Retrieval of a well-established skill is resistant to distraction: evidence from an implicit probabilistic sequence learning task

Teodóra Vékony[1] *, Lilla Török[2] *, Felipe Pedraza[3], Kate Schipper[4], Claire Plèche[4], László Tóth[2], Karolina Janacsek[5,6,7], and Dezso Nemeth[4,5,6]

[1] Department of Neurology, University of Szeged, Szeged, Hungary

[2] Department of Psychology and Sport Psychology, University of Physical Education, Budapest, Hungary

[3] Université Lumière - Lyon 2, Lyon, France

[4] Lyon Neuroscience Research Center (CRNL), INSERM, CNRS, Université Claude Bernard Lyon 1, Bron, France

[5] Institute of Psychology, ELTE Eötvös Loránd University, Budapest, Hungary

[6] Brain, Memory and Language Research Group, Institute of Cognitive Neuroscience and Psychology, Research Centre for Natural Sciences, Hungarian Academy of Sciences, Budapest, Hungary

[7] School of Human Sciences, Faculty of Education, Health and Human Sciences, University of Greenwich, London, United Kingdom

Corresponding author: Dezso Nemeth, Lyon Neuroscience Research Center, Université Claude Bernard Lyon 1, 95 Boulevard Pinel, 69675 Bron, France. Email: dezso.nemeth@univ-lyon1.fr, tel +33766807005

Conflict of interest: The authors report no conflict of interest.

***** These authors contributed equally to this work.

Abstract

The characteristics of acquiring new sequence information under dual-task situations have been extensively studied until now. Performing such a concurrent task has often been found to affect performance. In real life, however, we mostly perform a secondary task when the primary one is already well-acquired. The effect of a dual-task on the ability to retrieve well-established sequence representations remains elusive. The present study investigated whether accessing a well-acquired probabilistic sequence knowledge is affected by a concurrent task. Participants acquired non-adjacent regularities in an implicit probabilistic sequence learning task. After a 24-hour offline period, we tested the participants on the sequence learning task under dual-task or single-task conditions. Here we show that although the dual-task condition significantly prolonged the overall reaction times in the primary task, the access to the previously learned probabilistic representations remained intact. Moreover, we found an inverse relationship between the ability to successfully retrieve sequence knowledge and the accuracy of the secondary task, which fits in well with the hypothesis of competition between model-based and model-free processes. Our results highlight the importance of studying the dual-task effect not only in the learning phase but also during memory access.

Keywords: implicit sequence learning, statistical learning, memory retrieval, dual-task, divided attention

Introduction

Implicit sequence learning is a fundamental function of the brain, which underlies the acquisition of motor, cognitive and social skills. These skill-based actions, such as driving a car or doing sports, usually become more automatic with extensive practice. In everyday life, these skills are generally not performed in isolation but simultaneously with other actions. Therefore, the effect of a secondary task on implicit sequence learning performance has been studied extensively in the last few decades. Depending on the characteristics of the primary and the secondary task, evidence for impaired (e.g. Franklin, Smallwood, Zedelius, Broadway, & Schooler, 2016; Hemond, Brown, & Robertson, 2010; Jiménez & Vázquez, 2005; Nissen & Bullemer, 1987; Röttger, Haider, Zhao, & Gaschler, 2017; Shanks, Rowland, & Ranger, 2005; Wierzchon, Gaillard, Asanowicz, & Cleeremans, 2012), intact (e.g. Jiménez & Méndez, 1999; Jiménez & Vázquez, 2005; Nemeth et al., 2011; Röttger et al., 2017; Rowland & Shanks, 2006) or even improved performance (Hemond et al., 2010) was found during the acquisition of implicit sequence knowledge. However, the effect of a secondary task on the *retrieval* of a complex, well-established skill is rarely studied. In everyday life, we mostly perform a secondary task when the primary one is well-learned. For example, when we are learning how to drive, our entire attention is on the primary task, and we refrain from chatting with our co-pilot. However, after mastering this skill, we easily engage in conversations during the primary (driving) task. Therefore, answering the question of whether our performance is affected in such cases is crucial for the understanding of the effects of dual-tasks on real-life performances. Here we present a study where we test the effect of a secondary task *on the retrieval* of implicit sequence knowledge.

What do we know about the effect of a dual-task condition on the retrieval of the learned sequence? Sequence learning is typically measured by asking participants to respond to a series of stimuli that follow a predetermined sequence order (Nissen & Bullemer, 1987). By contrasting the performance to stimuli presented in a random order, the degree of sequence knowledge can be measured. Using short single-task practice periods and immediate retrieval, early studies have found impaired (Curran & Keele, 1993) or intact retrieval of sequence knowledge under dual-task conditions (Frensch, Lin, & Buchner, 1998; Shanks & Channon, 2002). The latter results were often interpreted as the secondary task affecting only the performance measured at the time of testing (i.e. the expression of knowledge), but not the underlying representations (Frensch et al., 1998). Cohen and Poldrack (2008) have investigated the effect of extended dual-task practice, and have demonstrated that the impairing effect of the dual-task decreases with extended practice on a serial reaction time task. These results raise the possibility that sequence learning knowledge remains intact after gaining experience on the primary task. These studies (along with most of the previous experiments testing the effect of dual-tasking on the initial learning) used fixed (deterministic) sequence learning tasks, the implicitness of which is questionable (Cohen & Poldrack, 2008; Frensch et al., 1998; Robertson & Cohen, 2006). For this reason, probabilistic sequence learning tasks are often used instead of deterministic ones. In a typical probabilistic sequence learning task, certain sequence elements occur at a higher frequency than others, and the participants learn to answer faster and more accurately for the more than the less probable ones (J. H. Howard & Howard, 1997). The probabilistic learning results in a robust knowledge (Kobor, Janacsek, Takacs, & Nemeth, 2017), and it is more likely to find intact or impaired sequence learning during a concurrent secondary task in the initial learning phase (Jiménez & Méndez, 1999, 2001); however, others did claim detrimental effects (Shanks et al., 2005). An early study of Schvaneveldt and Gomez (1998) have found that probabilistic sequence knowledge learned without a secondary task cannot be transferred to a dual-task condition. On the contrary, transfer from dual-task learning to single-task performance did occur, and the authors concluded that the impaired performance was due to problems in the expression of knowledge, and not to the impaired learning itself. However, what do we know about the effects of a secondary task after extended practice? There is a long history of studying the dual-task effects on automatic behaviours, mostly using non-complex, choice-response tasks. Most of the results support the claim that an already automatized behaviour is resistant to concurrently performing a secondary task (e.g. Logan, 1979) and that the dual-task cost on *general skill learning* (i.e. on the increasing speed due to practice independently from the sequence structure) tend to decrease with extended practice (e.g. Brown & Bennett, 2002; Hazeltine, Teague, & Ivry, 2002; Ruthruff, Johnston, & Van Selst, 2001; Ruthruff, Van Selst, Johnston, & Remington, 2006; Van Selst, Ruthruff, & Johnston, 1999). Nevertheless, to the best of our knowledge, no study has compared directly the accessibility of more complex, well-acquired probabilistic sequence knowledge (learned without a secondary task) between single and dual-task testing conditions.

In the present study, we aimed to investigate the effect of a concurrent secondary task on the retrieval of implicit probabilistic sequence knowledge. So far, studies have investigated the effect of single-task practice on immediate retrieval of the sequence knowledge. However, we do not know whether a newly-introduced secondary task might disrupt the retrieval of a consolidated sequence knowledge, although it resembles how we pursue dual-task situations in everyday life. Therefore, we investigated the effect of a secondary task on the retrieval of a well-learned probabilistic sequential knowledge after extended practice, a 24-hour offline period and a reactivation phase. First, we trained all participants on an implicit probabilistic sequence learning task in a single-task condition. After a 24-hour offline period, participants were tested again; however, at this stage, they performed the task with or without a stimulus-counting secondary task. To control for the expression of knowledge vs. acquired knowledge problem (Jiménez & Méndez, 1999; Vekony et al., 2019), we inserted probe blocks (i.e. blocks without a secondary task for both groups) into the second session. Based on previous findings, we hypothesized that the secondary task will not limit the accessibility of the well-learned probabilistic sequence knowledge.

Methods

Participants

The participants were selected from a pool of eighty-one participants. One participant who did not follow the dual-task instruction in the Retrieval Phase (see Procedure section in the Methods) was excluded from the pool. Two more participants were excluded because their age was more than 3 standard deviation than the average of the participants. From the rest of the pool, we selected the participants for the two groups to be matched by age, years of education, handedness, short term and working memory performance (Table 1), the learning performance in the first session, the level of consolidation (change between the end of the Learning Phase and the beginning of the Retrieval Phase), and by the performance in the beginning of Retrieval Phase (still without dual-task, see Procedure section in the Methods). By that, we were able to make sure that the observed effects were not due to differences in demographical characteristics, in general cognitive functions or in the level of initial learning or consolidation. The final analysis was carried out on 68 participants ($N_{\text{Single-task group}}$ = 34, $N_{\text{Dual-Task Group}}$ = 34). The participants were between 18 and 33 years of age (M_{age} = 22.91, SD_{age} = 3.48). The average years of education ranged between 10 and 20 ($M_{\text{years of education}} = 14.66$, $SD_{\text{years of education}} = 2.36$). Handedness was measured by the Edinburgh Handedness Inventory (Oldfield, 1971). The Laterality Quotient (LQ) of the sample varied between −53.85 and 100 (−100 means complete left-handedness, 100 means complete right-handedness; $M_{LQ} = 44.79$, $SD_{LQ} = 34.22$). All participants had normal or corrected to normal vision, none had a history of any neurological and/or psychiatric disorder or reported taking any psychoactive medication at the time of the experiment. They performed in the normal range on the Counting Span task ($RNG_{\text{Counting Span}}$: 2.33-6, $M_{\text{Counting Span}} = 3.81$, $SD_{\text{Counting Span}} = 0.89$) and on the Digit Span task ($RNG_{\text{Counting Span}}$: 5-9, $M_{\text{Counting Span}} = 6.28$, $SD_{\text{Counting Span}} = 1.13$). The Research Ethics Committee of the Eötvös Loránd University (Budapest, Hungary) approved the study and it was conducted in accordance with the Declaration of Helsinki.

Table 1. Comparison of the two groups on age, years of education, handedness, short-term, and working memory performance

	Dual-Task Group M(SD)	*Single-Task Group M(SD)*	*Group comparison (t-test results)*
Age (years)	23.03 (3.42)	22.79 (3.60)	$t(66) = -0.276$, $p = 0.783$
Education (years)	14.50 (2.27)	14.82 (2.47)	$t(66) = 0.562$, $p = 0.576$
Handedness (LQ)	50.86 (35.07)	38.73 (32.80)	$t(66) = -1.474$, $p = 0.145$
Counting Span Score	3.92 (0.96)	3.69 (0.81)	$t(66) = -1.045$, $p = 0.300$
Digit Span Score	6.32 (1.15)	6.24 (1.13)	$t(66) = -0.320$, $p = 0.750$

Alternating Serial Reaction Time task

We used the Alternating Serial Reaction Time (ASRT) task (J. H. Howard & Howard, 1997; Nemeth et al., 2010). Four empty circles (300 pixels each) were presented continuously in front of a white background arranged horizontally in the middle of a computer screen. A target stimulus (a black drawing of a dog's head, 300 pixels) was presented sequentially in one of the four empty circles. Participants were asked to respond with their middle and index fingers of both hands by pressing the button corresponding to the target position on a keyboard with four heightened keys (Z, C, B, and M on a QWERTY keyboard), each of the four keys corresponding to the circles in a horizontal arrangement. Participants were instructed to be as fast and as accurate as possible in the task (Figure 1A).

The serial order of the four possible locations (coded as 1, 2, 3, and 4) in which target stimuli could appear was determined by an eight-element probabilistic sequence. In this sequence, every first element appeared in the same order as the task progressed, while the second elements' positions were randomly chosen out of the four possible locations (e.g., 2r4r3r1r; where r indicates a random position). Thus, some combinations of three consecutive trials (*triplets*) occur with a greater probability than other combinations. For example, 2_**4**, 4_**3**, 3_**1**, and 1_**2** (where ''_" indicates any possible middle element of the triplet) would occur with high probability because the third element (bold numbers) could be derived from the sequence (or occasionally could be a random element as well). In contrast, 1_**3** or 4_**2** would occur with less probability because they could never be obtained consisting of two consecutive sequence elements. Therefore, the third element of a high-probability triplet is more predictable from the first element when compared to a low-probability triplet. There were 64 possible triplets in the task altogether (four stimuli combined for three consecutive trials). Sixteen of them were high-probability triplets, each of them occurring in approximately 4% of the trials, about five times more often than the low-probability triplets (0.8%). Overall, high-probability triplets occur with approximately 62.5% probability during the task, while the remaining 48 low-probability triplets only occur with 37.5% probability (Figure 1B). As the participants practice the ASRT task, their responses become faster and more accurate to the high-probability triplets compared to the low-probability ones, revealing statistical learning throughout the task (J. H. Howard & Howard, 1997; Kobor et al., 2017; Song, Howard, & Howard, 2007; Unoka et al., 2017).

The ASRT task was completed in blocks, and each block contained 85 button presses (5 random elements at the beginning of the block, then the 8-element alternating sequence was repeated 10 times). At the beginning of each block, four empty circles appeared horizontally on the screen for 200 ms, and after that, the first target stimulus turned up. The target stimulus remained on the screen until the first correct response. Between blocks, the participants could rest a bit, and along with this, feedback about their performance (average RTs and accuracy) appeared on the screen. After five blocks, a longer (5 minutes) mandatory pause was inserted. We defined each trial as the third element of a high or low-probability triplet. Trills (e.g. 1-2-1) and repetitions (e.g. 1-1-1) were eliminated from the analysis because participants may show pre-existing response tendencies for these types of triplets (D. V Howard et al., 2004; Janacsek, Borbély-Ipkovich, Nemeth, & Gonda, 2018; Takács et al., 2018; Unoka et al., 2017). The first button presses were also excluded from the analysis (first 5 random button presses, and the 6th and the 7th, as they cannot be evaluated as the third element of a triplet). To facilitate data analysis and to increase the signal-to-noise ratio, every five blocks were collapsed into a larger unit of analysis.

Procedure

The experiment consisted of two sessions (Figure 1C). In the Learning Phase, participants completed 45 blocks of ASRT (45 blocks divided into 9 units of analysis: Blocks 1-5, Blocks 6-10, Blocks 11-15, Blocks 16-20, Blocks 21-25, Blocks 26-30, Blocks 31-35, Blocks 36-40 and Blocks 41-45), which is long enough to acquire a stable statistical knowledge (e.g. Vekony et al., 2019). In the Retrieval Phase, which was held 24 hours after the Learning Phase, participants completed 23 blocks of ASRT with the same sequence that they learned the previous day (20 blocks divided into 4 units of analysis and 3 separate blocks: Blocks 1-5, Blocks 6-10, Blocks11, Blocks 12-16, Block 17, Blocks 18-22, Block 23). In Blocks 1-5, the instructions were completely the same as in the previous day. This phase was included to strengthen the knowledge acquired on the previous day and to make sure that the two groups consolidated the knowledge to a similar level (see Supplementary Material). However, in Blocks 6-10, Blocks 12-16 and Blocks 18-22, a random number of stimuli (between 40-45 out of the 85 appearing stimuli in one block) was colored in yellow. The Dual-Task Group was instructed to count the number of yellow dogs through the block. After the completion of the block, the participants had to type in the number of yellow dogs they counted. The yellow-colored stimuli appeared for the Single Task Group as well, however, they were instructed to carry on with the task without paying particular attention to the differently colored stimuli. The performance in the secondary task was evaluated by calculating the difference from the correct number of yellow stimuli divided by the total number of yellow stimuli for each unit of five blocks (thus, resulting in a percentage score of correctly counted yellow stimuli relative to the total number of yellow stimuli). Using this measure, we were able to control for the differences in the exact number of yellow stimuli across participants and for the fact that counting more stimuli is also considered as bad erroneous counting. If a participant's overall differences score was over 15%, the participant was considered not to properly follow the instructions and was excluded from the analysis (1 participant). Two probe blocks were inserted between the 3 dual-task phases (Block 11 and Block 17), and one more probe block at the end of the session (Block 23). In these blocks, there were no yellow stimuli for the Dual-Task, nor for the Single Task Group.

Following the ASRT task, the Inclusion-Exclusion task was administered to check whether the participants developed conscious knowledge about the learned probabilistic regularities (see Supplementary Materials for details).

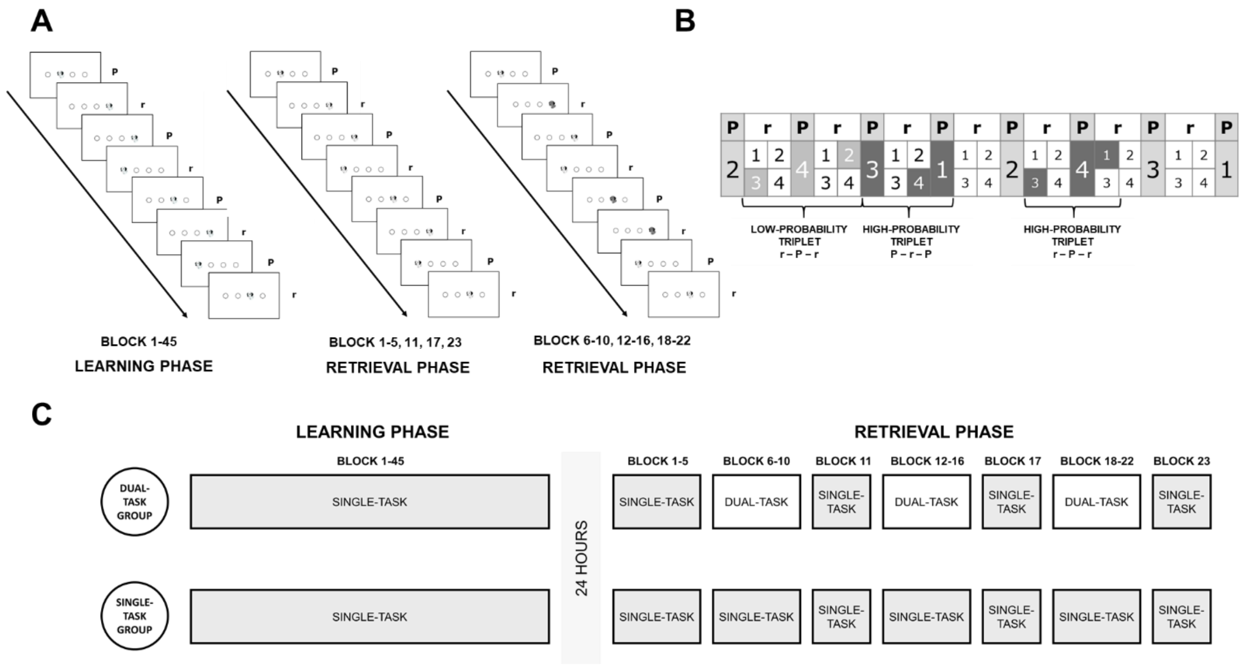

Figure 1. The structure of the ASRT and the experiment. (A) A dog's head appeared on one of the possible positions. The participants' task was to press the corresponding button as fast as possible. In the Learning Phase, only black-and-white stimuli appeared. During the Retrieval Phase, in Blocks 6-10, 12-16 and 18-22, sometimes yellow stimuli appeared on the screen, which the participants in the Dual-Task Group had to count, while the participants in the Single-Task Group were told to ignore. (B) High- and low-probability triplets. Some runs of consecutive trials (triplets) occurred with a greater probability than other elements. Every stimulus was defined as the third element of a high- or a low-probability triplet, based on the two preceding trials. High-probability triplets can be formed by two patterns and one random element and – although more rarely- by two random and one pattern elements. (C) Experimental design. In the Learning Phase, both groups of participants practiced single-task ASRT throughout 45 blocks. After 24 hours of rest, in the Retrieval Phase, participants completed 5 more blocks of single-task ASRT, and after that, 18 blocks of ASRT with or without dual-tasking (Dual Task Group and Single Task Group, respectively). However, 3 of them were probe blocks for both groups (Block 11, 17 and 23).

Statistical analysis

To evaluate the performance in the ASRT task, we calculated the median reaction times (RTs) and the mean accuracy scores separately for the high- and low-probability triplets in each 5 blocks in the Learning Phase (45 blocks: Blocks 1-5, Blocks 6-10, Blocks 11-15, Blocks 16-20, Blocks 21-25, Blocks 26-30, Blocks 31-35, Blocks 36-40 and Blocks 41-45), in the Retrieval Phase (20 blocks: Blocks 1-5, Blocks 6-10, Blocks 12-16, Blocks 18-22), and in each 3 probe blocks in the Retrieval Phase (3 blocks: Block 11, Block 17 and Block 23). Only correct responses were included in the RT analysis.

To test the effects of dual-tasking, we compared (1) the performance *while* the Dual-Task Group performed the dual-task (test) blocks, and (2) the performance *between* the dual-task phases (probe blocks). For the former, we performed a mixed-design ANOVA with Triplet (high- vs. low-probability) and Blocks (Blocks 6-10, Blocks 12-16 and Blocks 18-22 in the Retrieval Phase) as within-subject factors, and with Group (Dual-Task Group vs. Single-Task Group) as a between-subject factor. For the latter, we ran a mixed-design ANOVA with Triplet (high- vs. low-probability) and Probe block (Block 11 vs. Block 17 vs. Block 23 in the Retrieval Phase) as within-subject factors, and with Group (Dual-Task Group vs. Single-Task Group) as a between-subject factor. This analysis was carried out both on the RT and on the accuracy scores.

As the dual-tasking caused major differences in median RTs and average accuracy between the two groups in the Dual-Task Blocks, we wanted to make sure that the results found about the sequence knowledge were not due to the changes in the overall speed and accuracy (i.e., because of the dual-task effect on the *general skill learning*). To this end, we performed an additional analysis of the data with standardized scores. The corrected RT scores were calculated by dividing the learning scores (median RTs for low-probability triplets minus median RTs for high-probability triplets) by the *average* RTs of the given five blocks for each participant and for each unit of five blocks. We standardized the accuracy scores similarly, except that here the learning scores were defined by subtracting the accuracy for the low-probability triplets from the accuracy for the high-probability ones. The standardized scores were compared between groups with a mixed-design ANOVA with Block (Retrieval Phase 6-10 vs. Retrieval Phase 12-16 vs. Retrieval Phase 18-22) as a within-subject factor, and with Group (Dual Task Group vs. Single Task Group) as a between-subject factor. Please note that in this analysis, the dependent variable is a *difference score* of the high- and low-probability triplets (i.e., learning score).

For all ANOVAs, the Greenhouse-Geisser epsilon (ε) correction was used if necessary. Corrected *df* values and corrected *p* values are reported (if applicable) along with partial etasquared (η_p2) as the measure of effect size. LSD (Least Significant Difference) tests were used for pair-wise comparisons. Moreover, we conducted Bayesian ANOVAs on the *learning scores* to gain statistical evidence for null-results. Bayes Factor (BF) reflects how well a model behaves compared to the null-model. In the case of BF_{01}, the smaller the value is, the better the model predicts the data. The BF_{01} values of the null model are always 1 (containing the grand mean only) (Jarosz & Wiley, 2014). We report the BF_{01} values of the best model and its BF_M scores, which contrasts the predictive value of the model against all the other models (i.e. the degree to which the data have changed from the prior to posterior model odds) (Wagenmakers et al., 2018). All of the frequentist analysis of the ASRT task was carried out by using IBM SPSS Statistics 25, and the Bayesian analyses were run in JASP (version 0.10, JASP Team, 2019).

Results

Did the two groups perform equally before the dual-task phase?

As expected, both groups showed statistical learning in the Learning Phase, and the degree of learning did not differ between groups, as reflected by RTs (see 2nd paragraph of the Supplementary Materials) and by accuracies (see 3rd paragraph of the Supplementary Materials). The level of consolidation proved to be similar in the two groups (see 4th and 5th paragraph of the Supplementary Materials).

Did the two groups differ in the test blocks?

First, we wanted to find out whether the two groups performed differently in the test blocks (when the Dual-Task Group performed the secondary task). A mixed-design ANOVA on the RT scores with Triplet (high- vs. low-probability) x Block (Retrieval Phase Blocks 6-10 vs. Retrieval Phase Blocks 12-16 vs. Retrieval Phase Blocks 18-22) x Group (Dual-Task Group vs. Single-Task Group) factors revealed a significant main effect of Block ($F(1.665, 109.864) = 11.004$, $p < 0.001$, $n_p^2 = 0.143$), highlighting that the RTs on average accelerated during the Retrieval Phase. The significant main effect of Triplet ($F(1, 66) = 254.192$, $p < 0.001$, $n_p^2 = 0.794$) showed that the statistical knowledge was still detectable during the test blocks. Block × Triplet interaction was not significant ($F(2, 132) = 2.259$, $p = 0.108$, $n_p^2 = 0.033$) which means that the degree of the statistical learning did not change throughout the test blocks. The main effect of Group was significant ($F(1, 66) = 13.107$, $p < 0.001$, $n_p^2 = 0.166$), revealing that the average RTs were higher in the Dual-Task Group than in the Single-Task Group. This means that completing the secondary task during the ASRT caused a general slowing down in the task. The Block × Group interaction was significant ($F(2, 132) = 23.402$, $p < 0.001$, $n_p^2 = 0.262$), signaling that the general RT decrease was only detectable in the Dual-Task Group; the Single Task Group did not show acceleration following the Learning Phase. Importantly, the Triplet × Group interaction was not significant ($F(1, 66) = 0.521$, $p = 0.473$, $n_p^2 = 0.008$), which means that there was no difference between groups in terms of the degree of statistical knowledge. This lack of difference did not change throughout the blocks, as revealed by a non-significant Block × Triplet × Group interaction ($F(2, 132) = 0.211$, $p = 0.810$, $n_p^2 = 0.003$) (Figure 2). The Bayesian mixed-design ANOVA (containing only the Block and Group factors, as it was carried out on the *learning scores*, see the Statistical analysis section) favoured the null-model ($BF_{01} = 1$, $BF_M = 5.044$), suggesting similar level of statistical knowledge regardless of the block (Blocks 6-10 vs. Blocks 12-16 vs. Blocks 18-22) or the group (Single-Task Group or Dual-task Group).

The Block × Triplet × Group ANOVA of the accuracy scores also revealed a significant main effect of Block ($F(2, 132) = 6.541$, $p = 0.002$, $n_p^2 = 0.090$): accuracy scores on average became lower during the blocks of the Retrieval Phase. The significant main effect of Triplet ($F(2, 132) = 6.541$, $p = 0.002$, $n_p^2 = 0.090$) showed that statistical knowledge was observable also during this phase. However, the interaction of these two factors was also significant (Block × Triplet: $F(2, 132) = 4.399$, $p = 0.014$, $n_p^2 = 0.062$). The post hoc test revealed that although statistical knowledge was detectable in all units of five blocks, participants became less accurate for low-probability triplets during the Retrieval Phase. The main effect of Group was not significant ($F(1, 66) = 2.076$, $p = 0.154$, $n_p^2 = 0.031$), meaning that the average accuracy did not differed between groups. The Triplet × Group interaction was also not significant ($F(1, 66) = 1.744$, $p = 0.191$, $n_p^2 = 0.026$), which means that the degree of statistical knowledge was not affected by the dual-tasking. The Block × Triplet × Group interaction was also not significant ($F(2, 132) = 1.114$, $p = 0.331$, $n_p^2 = 0.017$), indicating that this lack of difference between group was stable across the test blocks (Figure 3). The Bayesian mixed-design ANOVA (Block × Group) revealed that the best fitting model contained the Block factor only ($BF_{01} = 0.422$, $BF_M = 3.202$). The other models had weaker predictive power, suggesting that the Group factor (Single-Task Group or Dual-task Group) did not ameliorate the model.

Did the two groups differ in the probe blocks?

We checked if the groups performed differently in the probe blocks inserted into the test blocks (Block 11, Block 17 and Block 23). The Block × Triplet × Group ANOVA of the RT scores revealed no difference of RTs between blocks ($F(2, 132) = 0.205$, $p = 0.815$, $n_p^2 = 0.003$), meaning that participants' RTs on average remained stable across the three probe blocks. The main effect of Triplet was found to be significant ($F(1, 66) = 114.514$, $p < 0.001$, $n_p^2 = 0.634$), revealing a stable statistical knowledge. The Block × Triplet interaction was also significant ($F(1, 66) = 4.491$, $p = 0.013$, $n_p^2 = 0.064$); however, as one probe block alone may be considered relatively noisy (as it contains only 85 button presses) compared to the average of three blocks, we did not consider these differences to be reliable. The main effect of Group was not significant ($F(1, 66) = 2.694$, $p = 0.105$, $n_p^2 = 0.039$), which means that the delaying effect of the dual-tasking did not affect the performance on the probe blocks. Most importantly, the Triplet × Group interaction was not significant ($F(1, 66) = 0.128$, $p = 0.721$, $n_p^2 = 0.002$), signaling a lack of difference between groups in terms of the degree of statistical knowledge also in the inserted probe blocks (Figure 4). The Block × Triplet × Group interaction was also not significant; however ($F(2, 132) = 1.935$, $p = 0.149$, $n_p^2 = 0.028$), again, this result is not considered reliable because of the lower signal-to-noise ratio compared to the results on the average of three blocks (Figure 2). The Bayesian mixed-design ANOVA revealed that the best fitting model contained the Block factor only ($BF_{01} = 0.411$, $BF_M = 4.769$). The other models had weaker predictive power, suggesting that the inclusion of the Group factor (Single-Task Group or Dual-task Group) did not ameliorate the model.

The Block × Triplet × Group ANOVA of the accuracy scores revealed no difference in terms of the average accuracy between blocks (Block: $F(2,132) = 2.027$, $p = 0.136$, $n_p^2 = 0.030$). Again, a stable statistical knowledge was proven by the Triplet main effect ($F(1, 66) = 75.822$, $p < 0.001$, $n_p^2 = 0.535$). The non-significant Block × Triplet interaction revealed no change of statistical knowledge during the blocks ($F(1.798, 118.696) = 1.403$, $p = 0.250$, $n_p^2 = 0.021$); however, again, this comparison may not be reliable because of the high signal-to-noise ratio. The groups did not differ in terms of the overall accuracy (Group: $F(1, 66) = 0.023$, $p = 0.880$, $n_p^2 < 0.001$). Most importantly, the degree of statistical knowledge was similar between groups in the probe blocks (Triplet × Group: $F(1, 66) = 0.929$, $p = 0.339$, $n_p^2 = 0.014$) (Figure 5). Again, this non-difference did not change across the blocks (Block × Triplet × Group: $F(2, 132) = 0.095$, $p = 0.910$, $n_p^2 = 0.001$) (Figure 3). The Bayesian mixed-designed ANOVA showed that the best fitting model was the null-model ($BF_{01} = 1$, $BF_M = 7.624$), suggesting that models the Block (Block 11 vs. Block 17 vs. Block 23) and/or the Group factor (Single-Task Group or Dual-task Group) does not predict the degree of statistical knowledge better.

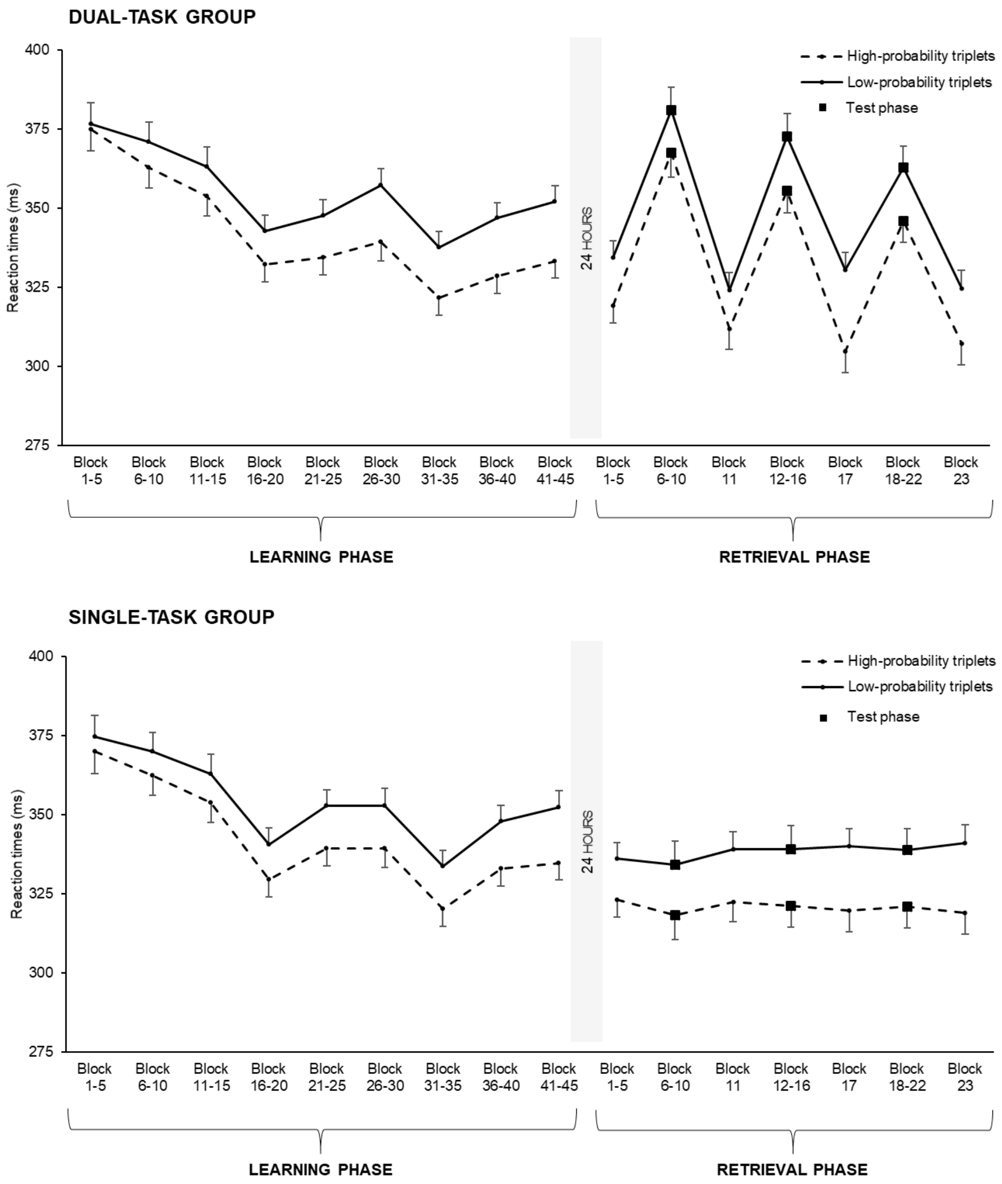

Figure 2. The RTs for the high- and low-probability triplets during the Learning and the Retrieval Phase, separately for the two groups. On the vertical axes, we can see the median RTs in milliseconds, and on the horizontal axis, the nine units of five blocks of the Learning Phase (ST 1 – ST 45) and the four units of 5 blocks and the 3 probe blocks of the Retrieval Phase (Blocks 1-5, Blocks 6-10, Block 11, Blocks 12-16, Block 17, Blocks 18-22, Block 23). The solid line represents the RTs for the low-probability triplets and the dashed line the RTs for the high-probability ones. The black squares represent the test phase when the Dual-Task Group performed the secondary task too. The error bars represent the standard error. The average RTs in the task became smaller over the Learning Phase, and the difference between the high- and low-probability triplets more pronounced. At the beginning of the Retrieval Phase, stable statistical knowledge was detected in Blocks 1-5. The dual-tasking slowed down the RTs of the participants, but the statistical knowledge remained stable even in these phases. We found similar results with standardized scores.

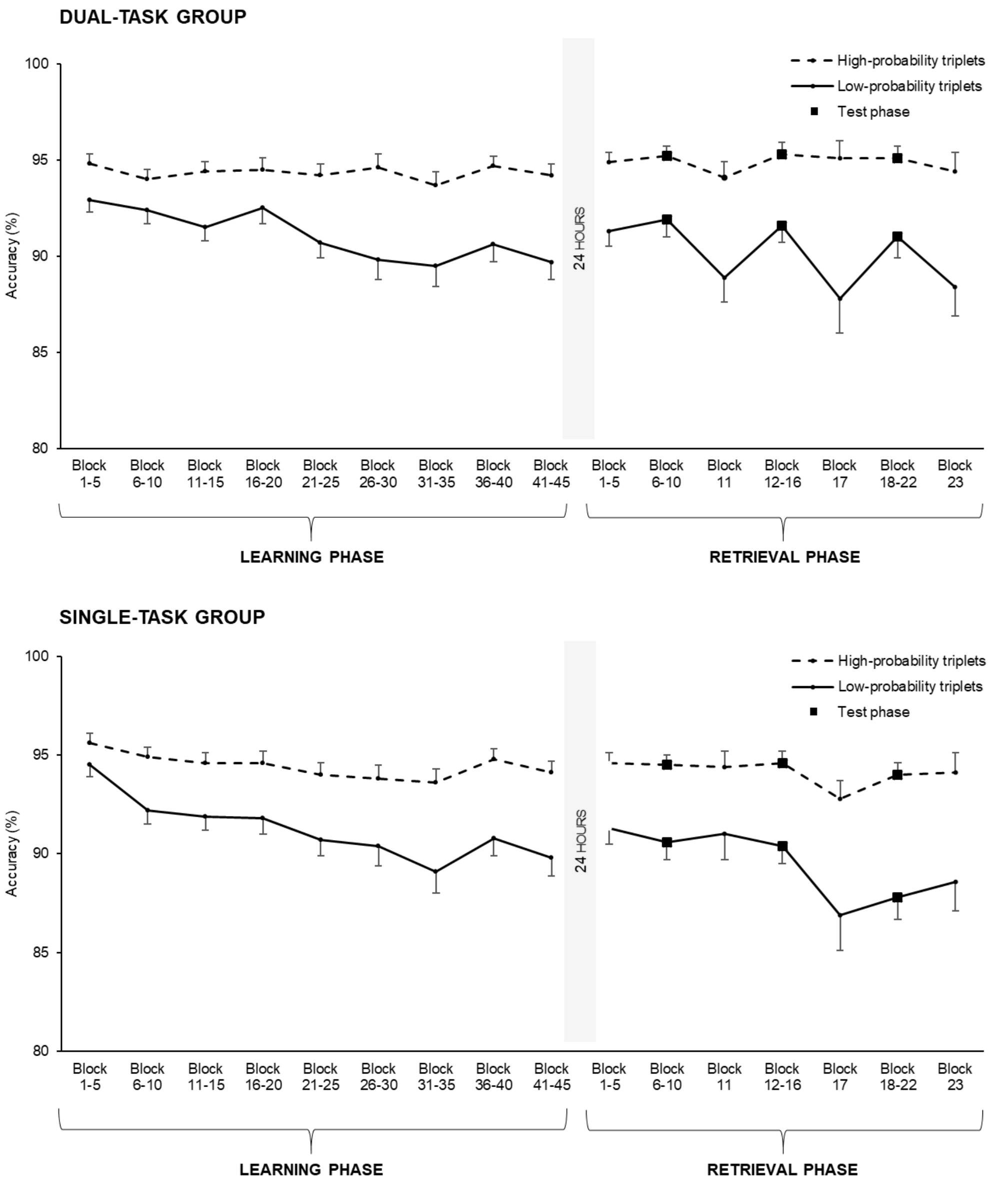

Figure 3. The accuracies for the high- and low-probability triplets during the Learning and the Retrieval Phase, separately for the two groups. On the vertical axes, we can see the mean accuracies in percentage, and in the horizontal axis, the nine units of five blocks of the Learning Phase (Blocks 1-45) and the four units of 5 blocks and the 3 probe blocks of the Retrieval Phase (Blocks 1-5, Blocks 6-10, Block 11, Blocks 12-16, Block 17, Blocks 18-22, Block 23). The solid line represents the RTs for the low-probability triplets and the dashed line the RTs for the high-probability ones. The black squares represent the test phase when the Dual-Task Group performed the secondary task. The error bars represent the standard error. The difference between the high- and low-probability triplets increased throughout the Learning Phase. At the beginning of the Retrieval Phase, stable statistical knowledge was detected, and it remained comparable even in the dual-task phases. We found similar results with standardized scores.

Did the two groups differ in the test blocks in terms of standardized scores?

In the Block × Group ANOVA on the *standardized RT learning scores* the main effect of Block was not significant ($F(2, 132) = 1.386$, $p = 0.254$, $n_p^2 = 0.021$), suggesting that the learning scores did not change during the test blocks. Importantly, consistent with the results without correction, no group difference was found in terms of the degree of statistical learning (Group: $F(1, 66) = 1.325$, $p = 0.254$, $n_p^2 = 0.020$). This lack of difference remained stable throughout the test blocks, as revealed by a non-significant Block × Group interaction ($F(2, 132) = 0.018$, $p = 0.982$, $n_p^2 < 0.001$). Similarly to the non-standardized results, the Bayesian mixed-design ANOVA suggested that the null-model is the best-fitting ($BF_{01} = 1$, $BF_M = 2.482$), revealing a similar level of statistical knowledge regardless of the block (Blocks 6-10 vs. Blocks 12-16 vs. Blocks 18-22) or the group (Single-Task Group or Dual-task Group).

The Block × Group ANOVA on the *standardized accuracy learning scores* revealed a significant main effect of Block ($F(2, 132) = 4.688$, $p = 0.011$, $n_p^2 = 0.066$): the learning score of Block 18-22 was overall higher than that of Blocks 6-10 ($p = 0.008$) and the Blocks 12-16 ($p = 0.027$). Similarly to the results of the non-standardized scores, no group differences were found in terms of the degree of the statistical knowledge (Group: $F(1, 66) = 1.678$, $p = 0.200$, $n_p^2 = 0.025$). This lack of group difference was consistent across the test blocks (Block × Group: $F(2, 132) = 1.133$, $p = 0.325$, $n_p^2 = 0.017$). The Bayesian mixed-design ANOVA showed that the best fitting model contained the Block factor only ($BF_{01} = 0.427$, $BF_M = 3.122$). The other models had weaker predictive power, suggesting that the Group factor (Single-Task Group or Dual-task Group) did not result in a better model about the degree of statistical knowledge.

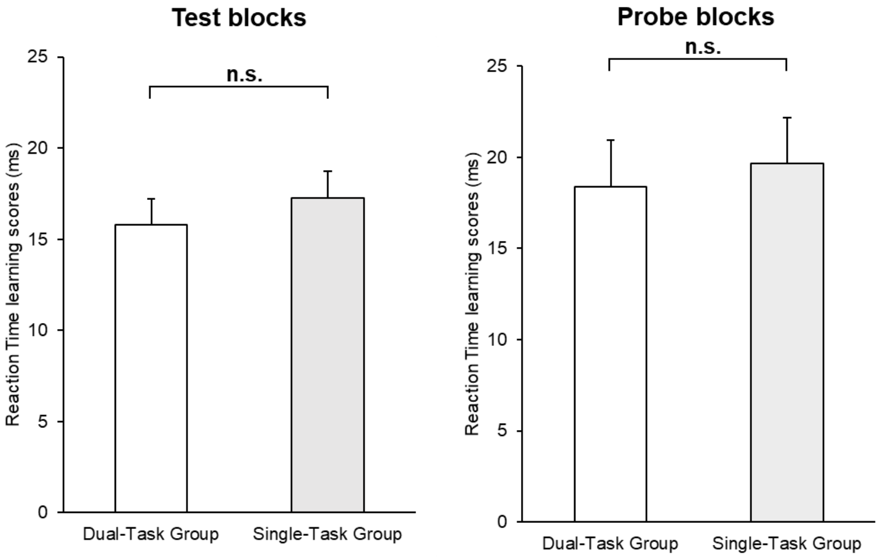

Figure 4. The RT learning scores of the target blocks and the probe blocks of the Retrieval Phase. The horizontal axis represents the two groups, and the vertical axis the RT learning scores (the RTs for the low-probability triplets *minus* the RTs for the high-probability triplets, the blocks collapsed together). The error bars signal the standard error. The learning scores of the two groups differed neither in the test blocks nor in the probe blocks. We found similar results with standardized scores.

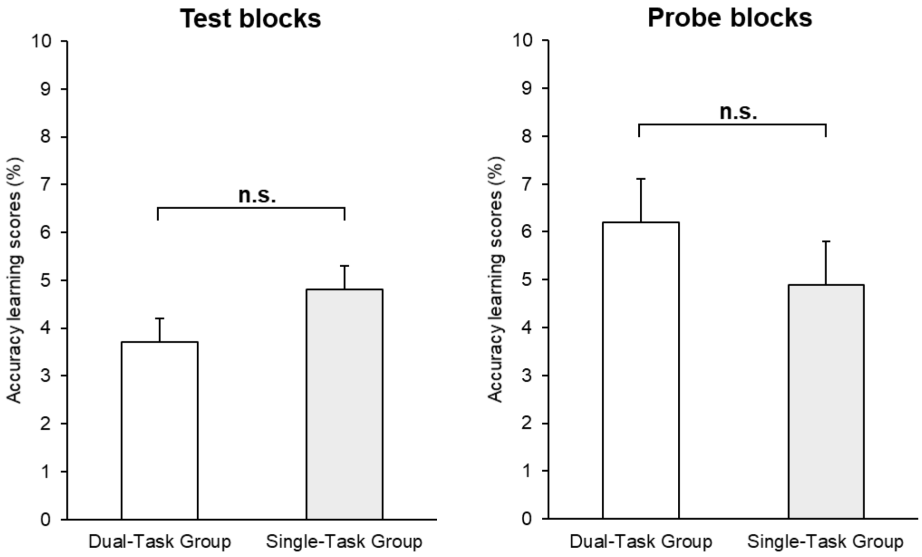

Figure 5. The accuracy learning scores of the test blocks and the probe blocks of the Retrieval Phase. The horizontal axis represents the two groups, and the vertical axis the accuracy learning scores (the accuracy for the high-probability triplets *minus* the accuracies for the low-probability triplets, the blocks collapsed together). The error bars signal the standard error. The learning scores of the two groups differed neither in the test blocks nor in the probe blocks. We found similar results with standardized scores.

Was the secondary task performance associated with the statistical learning performance in the Dual-Task Group?

The secondary task performance was evaluated by the difference between the number of appearing target stimuli and the reported number expressed as a percentage (see Methods for details). The mean difference was 4.98 % (SD = 4.2 %, RNG = 0.47-14.69 %). We explored for correlations between the secondary task performance of the Dual-Task Group with the ASRT learning scores in Blocks 6-10, Blocks 12-16 and Blocks 18-22 separately and also merged together (RTs and accuracy). We found positive correlation between the ASRT accuracy learning score in the last test blocks (Blocks 18-22) and the secondary task performance in Block 12-16 ($r(34) = 0.457$, $p = 0.007$), in the Blocks 18-22 ($r(34) = 0.486$, $p = 0.004$), and with the overall secondary task performance (Blocks 6-10, Blocks 12-16 and Blocks 18-22 merged together: $r(34) = 0.402$, $p = 0.018$). The correlation remained significant for the standardized blocks as well (secondary task Blocks 18-22 vs. ASRT ACC learning score in Blocks 12-16: $r(34) = 0.457$, $p = 0.007$; Block 18-22: $r(34) = 0.486$, $p = 0.004$; Overall: $r(34) = 0.402$, $p = 0.018$). It means that the worst secondary task performance (a bigger difference from the correct response in the last part of the dual-task phase) was associated with better sequence knowledge retrieval. No correlations were found with the RT learning scores (all $p > 0.545$) and correlation was not found with the accuracy learning scores in Blocks 6-10 and Blocks 12-16 either (all $p > 0.054$).

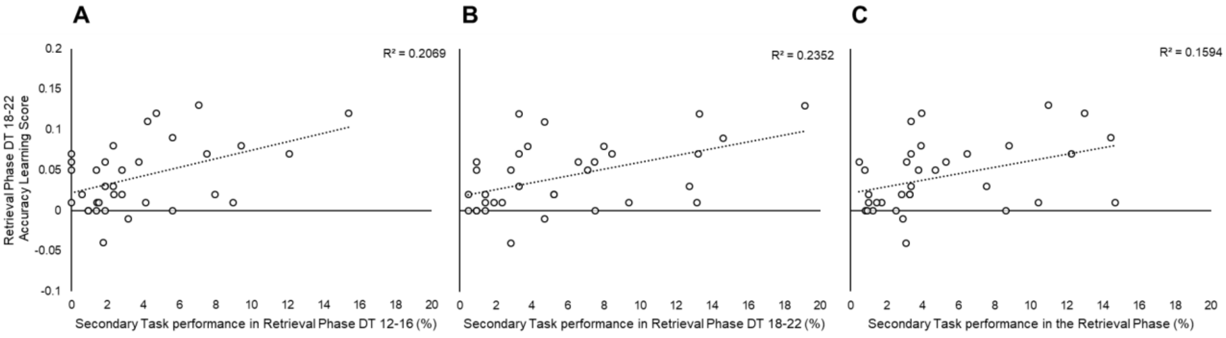

Figure 6. Correlation between the ASRT (primary task) accuracy learning scores of the last five dual-task blocks and the secondary task performance of Blocks 12-16 (A), Blocks 18-22 (B) and the overall secondary task performance (C). There was a positive correlation between the learning performance in Blocks 18-22 and the difference from the correct response in the secondary task. The thus higher level of learning was associated with worse performance in the secondary task.

Discussion

Here we investigated the effect of a secondary task on the retrieval of a well-established implicit sequence knowledge. Our participants practiced a sequence learning task with non-adjacent second-order dependencies through 45 blocks to master the ability to perform the sequence. After a 24-hour offline period, the participants were tested again on the task with or without a concurrent stimulus-counting task. Participants who were examined under dual-task conditions have retrieved their statistical knowledge to the same level as the participants with only single-task testing conditions. Moreover, this similarity proved to be true during blocks where both groups retrieved their knowledge under single-task conditions. These results remained the same even when we controlled for the differences in average RTs between groups. Moreover, the lack of difference between groups in terms of the implicit sequence learning performance was confirmed also by Bayesian statistical methods. Additionally, in accuracy measures, we found an inverse relationship between the performance on the primary implicit sequence learning task and on the secondary stimulus-counting task. It suggests that worst performance on the secondary task is associated with increased retrieval of probabilistic sequence knowledge. Our results went beyond previous literature by showing that well-learned, non-adjacent probabilistic sequence knowledge is resistant to a concurrent secondary task.

Our study could have had three possible outcomes: impaired, intact, or improved retrieval of the learned probabilities under dual-task conditions compared to single-task retrieval. Expecting impaired performance can seem reasonable at first, as the majority of previous literature reported deteriorating effects of a secondary task on sequence learning abilities (Franklin et al., 2016; Hemond et al., 2010; Jiménez & Vázquez, 2005; Röttger et al., 2017; Shanks et al., 2005; Wierzchon et al., 2012). This deteriorating effect on learning was explained by numerous theories such as the suppression hypothesis (Frensch et al., 1998; Frensch, Wenke, & Rünger, 1999), task integration (Schmidtke & Heuer, 1997), organizational hypothesis (Stadler, 1995) or the response selection hypothesis (Schumacher & Schwarb, 2009). However, other studies did reveal intact or even improved learning, especially in the case of more complex statistical regularities (Jiménez & Méndez, 1999; Jiménez & Vázquez, 2005; Nemeth et al., 2011; Schvaneveldt & Gomez, 1998). An important difference between our and most previous studies is that we introduced the secondary task after a considerable amount of practice on the primary task, and with this modification, we did not find evidence for impaired efficiency of the retrieval of the learned information under dual-task conditions. Contrary to our results, an early study of Schvaneveldt and Gomez (1998) have found impaired probabilistic sequence knowledge in a dual-task condition after initial single-task learning, but not on a single-task condition after initial dual-task learning. They have concluded that what is learned initially in single-task situations cannot be applied to a dual-task condition. However, in their study, they tested the transfer to dual-task conditions within one session (with less practice), while in our design, we implemented a longer practice period, a 24-hour consolidation and a reactivation period to make sure that the sequence is well-consolidated before the retrieval. This suggests that to acquire the sequence knowledge to a great extent might help maintain a good level of retrieval during a subsequent dual-task condition.

Apart from the potentially disruptive effect of the secondary task, another possible outcome of the study was that the concurrent task will leave access to the sequence knowledge intact. This would mean that the processes behind the sequence knowledge retrieval and the stimulus-counting secondary task are independent from each other, similarly to how performance becomes automatized and resistant to dual-tasking on non-complex choice-response tasks (e.g. Brown & Bennett, 2002; Hazeltine et al., 2002; Logan, 1979; Ruthruff et al., 2001, 2006; Van Selst et al., 1999). As knowledge becomes skill-like, it ceases to rely on the same resources. Our main results are in line with this prospect: the degree of statistical knowledge remained the same compared to the single task retrieval during dual-tasking, as well as in the probe blocks, where both groups accessed their statistical knowledge without a concurrent counting task (please note that here the sequence knowledge is the one that becomes automatized but not the perceptuo-motor improvement, see below). Moreover, the lack of differences persisted even after the normalization of the baseline scores. The fact that the sequence knowledge was comparable to the single-task group in both phases (in the dual-task blocks and in the inserted probe blocks) indicates that the secondary task affected neither the performance nor the competence of the primary task (Kiss, Nemeth, & Janacsek, 2019; Vekony et al., 2019). These results are also in harmony with previous research that found intact implicit sequence knowledge after practicing the primary task in single-task conditions (Frensch et al., 1998; Shanks & Channon, 2002). However, in these studies, the presentation of the dual-task blocks followed immediately the few learning blocks, which makes it harder to draw conclusions about the long-term retrieval. Therefore, our results extend this knowledge by providing evidence for three additional aspects. First, the retrieval of sequence knowledge remains resistant to a concurrent task even after a 24-hour offline period, which underlies the robust nature of probabilistic learning (Kobor et al., 2017; Nemeth & Janacsek, 2011). Second, the retrieval of clearly implicit probabilistic representations (see Supplementary Material) remain intact after extended practice. Third, neither the competence nor the performance of a well-acquired sequence knowledge can be disrupted by a secondary task.

The third potential outcome of the study was that the secondary task would improve the retrieval of the memory representations of the primary sequence learning task. This possibility would fit in well with the hypothesis of competition between the frontal-lobe mediated control functions and the model-free processes. Several studies have detected negative correlations between frontal functions and probabilistic statistical learning by means of behavioral and electrophysiological measurements in healthy subjects (Filoteo, Lauritzen, & Maddox, 2010; Tóth et al., 2017), by studying the developmental aspects of learning (Janacsek, Fiser, & Nemeth, 2012) or populations with hypofrontality (Virag et al., 2015) by using hypnosis to reduce frontal functions (Nemeth, Janacsek, Polner, & Kovacs, 2013), or by disrupting it with brain stimulation techniques (Ambrus et al., 2019). Here we did not find improved performance in the dual-task condition as would be predicted by the competition theory. However, we did find evidence for a negative relationship between the performance on the secondary task and the successful retrieval of the sequence knowledge. Namely, a negative relationship emerged between the accuracy of counting the target stimuli and the difference between the high- and low-probability triplets in the later parts of the dual-task blocks. In other words, long-lasting dual-tasking might make the frontal-lobe exhausted, and thus the retrieval of the implicit sequence knowledge becomes better. This explanation harmonizes with the argument that cognitive fatigue enhances procedural sequence learning (Borragan, Slama, Destrebecqz, & Peigneux, 2016). Most of the previous studies investigating the implicit sequence learning performance under dual-tasking did not report the relationship between the primary and secondary task performance (and used secondary task performance mostly as a criterion to exclude participants who did not follow the instructions). The above-mentioned correlational finding opens up many possible research questions regarding a potential negative relationship between frontal functions and the successful performance on sequence learning under dual-task conditions. Taken together, two out of three possible outcomes were supported by our data, namely the hypothesis of *independent processes* and the *competition*. Future studies revealing under which circumstances they are valid seems warranted.

Beyond the theoretical explanations of our results, methodological aspects can also account for them. The ASRT task allows us to disentangle the general skill-related processes and sequence-specific knowledge. It was not taken into account by many previous studies, which makes it harder to reveal the underlying mechanism behind dual-task effects. In our study, the general skill-related processes are shown by the change of the overall RTs and accuracy in time (perceptuo-motor adaptation), while the sequence-specific knowledge is considered to be the emergence of difference between high and low-probability triplets (statistical learning). It is important to note that the increased overall RTs during the retrieval phase under dual-task conditions did not reveal impaired *sequence knowledge:* they indicate altered *general skill retrieval* on the primary task due to the dual-task constraint. The dual-task slowed down the overall perceptuo-motor adaptation performance, which means that in this aspect, the performance was not automatized until this point. However, the statistical knowledge that emerged during the learning phase became robust enough to persist under dual-tasking, thus we found a dissociation between the two processes. The dissection of the perceptuo-motor adaptation and the statistical aspect of the learning process was also supported by the fact that after the normalization of the baseline RTs, the lack of differences in statistical learning between the single- and the dual-task group remained. As the development of general skill and statistical learning might follow different trajectories (Juhasz, Nemeth, & Janacsek, 2019), this result is crucial for stating that the sequence knowledge is the same between groups. Previous inconsistencies in the literature might also originate from differences in the proportion of general skill- and statistical learning-related factors of the used task. Therefore, future studies investigating the process of sequence learning or the retrieval of the learned knowledge under dual-task conditions could benefit from considering these aspects as potential confounding factors.

Previous studies have tried to determine which characteristics of the secondary tasks are crucial for disrupting the learning process, such as the correlation between the primary and the secondary task events or the features of the required response (e.g. Röttger et al., 2017). In our study, we chose a visual secondary task implemented in the stimuli stream, which does not break the stimulus-response interval causing interference as tone counting tasks might do (Jiménez & Méndez, 1999). However, we do not know if different secondary tasks involving functionally distant cognitive processes affect the sequence retrieval to a similar extent. For example, sentence processing has been found to impair probabilistic sequence learning, while mathematical tasks, word processing have not revealed disruptive effect (Nemeth et al., 2011). This result can be explained by the fact that language processing also relies on non-adjacent dependencies, similarly to the probabilistic sequence learning task used in the study. On the contrary, sequence learning has also been shown to be boosted when the secondary task involved similar sequence-learning mechanisms as the primary adjacent serial reaction time task did (Hemond et al., 2010). Therefore, studies revealing which cognitive processes can interfere with the retrieval of sequence information and which do not, seem warranted.

In everyday life, we mostly perform a secondary task when the primary one is well-acquired. In spite of this fact, no study has investigated the effect of a secondary task in the case of well-acquired, complex statistical regularities. With the aim of filling this gap of knowledge, we performed a dual-task experiment after extensive practice, a 24-hour offline period and a reactivation phase. We found evidence for intact (independent processes theory) and partially for improved performance (competition theory) under dual-tasking. Two main conclusions can be drawn from our study. First, we found an intact retrieval of implicit probabilistic sequence knowledge, providing evidence that complex probabilistic representations are robust against dual-tasking even if the general perceptuo-motoric aspect of the primary task is affected. This might mean that complex statistical representations become resistant earlier than the perceptuo-motoric processes and that we are able to correctly utilize them if we are performing a new, non-statistical secondary task. Second, we also detected a negative relationship between more successful retrieval and secondary task performance. This result fits in well with the competition hypothesis of model-based and model-free processes, thus, more studies revealing the nature of the relationship between control functions and the ability to retrieve sequence-specific knowledge seem necessary.

Acknowledgments

This research was supported by the National Brain Research Program (project 2017-1.2.1-NKP-2017-00002); Hungarian Scientific Research Fund (NKFIH-OTKA K 128016, PI: D.N., NKFIH-OTKA PD 124148, PI: K.J.); Janos Bolyai Research Fellowship of the Hungarian Academy of Sciences (to K.J.); IDEXLYON Fellowship (to D.N). Thanks to Lison Fanuel for the comments and suggestions on the manuscript.

References

Ambrus, G. G., Vekony, T., Janacsek, K., Trimborn, A.-B. C., Kovacs, G., & Nemeth, D. (2019). When less is more: enhanced statistical learning of non-adjacent dependencies after disruption of bilateral DLPFC. *BioRxiv*. https://doi.org/10.1101/198515

Borragan, G., Slama, H., Destrebecqz, A., & Peigneux, P. (2016). Cognitive fatigue facilitates procedural sequence learning. *Frontiers in Human Neuroscience*, *10*(MAR2016). https://doi.org/10.3389/fnhum.2016.00086

Brown, S. W., & Bennett, E. D. (2002). The role of practice and automaticity in temporal and nontemporal dual-task performance. *Psychological Research*, *66*(1), 80–89. https://doi.org/10.1007/s004260100076

Cohen, J. R., & Poldrack, R. A. (2008). Automaticity in motor sequence learning does not impair response inhibition. *Psychonomic Bulletin and Review*, *15*(1), 108–115. https://doi.org/10.3758/PBR.15.1.108

Curran, T., & Keele, S. W. (1993). Attentional and Nonattentional Forms of Sequence Learning. In *Journal of Experimental Psychology: Learning, Memory, and Cognition* (Vol. 19).

Filoteo, J. V., Lauritzen, S., & Maddox, W. T. (2010). Removing the frontal lobes: The effects of engaging executive functions on perceptual category learning. *Psychological Science*, *21*(3), 415–423. https://doi.org/10.1177/0956797610362646

Franklin, M. S., Smallwood, J., Zedelius, C. M., Broadway, J. M., & Schooler, J. W. (2016). Unaware yet reliant on attention: Experience sampling reveals that mind-wandering impedes implicit learning. *Psychonomic Bulletin and Review*, *23*(1), 223–229. https://doi.org/10.3758/s13423-015-0885-5

Frensch, P. A., Lin, J., & Buchner, A. (1998). Learning versus behavioral expression of the learned: The effects of a secondary tone-counting task on implicit learning in the serial reaction task. *Psychological Research*, *61*(2), 83–98. https://doi.org/10.1007/s004260050015

Frensch, P. A., Wenke, D., & Rünger, D. (1999). A Secondary Tone-Counting Task Suppresses Expression of Knowledge in the Serial Reaction Task. *Journal of Experimental Psychology: Learning Memory and Cognition*, *25*(1), 260–274. https://doi.org/10.1037/0278-7393.25.1.260

Hazeltine, E., Teague, D., & Ivry, R. B. (2002). Simultaneous dual-task performance reveals parallel response selection after practice. *Journal of Experimental Psychology: Human Perception and Performance*, *28*(3), 527–545. https://doi.org/10.1037/0096-1523.28.3.527

Hemond, C., Brown, R. M., & Robertson, E. M. (2010). A distraction can impair or enhance motor performance. *Journal of Neuroscience*, *30*(2), 650–654. https://doi.org/10.1523/JNEUROSCI.4592-09.2010

Howard, J. H., & Howard, D. V. (1997). Age differences in implicit learning of higher order dependencies in serial patterns. *Psychology and Aging*, *12*(4), 634–656. https://doi.org/10.1037/0882-7974.12.4.634

Howard, D. V, Howard, J. H., Japikse, K., DiYanni, C., Thompson, A., & Somberg, R. (2004). Implicit sequence learning: effects of level of structure, adult age, and extended practice. *Psychology and Aging*, *19*(1), 79–92. https://doi.org/10.1037/0882-7974.19.1.79

Janacsek, K., Borbély-Ipkovich, E., Nemeth, D., & Gonda, X. (2018). How can the depressed mind extract and remember predictive relationships of the environment? Evidence from implicit probabilistic sequence learning. *Progress in Neuro-Psychopharmacology and Biological Psychiatry*, *81*, 17–24. https://doi.org/10.1016/j.pnpbp.2017.09.021

Janacsek, K., Fiser, J., & Nemeth, D. (2012). The best time to acquire new skills: Age-related differences in implicit sequence learning across the human lifespan. *Developmental Science*, *15*(4), 496–505. https://doi.org/10.1111/j.1467-7687.2012.01150.x

Jarosz, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting bayes factors. *Journal of Problem Solving*, *7*(1), 2–9. https://doi.org/10.7771/1932-6246.1167

JASP Team. (2019). *JASP (Version 0.10)*.

Jiménez, L., & Méndez, C. (1999). Which Attention Is Needed for Implicit Sequence Learning? *Journal of Experimental Psychology: Learning Memory and Cognition*, *25*(1), 236–259. https://doi.org/10.1037/0278-7393.25.1.236

Jiménez, L., & Méndez, C. (2001). Implicit sequence learning with competing explicit cues. *Quarterly Journal of Experimental Psychology Section A: Human Experimental Psychology*, *54*(2), 345–369. https://doi.org/10.1080/713755964

Jiménez, L., & Vázquez, G. A. (2005). Sequence learning under dual-task conditions: Alternatives to a resource-based account. *Psychological Research*, *69*(5–6), 352–368. https://doi.org/10.1007/s00426-004-0210-9

Juhasz, D., Nemeth, D., & Janacsek, K. (2019). Is there more room to improve? The lifespan trajectory of procedural learning and its relationship to the between- and within-group differences in average response times. *BioRxiv*, 593582. https://doi.org/10.1101/593582

Kiss, M., Nemeth, D., & Janacsek, K. (2019). Stimulus presentation rates affect performance but not the acquired knowledge – Evidence from procedural learning. *BioRxiv*, 650598. https://doi.org/10.1101/650598

Kobor, A., Janacsek, K., Takacs, A., & Nemeth, D. (2017). Statistical learning leads to persistent memory: Evidence for one-year consolidation. *Scientific Reports*, *7*(1). https://doi.org/10.1038/s41598-017-00807-3

Logan, G. D. (1979). On the use of a concurrent memory load to measure attention and automaticity. *Journal of Experimental Psychology: Human Perception and Performance*, *5*(2), 189–207. https://doi.org/10.1037/0096-1523.5.2.189

Nemeth, D., & Janacsek, K. (2011). The dynamics of implicit skill consolidation in young and elderly adults. *Journals of Gerontology - Series B Psychological Sciences and Social Sciences*, *66 B*(1), 15–22. https://doi.org/10.1093/geronb/gbq063

Nemeth, D., Janacsek, K., Csifcsak, G., Szvoboda, G., Howard, J. H., & Howard, D. V. (2011). Interference between sentence processing and probabilistic implicit sequence learning. *PLoS ONE*, *6*(3). https://doi.org/10.1371/journal.pone.0017577

Nemeth, D., Janacsek, K., Londe, Z., Ullman, M. T., Howard, D. V., & Howard, J. H. (2010). Sleep has no critical role in implicit motor sequence learning in young and old adults. *Experimental Brain Research*, *201*(2), 351–358. https://doi.org/10.1007/s00221-009-2024-x

Nemeth, D., Janacsek, K., Polner, B., & Kovacs, Z. A. (2013). Boosting human learning by hypnosis. *Cerebral Cortex*, *23*(4), 801–805. https://doi.org/10.1093/cercor/bhs068

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology*, *19*(1), 1–32. https://doi.org/10.1016/0010-0285(87)90002-8

Oldfield, R. C. (1971). The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia*, *9*(1), 97–113. https://doi.org/10.1016/0028-3932(71)90067-4

Robertson, E. M., & Cohen, D. A. (2006, June 29). Understanding consolidation through the architecture of memories. *Neuroscientist*, Vol. 12, pp. 261–271. https://doi.org/10.1177/1073858406287935

Röttger, E., Haider, H., Zhao, F., & Gaschler, R. (2017). Implicit sequence learning despite multitasking: the role of across-task predictability. *Psychological Research*, *83*(3), 526–543. https://doi.org/10.1007/s00426-017-0920-4

Rowland, L. A., & Shanks, D. R. (2006). Sequence learning and selection difficulty. *Journal of Experimental Psychology: Human Perception and Performance*, *32*(2), 287–299. https://doi.org/10.1037/0096-1523.32.2.287

Ruthruff, E., Johnston, J. C., & Van Selst, M. (2001). Why practice reduces dual-task interference. *Journal of Experimental Psychology: Human Perception and Performance*, *27*(1), 3–21. https://doi.org/10.1037/0096-1523.27.1.3

Ruthruff, E., Van Selst, M., Johnston, J. C., & Remington, R. (2006). How does practice reduce dual-task interference: Integration, automatization, or just stage-shortening? *Psychological Research*, *70*(2), 125–142. https://doi.org/10.1007/s00426-004-0192-7

Schmidtke, V., & Heuer, H. (1997). Task integration as a factor in secondary-task effects on sequence learning. *Psychological Research*, *60*(1–2), 53–71. https://doi.org/10.1007/BF00419680

Schumacher, E. H., & Schwarb, H. (2009). Parallel Response Selection Disrupts Sequence Learning Under Dual-Task Conditions. *Journal of Experimental Psychology: General*, *138*(2), 270–290. https://doi.org/10.1037/a0015378

Schvaneveldt, R. W., & Gomez, R. L. (1998). Attention and probabilistic sequence learning. *Psychological Research*, *61*(3), 175–190. https://doi.org/10.1007/s004260050023

Shanks, D. R., & Channon, S. (2002). Effects of a secondary task on "implicit" sequence learning: Learning or performance? *Psychological Research*, *66*(2), 99–109. https://doi.org/10.1007/s00426-001-0081-2

Shanks, D. R., Rowland, L. A., & Ranger, M. S. (2005). Attentional load and implicit sequence learning. *Psychological Research*, *69*(5–6), 369–382. https://doi.org/10.1007/s00426-004-0211-8

Song, S., Howard, J. H., & Howard, D. V. (2007). Sleep does not benefit probabilistic motor sequence learning. *Journal of Neuroscience*, *27*(46), 12475–12483. https://doi.org/10.1523/JNEUROSCI.2062-07.2007

Stadler, M. A. (1995). Role of Attention in Implicit Learning. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *21*(3), 674–685. https://doi.org/10.1037/0278-7393.21.3.674

Takács, Á., Kóbor, A., Chezan, J., Éltető, N., Tárnok, Z., Nemeth, D., … Janacsek, K. (2018). Is procedural memory enhanced in Tourette syndrome? Evidence from a sequence learning task. *Cortex*, *100*, 84–94. https://doi.org/10.1016/j.cortex.2017.08.037

Tóth, B., Janacsek, K., Takács, Á., Kóbor, A., Zavecz, Z., & Nemeth, D. (2017). Dynamics of EEG functional connectivity during statistical learning. *Neurobiology of Learning and Memory*, *144*, 216–229. https://doi.org/10.1016/j.nlm.2017.07.015

Unoka, Z., Vizin, G., Bjelik, A., Radics, D., Nemeth, D., & Janacsek, K. (2017). Intact implicit statistical learning in borderline personality disorder. *Psychiatry Research*, *255*, 373–381.

https://doi.org/10.1016/j.psychres.2017.06.072

Van Selst, M., Ruthruff, E., & Johnston, J. C. (1999). Can practice eliminate the psychological refractory period effect? *Journal of Experimental Psychology: Human Perception and Performance*, *25*(5), 1268–1283. https://doi.org/10.1037/0096-1523.25.5.1268

Vekony, T., Marossy, H., Must, A., Vecsei, L., Janacsek, K., & Nemeth, D. (2019). Disentangling competence from performance in behavioral measures of learning: A lesson for cognitive neuroscience. *BioRxiv*, 726315. https://doi.org/10.1101/726315

Virag, M., Janacsek, K., Horvath, A., Bujdoso, Z., Fabo, D., & Nemeth, D. (2015). Competition between frontal lobe functions and implicit sequence learning: evidence from the long-term effects of alcohol. *Experimental Brain Research*, *233*(7), 2081–2089. https://doi.org/10.1007/s00221-015-4279-8

Wagenmakers, E. J., Love, J., Marsman, M., Jamil, T., Ly, A., Verhagen, J., … Morey, R. D. (2018). Bayesian inference for psychology. Part II: Example applications with JASP. *Psychonomic Bulletin and Review*, *25*(1), 58–76. https://doi.org/10.3758/s13423-017-1323-7

Wierzchon, M., Gaillard, V., Asanowicz, D., & Cleeremans, A. (2012). Manipulating attentional load in sequence learning through random number generation. *Advances in Cognitive Psychology*, *8*(2), 179–195. https://doi.org/10.2478/v10053-008-0114-0