5h: χ^2(1)=13.49 P<0.001. One-way anova ($F_{(2,49)}$= 13.492, p=0.018, η^2=0.36) with pair-wise post-hoc Tukey comparisons showed that the consolidation gains of LIM-C-5m (2.61±2.25) were significantly larger than those of LIM-S (0.56±2.00) (p=0.014) and those of LIM-C-5h (-0.89±1.81) (p<0.001); the consolidation gains of LIM-S and LIM-C-5h did not differ (p=0.104) (Figure 5b).

Number of Errors. An rm-GLM analysis with 3 time-points (Pre, Post and 24h) showed that there was no significant effect of training condition ($F_{(2,49)}$=0.166, p=0.847, η^2=0.007); there was, however, a significant time-point effect in the number of errors committed (**(**$F_{(2,98)}$=3.053, p=0.052, η^2=0.059**)**, and a significant time-point* training condition interaction effect ($F_{(4,98)}$= 5.681, p<0.001, η^2=0.188), reflecting the different rates of improvement in the groups (Figure 5a). Paired t-tests analyses conducted for LIM-S (see Data S1) and LIM-C-5h (see Data S8) showed that both sequences resulted in no significant difference in the number of errors committed across the practice session (within gains), while showing significant gains evolved overnight, in the 24h post-training interval (between-sessions gains). In contrast, LIM-C-5m sequence showed significant increase in the number of errors committed across the practice session (within gains) (t(17)=-2.832, p=0.012, d'=0.72), and a significant decrease overnight (between gains) (t(17)= 2.600, p=0.019, d'=0.74), i.e., no overnight gains, (Figure 5a).

An rm-GLM analysis with 2 time-points (Pre, Post) showed that there was no significant time-point effect in the number of errors committed ($F_{(1,49)}$=0.314, p=0.578, η^2=0.006) and no significant effect of training condition ($F_{(2,49)}$=0.456, p=0.636, η^2=0.018). However, there was a significant time-point* training condition interaction effect ($F_{(2,49)}$= 6.161, p=0.004, η^2=0.201), reflecting the different rates of improvement in the groups (Figure 5a). An rm-GLM analysis with 2 time-points (Post, 24h post training) showed that there was no significant time-point effect in the number of errors committed ($F_{(1,49)}$=3.591, p=0.064, η^2=0.068) and no significant effect of training condition ($F_{(2,49)}$=0.014, p=0.986, η^2=0.001). However, there was a significant time-point*training condition interaction effect ($F_{(2,49)}$= 11.520, p<0.001, η^2=0.320), reflecting the different rates of improvement in the groups (Figure 5a).

An rm-GLM analysis with 2 groups (LIM-C-5m, LIM-C-5h), and 2 time-points (24h post training, 1 week post training) showed that there was no significant effect of time-point the number of errors committed **(**$F_{(1,34)}$= 3.066, p =0.089, η^2=0.083), no significant effect of time-point* training condition interaction effect ($F_{(1,34)}$=1.470, p=0.234, η^2=0.041). However, there was a significant training condition effect ($F_{(1,34)}$= 5.809, p=0.021, η^2=0.146) (Figure 5a).

Supplementary Information 8. A comparison between the limited and extensive sequences in the Combined protocol: LIM-C and EXT-C (comparing the 4 test-blocks average)

Number of correct sequences. An rm-GLM analysis with An rm-GLM analysis with 3 time-points (Pre, Post, and 24h post training) showed a significant effect of time-point across the 2 sequences **(**$F_{(2,60)}$= 106.804, p <.001, η^2=.781). There was also a significant practice condition effect ($F_{(1,30)}$=6.636, p<.05, η^2=.181), reflecting overall superior performance for the EXT-C condition over the course of learning. There was a significant time-point*practice condition interaction ($F_{(2,60)}$=12.063, p<.001, η^2=.287), reflecting the differences in the learning curves between the practice conditions (Fig. 2c upper panel). To explore this interaction, we conducted a paired t-test analysis for each practice condition of the LIM-EXT-C group separately. Surprisingly, significant gains were expressed overnight only for the LIM-C trained sequence (15.00±2.61, 18.66±3.23; post, 24h, respectively, p=.000, d'=0.05). There was little or no improvement during the overnight consolidation phase in the performance of the EXT-C condition (18.81±2.45, 18.97±3.38; post, 24h, respectively, p=.762). Additionally, in spite of the difference in the time-course of the evolution of the practice-related gains, by 24h post training, there were no significant differences in the performance of the 2 sequences trained in the LIM-EXT-C group (18.66±3.23; 18.97±3.38; LIM-C and EXT-C, respectively; t(30)=-.267, p=.791).

One-Way Anova showed total gains differences between the 4 sequences - LIM-S, EXT-S, LIM-C, EXT-C (F(3,63)= 4.251, p=0.009, ɳ2=0.18); post-hoc Tukey comparisons showed that the total gains of EXT-S (7.83±3.14) were significantly larger than those of LIM-S (4.50±2.75) (p=0.007) but there was no significant difference between the total gains of LIM-C (6.84±2.78) and EXT-C (5.78±2.35) (p=0.701).