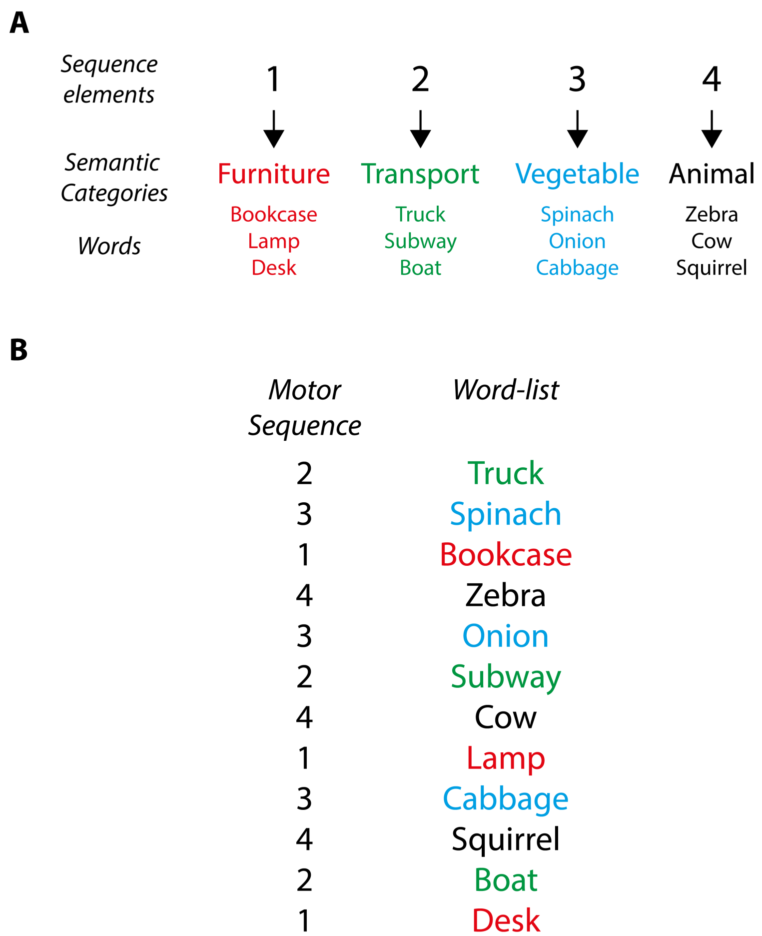

Figure S1. Related to Figure 1; Creating a common structure for a sequence of words and actions. (A) Each of the four elements within the sequence was mapped onto a semantic category, and in turn each category was mapped onto three words. **(B)** Using this mapping, we replaced each element of the motor sequence with a word. With three words in each semantic category, and four semantic categories there was a total of 12-words, which allowed a complete mapping of the motor sequence. When there was no consistent mapping between sequence element and semantic category, the motor sequence and word-list did not have a shared common structure.

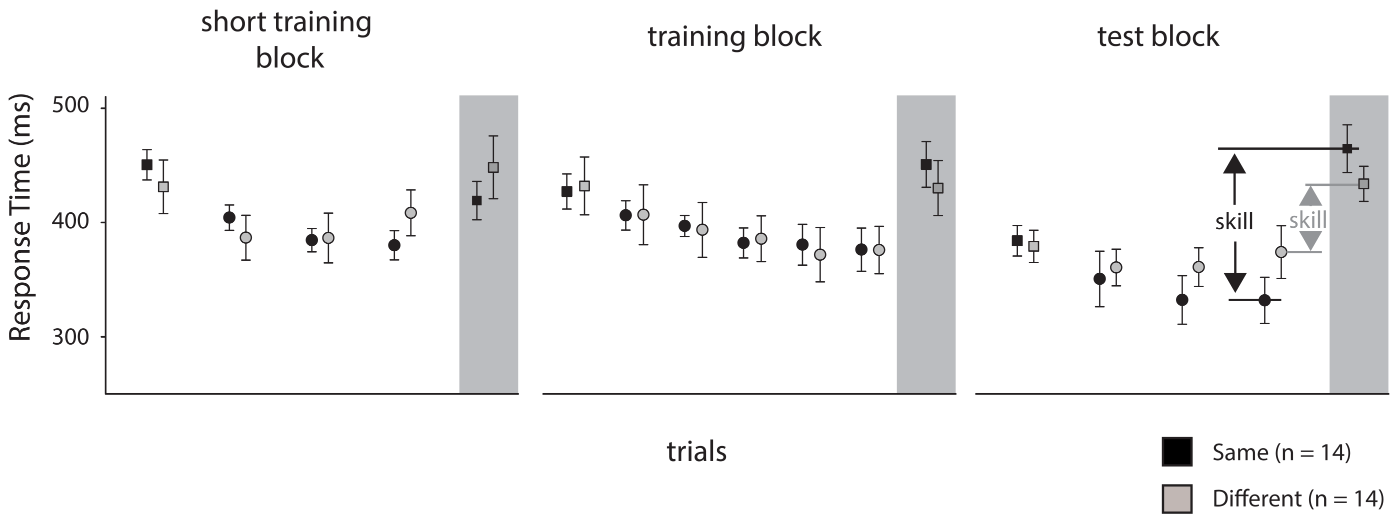

Figure S2, Related to Figure 1; Response times during motor sequence learning following word-list learning. The word-list and subsequent motor sequence task either had the same (black) or different high-level structures (grey). Motor sequence learning took place over an initial short training block, a training block and a subsequent test block. In each block, participants initially performed random trials (square symbol; 50 trial epochs; *mean* ± *sem*), sequential trials (circle symbol; 60 trial epochs; *mean* ± *sem*) and subsequent random trials (square symbol; 50 trial epochs; *mean* ± *sem*, highlighted in grey). The response time during the final sequential trials of the test block showed a trend towards a difference when the earlier word-list and subsequent motor sequence had the same rather than a different structures (unpaired t-test, $t(26) = 1.6$, $p = 0.12$). Response times during the sequential trails are not a specific measure of sequence learning because they are affected by multiple factors including familiarity with the task [S1]. By contrast, the difference in response time between the sequential and subsequent random trials provides both a sensitive and specific measure of sequence learning, which was substantially greater when the word-list and motor learning task had the same rather than different structures (unpaired test, $t(26) = 2.95$, $p = 0.006$). Thus, it was specifically the sequential aspects that were transferred and improved learning in the motor task.

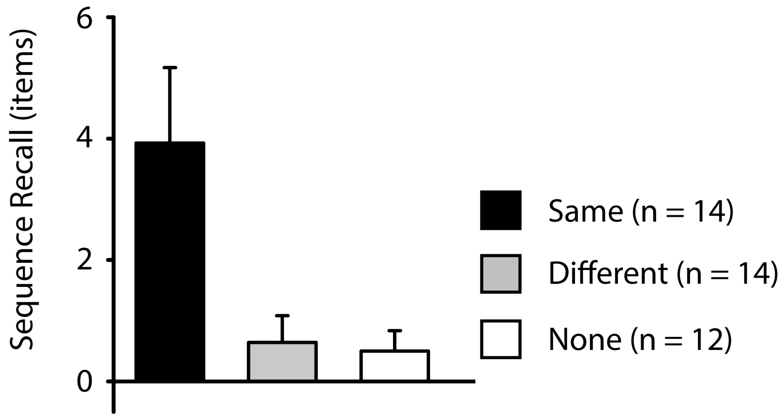

Figure S3, Related to Figure 1; Sequence recall following earlier word-list learning. Declarative recall of the sequence was substantially greater when the word-list and motor sequence shared a common structure than when they had different structures (see main text; unpaired t-test, 4±1.2 vs. 0.6±0.4 items; $t(26) = 2.5$, $p = 0.019$). Similarly, sequence recall was substantially greater when the memory tasks shared a common structure than when the word-list had no consistent structure (4±1.2 items vs. 0.5±0.3 items; $t(24) = 2.48$; $p = 0.02$). These differences remained even when the stringency to be included in the analysis was increased (from ≥3 to ≥4 items of the motor sequence; unpaired t-tests, $t(26) = 2.5$, $p = 0.019$ and $t(24) = 2.92$, $p = 0.007$; respectively). Potentially, recall for the motor sequence may have had a retrograde effect upon the word-list task, affecting subsequent word recall. However, there was no significant difference between the groups in either total or serial word recall (unpaired t-tests; serial recall; 5.7±0.9 vs. 5.3±0.9 words, $t(26) = 0.330$, $p = 0.744$, total recall; 10±0.3 vs. 9.8±0.5 words, $t(26) = 0.371$, $p = 0.713$). Even when we split the group showing the greater sequence recall (when both memory tasks had the same structure) into those with a high (>4-items) and a low sequence recall (≤4-items), those groups showed no significant difference in subsequent word recall (unpaired t-tests; serial recall; 5.6±1.8 vs. 6±1.14, $t(12) = 0.175$, $p = 0.864$, total recall; 10.1±0.4 vs. 10±0.7, $t(12) = 0.15$, $p = 0.876$). Finally, there was also no significant correlation between participants' recall for the sequence, and their recall of the word-list (serial recall; $R = 0.119$, $F(1,26) = 0.374$, $p = 0.546$, total recall; $R = 0.118$, $F(1,26) = 0.369$, $p = 0.549$). Thus, a difference in sequence recall did not lead to a difference in the subsequent recall of the word-list, and so it seems unlikely that sequence recall had an effect upon word-recall.

Supplemental Experimental Procedures

We examined how the transfer of skill or knowledge from one memory task to another (i.e., from task A to task B) is related to both the structural similarities between the tasks, and the interference between the tasks (i.e., impairment in task A recall). Each of our experiments had the same basic design. At 9am, participants learnt one memory task (task A) followed by another (task B), and 12-hrs later, at 9pm, they were retested on the initial memory task (i.e., task A). Learning transfer was measured at the same time of day (i.e., 9am), and the interference between the tasks, a change in performance between testing and retesting, was measured at the same times of day (i.e., 9am and 9pm). These measures may be affected by circadian factors for example, the amount of transfer may differ at different times-of-day. However, each group was subjected to the same circadian factors, and so by comparing across groups, we minimized the circadian affects.

The two memory tasks were of different types (i.e., motor vs. word learning). We compared task B performance across the tasks that either did or did not have the same structure, to assess transfer across the tasks; while, interference was measured as the difference in task A performance between initial testing and subsequent retesting. In our first set of experiments, participants learnt and then recalled a list of words, then acquired skill at performing a motor sequence, had their motor skill tested, and 12-hrs later had their word recall retested. While our second set of experiments was almost identical to the first, except that we reversed the order of the tasks: participants acquired skill at performing a motor sequence, had their motor skill tested, then learnt and recalled a list of words, and 12-hrs later had their motor skill and word recall retested.

Motor sequence learning task

We used a modified version of the serial reaction time task (SRTT; [S1,S2]). A solid circular visual cue (diameter 20mm, viewed from approximately 800mm) could appear at any one of four possible positions, designated 1 to 4, and arranged horizontally on a computer screen. Each of the four possible positions corresponded to one of the four buttons on a response pad (Cedrus, RB-410), upon which the participant's fingers rested. When a target appeared, participants were instructed to respond by pressing the appropriate button on the pad. If the participant made an incorrect response, the stimulus remained until the correct button was selected. Once the correct response was made, the cue on the screen disappeared and was replaced by the next cue after a delay of 400ms. Response time was defined as the interval between presentation of a stimulus and selection of the correct response.

Participants were introduced to the task as a test of reaction time; however, the position of the visual cue followed a repeating 12-item sequence (2-3-1-4-3-2-4-1-3-4-2-1). Participants were not told about the 12-item sequence, and there were no cues marking the introduction of the sequence. The first session consisted of three blocks: an initial, short training block containing 15 repetitions of the sequence, a longer training block with 25 repetitions of the sequence, and a test block containing 15 repetitions of the sequence. There was a second session that took place 12-hrs after the first session: participants performed a retest block containing 15 repetitions of the sequence.

Fifty random trials preceded and followed the sequential trials in the training, test and retest blocks. Within these random trials, there were no item repeats (for example, -1-1- was illegal), and each item had approximately the same frequency of appearance. Each set of random trials in each block was unique, which minimized the chance that participants might become familiar with the random trials.

We administered a free recall test when participants had completed the SRTT. Participants were asked if they had noticed a pattern to the visual cues of the task, and if so, to report verbally as many items of the sequence as possible [S3,S4].

Word-list task

A single word, from a list of 12 words (drawn from the California Verbal Learning Task), was presented on a computer screen for 2s. The word was then removed, and replaced by another word also drawn form the list of 12 words. This process continued until all 12 words had been presented. The same 12 words were presented individually and in the same order for five iterations for each participant. At the end of each of these presentations, participants were asked to recall in order as many of the words as possible (i.e., a serial recall). Participants were not prompted for particular words, nor were they told those words, if any, which they had failed to recall. Following the fifth recall, there was a ten-minute interval after which participants were again asked to recall in order as many of the words as possible. Participants were asked for a final serial recall 12-hrs later.

Motor sequence and word-list structure

We assigned each of the four elements (1, 2, 3, 4) within the motor sequence to one of the four semantic categories in the word-list (transport, vegetable, animal, furniture), which created a consistent shared relationship between sequence element and semantic category (see Figure S1). The 12-item motor sequence (2-3-1-4-3-2-4-1-3-4-2-1) was transformed into a word list (truck-spinach-bookcase-zebra-onion-subway-cow-lamp-cabbage-squirrel-boat-desk), and the two shared a common structure. In another group, there was no consistent mapping between sequence elements and the semantic categories of the words, and so the two tasks did not share a common structure. While in another group, the initial memory task, either the word-list (Experiment 1) or the motor task (Experiment 2), had no repeating structure; for the word list the order of words changed at each iteration, and for the motor task there was a different sequence of elements at each iteration. By comparing across these three groups, it was possible to test how learning one task with the same, different or no consistent sequential structure affected subsequent learning of another task with different elements (i.e., words vs. actions).

Interference between the memory tasks

We used another two groups to modify interference between the tasks. In one group we inserted a 2-hr interval between the tasks: the first task (i.e., task A) continued to be learnt at 9am, while, the second task (i.e., task B) was not learnt until 11am; nonetheless, subsequent retesting on task A continued to occur 12-hrs after its initial acquisition, at 9pm. Inserting an interval is a widely used method for reducing interference between memory tasks [S3,S5]. In the other group, we modified the sequence from being entirely high-order, with the next item of the sequence (n+1) being determined by the current (n) and the previous item (n-1), to being predominately low-order, with the next item (n+1) being determined two thirds of the time by just the current item (n; [S1])). This was achieved by swapping the 9th and 11th items of the sequence, which changed it from 2-3-1-4-3-2-4-1-<u>3</u>-4-<u>2</u>-1 to 2-3-1-4-3-2-4-1-<u>2</u>-4-<u>3</u>-1. Earlier work has shown that this change allows processing of the motor sequence to occur independently of the mediotemporal lobe (MTL), a brain area critical to word-list learning [S6]. Potentially, this reduces the overlap in the circuits supporting motor and word-list learning, which we predicted would reduce interference between the tasks ([S1,S6,S7]; for a review please see [S5]). But, the change or otherwise in the role of the MTL may have very little do with preventing interference between the memory tasks. The change in the transitions between items within the sequence (i.e., the transitional structure) may have other consequences that prevent interference between the tasks. Alternatively, changes in the order of items within the sequence (i.e., ordinal structure) may be responsible for preventing interference between the tasks [S8]. Our work did not set out to identify the mechanism responsible for preventing interference between the tasks by changing the sequence structure. We simply used the change in the structure of the sequence as another method, to complement the insertion of a 2-hr delay, in preventing interference, and so test its importance for learning transfer. To prevent interference from being affected by the motor and word-list task no longer sharing the same structure, we changed the order of elements within both tasks. Our strategy was to use two very different techniques that shared the key common attribute of modifying interference. Thus, any common effect they had upon learning transfer would be attributable to their common effect upon memory stability, and its susceptibility to interference.

In sum, in each experiment there were five groups in which: (a) the memory tasks had a common structure; (b) had a different structure; or (c) the initial memory task (i.e. task A) had no consistent structure (the sequence of elements or words changed at each iteration). While in a further two groups, we sought to alter the interference between tasks, when they had a common structure by: (d) inserting a 2-hr interval between the task or; (e) by modifying the structure of the tasks.

Participants

One hundred and thirty-six (53 male, 20.4±2.3 years; *mean ± std*), right-handed participants were recruited (defined by the Edinburgh handedness questionnaire [S9]). Only healthy participants, 18-35 years of age, with no medical, neurological or psychiatric history, and either normal or corrected to normal vision were recruited to the study. All participants provided informed consent for the study, which was approved by the local institutional review board. We randomly assigned 66 participants to the first set of experiments (Experiment 1; see Figure 1A) with 14 participants assigned to those groups when the tasks had the same, or different structures, or when a 2-hr interval was inserted between the memory tasks, while 12 participants were assigned to each of the remaining two groups. The other 70 participants were allocated to the second set of experiments (Experiment 2; see Figure 1C) with 14 participants being assigned to each of the five groups. All of these participants were included in our analyses. During the interval between testing and subsequent retesting in each of the experiments participants engaged in normal daily activities, but they refrained from napping.

Data Analysis

Response times were defined as the time to make a correct response. Any response time in the top one percentile (i.e., $\alpha = 0.01$) of a participant's data was identified using a Grubbs' Test and removed. We quantified the amount of sequence learning in the motor task (i.e., the SRTT) by subtracting the average response time of the final fifty sequential trials from the average response time of the random trials that immediately followed [S2,S10]. The difference between random and sequential RT is a widely used learning measure, which is both sensitive and specific to learning of the motor sequence (for example; [S2,S10,S11]; for review [S1]). We did not use accuracy as a measure of motor skill because even with limited experience error rates are extremely low (<2-4%, [S10,S12,S13]). The free recall of the motor sequence was scored as the longest, continuous and accurate verbally recalled segment of the sequence that was at least 3-items long (i.e., a triplet or more). For the word-list learning task, we analysed both the total number of words recalled regardless of the order of recall (i.e., total recall), and the longest number of consecutively recalled words in the correct order (i.e., serial recall).

We explored graphically all of the data in MATLAB. Specifically, we examined the distribution of the data using histograms, normal probability plots, and verified that the data followed a normal distribution using the Shapiro-Wilk test.

We used mixed repeated measures ANOVAs to compare learning across the groups. The sphericity of the data was examined using a Mauchy's test. If sphericity was violated, we used a Greenhouse-Geisser correction, which is shown in the main text as a correction to the degrees-of-freedom.

For the motor skill task, learning was the change in performance over the short training, long training and subsequent testing blocking (i.e. learning had three levels); whereas, for the word-list task, this was the change in performance over the five subsequent iterations of the list (i.e., learning had five levels). We also used repeated measures ANOVAs to compare the susceptibility of a memory to interference by comparing the change in performance between testing and subsequent retesting across the groups (i.e., $skill_1$ vs. $skill_2$ or $recall_1$ vs. $recall_2$; for both change in performance had two levels). We used further ANOVAs, when appropriate to compare across groups, and then unpaired t-tests to better understand the differences between the groups. We used paired t-tests to determine the significance of changes within groups. All the t-tests used in the analysis of this study were two-tailed.

Supplemental References

S1. Robertson, E.M. (2007). The Serial Reaction Time Task: Implicit Motor Skill Learning? J. Neurosci. *27*, 10073-10075.

S2. Nissen, M.J., and Bullemer, P. (1987). Attentional requirements of learning: evidence from performance measures. Cognitive Psychology *19*, 1-32.

S3. Brown, R.M., and Robertson, E.M. (2007). Off-Line Processing: Reciprocal Interactions between Declarative and Procedural Memories J. Neurosci. *27*, 10468-10475.

S4. Cohen, D.A., and Robertson, E.M. (2011). Preventing interference between different memory tasks. Nature neuroscience *14*, 953-955.

S5. Robertson, E.M. (2012). New insights in human memory interference and consolidation. Current biology : CB *22*, R66-71.

S6. Curran, T. (1997). Higher order associative learning in amnesia: Evidence from the serial reaction time task. Journal of Cognitive Neuroscience *9*, 522-533.

S7. Schendan, H.E., Searl, M.M., Melrose, R.J., and Stern, C.E. (2003). An FMRI study of the role of the medial temporal lobe in implicit and explicit sequence learning. Neuron *37*, 1013-1025.

S8. Song, S., and Cohen, L.G. (2014). Practice and sleep form different aspects of skill. Nature communications *5*, 3407.

S9. Oldfield, R. (1971). The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia *9*, 97-113.

S10. Willingham, D.B., Nissen, M.J., and Bullemer, P. (1989). On the development of procedural knowledge. J Exp Psychol Learn Mem Cogn *15*, 1047-1060.

S11. Boyd, L.A., and Winstein, C.J. (2001). Implicit motor-sequence learning in humans following unilateral stroke: the impact of practice and explicit knowledge. Neurosci Lett *298*, 65-69.

S12. Robertson, E.M., Pascual-Leone, A., and Miall, R.C. (2004). Current concepts in procedural consolidation. Nat. Rev. Neurosci. *5*, 576-582.

S13. Cohen, D.A., Pascual-Leone, A., Press, D.Z., and Robertson, E.M. (2005). Off-line learning of motor skill memory: A double dissociation of goal and movement. PNAS *102*, 18237-18241.