

Working memory affects anticipatory behavior during implicit pattern learning

Srdan Medimorec, Petar Milin and Dagmar Divjak

University of Birmingham

Corresponding author information:

Srdan Medimorec, Department of Modern Languages, University of Birmingham, Edgbaston,

Birmingham, B15 2TT, UK. Email: s.medimorec@bham.ac.uk

Abstract

We investigated the relation between implicit sequence learning and individual differences in working memory (WM) capacity. Participants performed an oculomotor version of the serial reaction time (SRT) task and three computerized WM tasks. Implicit learning was measured using anticipation measures only, as they represent strong indicators of learning. Our results demonstrate that anticipatory behavior in the SRT task changes as a function of WM capacity, such that it increases with decreased WM capacity. On the other hand, WM capacity did not affect the overall number of correct anticipations in the task. In addition, we report a positive relation between WM capacity and the number of consecutive correct anticipations (or chunks), and a negative relation between WM capacity and the overall number of errors, indicating different learning strategies during implicit sequence learning. The results of the current study are theoretically important because they demonstrate that individual differences in WM capacity could account for differences in learning processes, and ultimately change individuals' anticipatory behavior, even when learning is implicit, without intention and awareness.

Keywords: sequence learning; implicit learning; anticipation; the SRT task; working memory

Working memory affects anticipatory behavior during implicit pattern learning

Anticipation represents a ubiquitous and central characteristic of sequence learning (Cleeremans & McClelland, 1991; Dale, Duran, & Morehead, 2012). However, the exact nature of the relation between anticipation during sequence learning and memory remains poorly understood (Dale et al., 2012; Janacsek & Nemeth, 2013; Schwarb, & Schumacher, 2012). For example, working memory (WM) could be involved in sequence learning because WM is needed for retrieval of information from long term memory (Martini, Furtner, & Sachse, 2013). On the other hand, it is possible that WM is not related to sequence learning, because of the relative automaticity of this type of learning (e.g., Unsworth & Engle, 2005). Understanding how these systems interact appears to be critical in elucidating the mechanisms underlying sequential learning. To this end, we investigated whether individuals' anticipatory behavior changes as a function of working memory (WM) capacity during implicit sequence learning.

Sequence learning is one of the fundamental cognitive abilities that enables individuals to acquire representations of their environment. Specifically, sequence learning is a mechanism responsible for the acquisition of abstract knowledge of regularities present in the environment. Previous research has reported evidence that such distributional learning can take place implicitly - without awareness and explicit instructions (Stadler & Frensch 1998; Turk-Browne, Scholl, Chun, & Johnson, 2009). Implicit sequence learning underpins many aspects of human behavior, such as language and various motor skills (Cleeremans & McClelland, 1991; Conway, Bauernschmidt, Huang, & Pisoni, 2010; Hunt & Aslin, 2001; Masters, 1992). For example, Conway et al. (2010) demonstrated a correlation between performance on an implicit learning task and sensitivity to word predictability in speech. The authors suggested that increased

implicit learning capabilities result in more detailed representations of the word order probabilities, leading to improvements in speech perception.

The relation between implicit sequence learning (typically measured by reaction time (RT)), and WM capacity has been studied extensively (e.g., Bo, Jennett, & Seidler, 2012; Caljouw, Veldkamp, & Lamoth, 2016; Feldman, Kerr, & Streissguth, 1995; Guzmán, 2018; Kaufman et al, 2010; Unsworth & Engle, 2005; Weitz, O'Shea, Zook, & Needham, 2011; Yang & Li, 2012). However, previous studies reported mixed results regarding the relation between the two systems, potentially because researchers used different WM capacity tests (i.e., visuospatial, verbal, or numerical), but also different measures of learning in the SRT task (e.g., the difference in average RT between blocks with a training sequence and blocks with a random sequence, or the rate of RT improvement across blocks of the SRT task; for a review, see Janacsek & Nemeth, 2013).

For example, a number of previous studies found no relation between implicit sequence learning and WM capacity (Caljouw et al., 2016; Guzmán, 2018; Jimenez & Vazquez, 2005; Jongbloed-Pereboom, Nijhuis-van der Sanden, & Steenbergen, 2019; Masters, 1992; Meissner, Keitel, Südmeyer, & Pollok, 2016; Unsworth & Engle, 2005; Yang & Li, 2012). Unsworth and Engle (2005) reported that there were no differences in implicit learning in a manual version of the SRT between high and low WM capacity individuals. The authors measured implicit learning as a difference in average RT between sequence blocks and random trials, while WM capacity was measured using operation span (a numerical task). On the other hand, the authors reported WM capacity effects on explicit sequence learning, when participants were aware of the sequence and the learning goal (see also Bo, Borza, & Seidler, 2009; Weitz et al., 2011). Together, these findings were interpreted to indicate the importance of WM capacity in tasks that

require some form of control (i.e., explicit learning), compared to relative automaticity of implicit learning (see also Kaufman et al., 2010).

In contrast, there is evidence from different domains that support the likelihood that implicit learning and WM capacity could be related. For example, Bo et al. (2012) reported a significant positive correlation between visuospatial WM capacity and the rate of RT improvement in the SRT task. Visuospatial WM was measured using a change detection task, while the implicit learning was measured as a change in median RT across sequential blocks. The authors argued that learning in the SRT task relies on the number of items individuals can hold in WM.

Interestingly, there is also evidence that WM and implicit learning could be negatively correlated. Based on findings obtained with artificial neural network models of language learning, Elman (1993) reported that learning can be actually improved under conditions of limited WM capacity. Likewise, Newport (1990) hypothesized that developmentally immature language learners focus on simpler linguistic structures because of their limited WM capacity (see also Erickson & Thiessen, 2015). In this way, early learners or learners with limited WM capacity grasp only constituents of speech, and then combine them into more complex structures. Late learners or learners with larger WM capacity, conversely, can take in more complex structures, which appears to put them at a disadvantage for learning (we will return to this in our General Discussion). A series of recent behavioral and brain imaging studies has demonstrated that weaker executive functions can lead to better implicit learning (Nemeth, Janacsek, Polner, & Kovacs, 2013; Tóth et al., 2017; Virag et al., 2015). For example, Nemeth et al. (2013) reported increased sequence learning in hypnosis, compared to a waking alert condition. The authors attributed such improved learning to the disruption to the executive system caused by hypnotic

instructions. In addition, using EEG data, Tóth et al. (2017) found that a better statistical learning score was related to a lower strength of connectivity between the sensorimotor and cognitive control brain regions.

Moreover, recent studies have set out to demonstrate that WM capacity and implicit learning, at least partly, relied on shared brain networks thus implying the relation between the two systems (e.g., Hasson, Cashdollar, Weisz, & Ruhnau, 2016; Janacsek & Nemeth, 2013). Using magnetoencephalography recordings, Hasson et al. (2016) found increased neural activity in higher WM capacity individuals when visual stimuli occurred with greater statistical regularity. Moreover, the SRT task has been demonstrated to recruit a brain network that includes the dorsolateral prefrontal cortex (DLPFC), which also plays a role in conscious executive processes (Torriero et al., 2007).

Present Investigation

In the current study, we investigate the relation between WM capacity and anticipatory measures during implicit sequence learning. The WM capacity measure reported here represents a combined score of three different complex verbal and visuospatial span tasks. This more encompassing WM capacity measure should account for the potential differential correlations of different WM capacity measures with implicit sequence learning (cf., Janacsek & Nemeth, 2013). In addition, implicit learning was measured using anticipation measures only, as they represent strong indicators of learning (Dale, Duran, & Morehead, 2012; Nissen & Bullemer, 1987; Schvaneveldt & Gomez, 1998; Stadler, 1989). We used the oculomotor variation of the serial reaction time task (SRT task; Kinder, Rolfs, & Kliegl, 2008; Marcus, Karatekin, & Markiewicz, 2006; Vakil, Bloch, & Cohen, 2017) previously demonstrated to provide measures of anticipatory responses, a feature that is hard to examine in manual versions of the SRT task.

This is because manual responses in the SRT task are typically made after the stimulus onset, with only a limited number of pre-target button presses that would indicate anticipations (Marcus et al., 2006). In contrast, anticipatory eye movements appear at a relatively high rate in this task, and can be measured directly by recording eye movements (i.e., anticipations appear if participants transition gaze towards a next position before the subsequent stimulus appears; Marcus et al., 2006; Vakil et al., 2017). We investigate the relation between direct measures of anticipation, namely the overall number of anticipations and the number of correct anticipations, and individual differences in WM capacity. The anticipation measures reported here could be conceptualized as indicators of *learning strategies* (i.e., the overall number of anticipations), and *learning outcomes* (i.e., the number of correct anticipations) in the SRT task. Such conceptualization allows us to investigate how these processes unfold during sequence learning, and how they relate to individual differences in WM capacity. In addition, we provide a more nuanced analysis of learning efficiency, and by extension of learning strategies, by investigating the relation between WM capacity and individual differences in the number of consecutive correct anticipations (i.e., chunking patterns). Previous research investigating explicit sequence learning has reported a positive correlation between WM capacity and chunking, defined by faster RTs of groups of movements (Bo & Seidler, 2009; Kennerley, Sakai, & Rushworth, 2004; Sakai, Kitaguchi, & Hikosaka, 2003; Shea, Park, & Braden, 2006). This positive correlation is usually taken to indicate a WM imposed limit on sequence elements that can be considered during explicit learning (on average around three items in length; e.g., Bo & Seidler, 2009). Thus, lower and higher WM spans potentially use different sequence learning strategies. To our knowledge the relation between WM capacity and chunking, here directly measured by the

length of correct anticipation sequences, has not been reported in the context of implicit sequence learning.

During a typical SRT task, participants are asked to follow a target on the screen. Critically, unknown to the participants, the target presentation follows a fixed (i.e., to-be-learned) sequence, and at some point, a different (i.e., interfering) sequence is intercalated. We expect that the number of correct anticipations (henceforth *correct anticipations*) should increase across the learning blocks (i.e., during a presentation of a to-be-learned sequence), decrease in the interference block (i.e., during a presentation of a different sequence), and then again increase in the recovery block (i.e., another presentation of the learning sequence; Marcus et al., 2006; Vakil et al., 2017). In addition, we report the changes in overall number of anticipations (both correct and incorrect; henceforth *anticipations*).

Based on a number of previous studies (e.g., Caljouw et al., 2016; Guzmán, 2018; Kaufman et al., 2010; Unsworth & Engle, 2005), a relation between WM capacity and implicit sequence learning would not be expected. However, if WM capacity and implicit learning at least partly rely on shared mechanisms, we should expect that the two systems are related to some extent (Janacsek & Nemeth, 2013). As noted in the introduction, the potential direction of this relation remains a matter of debate. For example, participants with higher WM capacity could have additional resources available during the learning phase, which could lead to more anticipations and more correct anticipations compared to participants with lower spans. By this account, WM capacity imposes an upper bound on the number of items that can be considered during the SRT task. A positive relation between WM capacity and the number of consecutive correct anticipations (i.e., chunks) would further corroborate this account.

Method

Participants

Participants were 35 students and staff (28 female; mean age = 27.4; age range = 18 - 56) from the University of Sheffield. Participants received a £7 Amazon voucher. All participants had normal or corrected-to-normal vision. Two additional participants were excluded because they failed to complete the experiment. The study was a part of a larger research project on language learning.

Materials

The serial reaction time task. We introduced a rapid oculomotor version of a deterministic SRT task, in an attempt to attenuate potential explicit awareness of the task.¹ The oculomotor SRT task, based on the digital SRT task by Nissen and Bullemer (1987; see also Kinder et al., 2008; Marcus et al., 2006; Vakil, et al., 2017), was implemented using OpenSesame (Mathôt, Schreij, & Theeuwes, 2012). Eye movements were recorded using an EyeLink Portable Duo eye tracker (SR research, ON, Canada), tracking at a sampling rate of 500 Hz in the head-stabilized mode. Participants were calibrated by the 9-point calibration type. Tracking was monocular, using participant's dominant eye, while viewing was binocular. Overall, the right eye was recorded for 69 % of participants ($n = 24$). Stimuli consisted of five slides, with resolution of 1024 x 768 pixels. Each slide contained four white 65 x 65 mm squares on a grey background (the white squares were also our areas of interest (AOIs); see Figure 1). The target (a black circle) with a diameter of 20 mm appeared in different white squares across four slides, while the fifth slide was the "anticipatory" slide and contained blank squares only.

¹ More complex probabilistic sequences could likewise increase task reliability (Stark-Inbar, Raza, Taylor, & Ivry, 2016; West, Vellido, Shanks, & Hulme, 2017).

The slides were presented centrally on a 21 in. monitor (refresh rate: 60 Hz), 70 cm away from participants' eyes. Individual white squares subtended visual angles of 5.5° horizontally and vertically, while the target subtended visual angles of 1.7° . We used second-order conditional sequences (SOC; Gabriel et al., 2013; Vakil, et al., 2016; Wilkinson & Shanks, 2004), meaning a target location could be predicted only if the two preceding locations were considered. We used two sequences: "342312143241" and "341243142132" (adopted from Wilkinson & Shanks, 2004). Here, the numbers 1–4 correspond to the four positions: down, left, right, and up. Each sequence served either as the learning or the interfering sequence, and the order of sequences was counterbalanced across the participants.



Figure 1. Anticipatory slide (left panel), and one of the target slides (right panel).

Explicit knowledge questionnaire. To assess sequence awareness, participants were asked the following questions after they had completed the task: (1) Did you notice anything special about the experiment?; (2) Did you notice any patterns during the experiment?; (3) If so, could you explicitly recall the pattern?; (4) If so - please write the pattern down.

Working memory tasks. WM capacity was measured using automated versions of three complex span tasks: the operation span, reading span, and symmetry span. The tasks were

administered using Tatool (von Bastian, Locher, & Rufin, 2013), a Java based programming framework.

In the operation span task, participants were shown a random number that needed to be remembered. Each number was followed by a math problem (e.g., $3 \times 7 = 21$) and participants were asked to make a decision on the veracity of the provided answer (half of the answers were correct). At recall, participants were asked to type in the random numbers they had seen, in the order of presentation. Their final score was the number of correct items in the correct order. Set sizes (number - math problem pairs) ranged from 3 to 7, and each set was presented 3 times in random order.

In the reading span task, participants were presented with a number that needed to be remembered. Each number was followed by a sentence and participants were asked to determine whether the sentence made sense or not (half of the sentences made sense). At recall, participants were asked to report the presented numbers, in the order of presentation. The final score was the number of correct items in the correct order. Set sizes (number – sentence pairs) ranged from 2 to 6, and each set was presented 3 times in random order.

In the symmetry span task, participants saw a 4×4 grid with one of the cells filled in blue. This was followed by a presentation of an 8×8 grid where some squares were filled, and participants were asked to decide whether the filled square pattern was symmetrical about the vertical axis (the pattern was symmetrical half of the time). At recall, participants were asked to reconstruct the sequence of the previously filled cells, in the order of appearance. The final score was the number of cell locations recalled in the correct order. Set sizes ranged from 2 to 5, and each set was presented 3 times in random order.

We created a composite WM capacity score for each participant, by z -transforming complex span tasks and averaging them (e.g., Harrison, Shipstead, & Engle, 2015; Kane et al., 2007; Unsworth, 2017; the average correlation among WM capacity measures was .47; descriptive statistics for WM capacity measures are provided in Table S1 in the Supplementary Material).

Procedure

Participants were seated in front of the monitor, with head position controlled using a chinrest. The calibration was performed at the beginning of the experiment, while drift check and correction (if required) were performed at the beginning of each block. In the SRT task, participants were instructed to follow the target on the screen. The experiment began with twelve practice trials (randomly generated sequences). The experiment consisted of six blocks, each containing a 12-element sequence repeated five times (i.e., 60 trials within a block). At the beginning of each trial the anticipatory slide (i.e., blank squares) was presented for 500 ms, followed by the presentation of the target (i.e., a black circle inside the square) for 1100 ms. The first four blocks were learning blocks (Block 1 to Block 4). Each of these blocks started from a different point in the sequence. The learning blocks were followed by an interfering block, containing a different 12-element sequence (Block 5). Finally, the original sequence was reintroduced in a recovery block (Block 6). The SRT task, administered in one session, took approximately 15 minutes to complete. After the SRT task, participants filled out the explicit knowledge questionnaire. Finally, the WM capacity battery was administered.

Results

We used the **R** Environment for Statistical Computing (R Core Team, 2018) and **lme4** package (version 1.1-17; Bates, Maechler, Bolker, & Walker, 2015) to fit generalized linear

mixed-effects models with the Binomial link-function (i.e., Logistic GLMM). Type II Wald chi-square tests of models and parameter confidence intervals were obtained using the **car** package (Fox & Weisberg, 2011), while slope analyses were performed and plotted using the **jtools** package (Long, 2018). Additional data visualization was done using the **sjPlot** package (Lüdtke, 2018).

We fit a series of mixed effects logistic regression analyses to our two dependent variables of interest (DVs are analyzed separately): (1) *anticipations* – the overall number of anticipations (correct plus incorrect); and (2) *correct anticipations* – the number of correct anticipations. Anticipations appeared if participants transitioned their gaze towards a different (potential) target location during the presentation of the blank slide (i.e., during the first 500 ms of each trial); anticipations were correct if the participant's gaze remained within the correct AOI at the time the target appeared (otherwise anticipations were incorrect).

We entered the interaction of Block (factor) and WM capacity (covariate) as fixed effects. As random effects we entered intercept for participants: $DV \sim \text{BLOCK} * \text{WMC} + (1|\text{PARTICIPANT})$. The model fit to data was tested against a reduced model with no interaction term: $DV \sim \text{BLOCK} + \text{WMC} + (1|\text{PARTICIPANT})$, and against a null model, containing only a constant term (the intercept): $DV \sim 1 + (1|\text{PARTICIPANT})$. Fixed-effects structures were compared using the anova function and on the basis of the Akaike information criterion (AIC) decreasing with increased model fit. For additional information about the model selection, see Supplementary Material (AIC values across the models are presented in Table S2).

Each best fit model was also compared to a corresponding model including the test predictors (i.e., Block, WMC) and a set of additional control predictors (potential confounding factors). As control predictors we included age, education, and gender (for more details and an

additional control model containing a measure of executive function using a subset of participants, see Supplementary Material). Note that the best fit models outperformed the control models (see Tables S3 and S4).

Below we report the results from different phases of performance: learning (Blocks 1-4), interference (Block 4 vs. Block 5), recovery (Block 5 vs. Block 6), and baseline (Block 1 vs. Block 5).²

In addition, we used the **rle** (run length encoding) function from the **base** package (R Core Team, 2018) to compute the lengths of consecutive correct anticipations. This index represents a measure of chunking or grouping sizes of consecutive correct anticipations across the entire SRT task. The size of chunks (i.e., consecutive correct anticipations; the outcome) was regressed on WM capacity (i.e., the predictor), using generalized linear regression function (glm) from the **car** package (i.e., `CONSECUTIVE_CORRECT_ANTICIPATIONS ~ WMC`).

Anticipations

In our first set of analyses, we investigated the relation between WM capacity and anticipations (as indicators of learning processes) in the SRT task.

Learning (Blocks 1-4). There was an interaction between Block and WM capacity, such that predicted probabilities of anticipations changed for various combinations of Block and WM capacity: $\chi^2(3) = 31.58, p < .001$. (Slopes: Block 1: $b = -.56, SE = .19, p < .001$; Block 2: $b = -.41, SE = .19, p = .003$; Block 3: $b = -.31, SE = .19, p = .09$; Block 4: $b = -.09, SE = .19, p = .63$).

² Because our sample size was relatively limited and the size of the effects of interest could not be inferred from previous research, we addressed the potential power, and parametrization bias issues by conducting a corresponding set of analyses using Bayesian estimation. Note that the estimates from the frequentist and Bayesian models were largely comparable. Where there are differences in statistical significance of the estimates, we signal them in text. Further details about Bayesian analyses are provided in the Supplementary Material. For the full model frequentist and Bayesian estimate comparisons, see Tables S5-S12.

Figure 2 presents the predicted values of anticipations as a function of WM capacity across learning blocks.

Thus, the results of our first analysis suggest that individual differences in WM capacity could be related to implicit sequence learning strategies. Specifically, there was a negative relation between WM capacity and anticipations: anticipations decreased with increased WM span. The strength of this relation decreased across learning blocks, with eventual attenuation in Block 4.

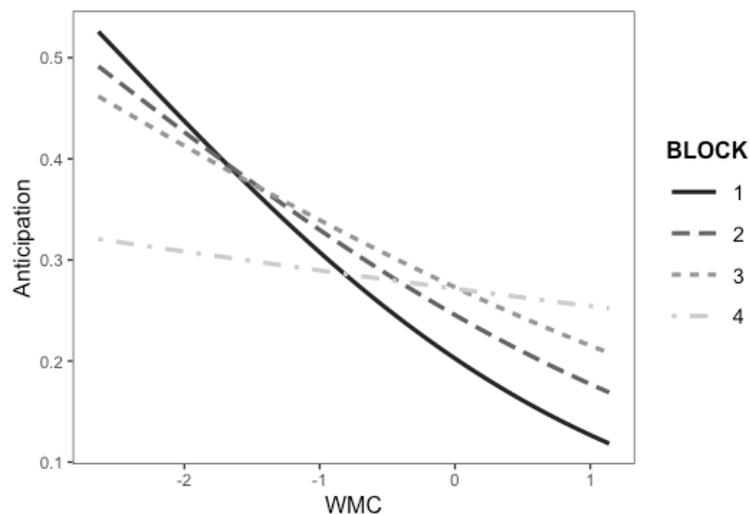


Figure 2. Interaction plot of predicted probabilities of anticipations calculated for working memory capacity levels across individual learning blocks (WMC = working memory capacity, z-score).

Interference (Block 4 vs. Block 5). There was an interaction between Block and WM capacity: $\chi^2(1) = 13.73, p < .001$ (Slopes: Block 4: $b = -.11, SE = .20, p = .58$; Block 5: $b = -.43, SE = .20, p = .03$). Thus, the results of the analysis revealed that WM capacity affected implicit sequence learning strategies. Specifically, there was no effect of WM capacity in the last learning

block (Block 4). However, when another sequence was introduced in the interference block (Block 5), there was a negative relation between WM capacity and anticipations, similar to the starting blocks of the learning phase (Blocks 1 and 2). Again, increased WM capacity predicted fewer anticipations.

Recovery (Block 5 vs. Block 6). There was an interaction between Block and WM capacity: $\chi^2(1) = 9.65, p = .002$ (Slopes: Block 5: $b = -.40, SE = .18, p = .03$; Block 6: $b = -.14, SE = .18, p = .44$). While increased WM capacity was related to a decrease in anticipations when another sequence was introduced (Block 5), there was no effect of WM capacity on anticipations in Block 6, when the original sequence was reintroduced. The Block 6 slope resembled the slopes from the late stages of the learning phase (Block 4 in particular).

Baseline (Block 1 vs. Block 5). There was a significant effect of Block on anticipations: $\chi^2(1) = 19.63, p < .001$, such that anticipations increased in the interference block compared to the baseline. Moreover, there was a significant effect of WM capacity: $\chi^2(1) = 7.81, p = .005$, such that anticipations decreased with increased WM capacity. There was no interaction. Consistent with the results from the other phases of the SRT task, there was a negative relation between WM capacity and anticipations.

The analyses demonstrated that anticipatory behavior changed dynamically across the SRT task as a function of WM capacity. In addition, WM capacity was negatively related to the overall number of anticipations across the task: the greater the WM capacity, the less probable the anticipation.

Correct Anticipations

Next, we investigated the relation between WM capacity and correct anticipations (or learning outcomes).

Learning (Blocks 1-4). Analysis demonstrated that correct anticipations increased across learning blocks: $\chi^2(3) = 17.94, p < .001$ (see Figure 3). There was no effect of WM capacity: $\chi^2(1) = .29, p = .587$, or interaction. These results indicate that, in contrast to the frequency of anticipations, correct anticipations were not affected by individual differences in WM capacity. On the other hand, correct anticipations increased gradually over the learning blocks.

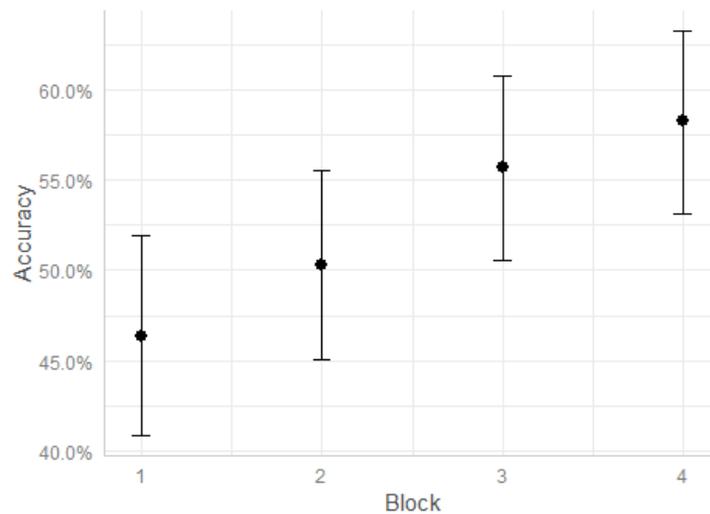


Figure 3. Predicted probabilities of correct anticipations across learning blocks.

Interference (Block 4 vs. Block 5). There was a significant effect of Block on correct anticipations: $\chi^2(1) = 6.27, p = .012$, such that predicted probabilities of correct anticipations decreased in the interference block. There was no effect of WM capacity: $\chi^2(1) = .53, p = .468$, or interaction. Thus, anticipation accuracy was not affected by WM capacity. Overall, participants' anticipations were more accurate in the last learning block, compared to the interference block, where another sequence was introduced.

Recovery (Block 5 vs. Block 6). There was no effect of Block: $\chi^2(1) = 3.80, p = .051$.³ There was no WM capacity effect: $\chi^2(1) = .75, p = .386$ on correct anticipations, and no interaction. Similar to previous phases of the SRT task, WM capacity did not affect anticipation accuracy.

Baseline (Block 1 vs. Block 5). None of the models outperformed the null model. Consistent with the results from the other stages of the SRT task, WM capacity did not affect anticipation accuracy.

Taking all results together, participants clearly demonstrated learning in the SRT task, as indexed by an increase in anticipation accuracy over the four learning blocks, followed by reduced accuracy in the interference block. These results are consistent with previous studies (Marcus et al., 2006; Vakil, et al., 2017). Critically, changes in anticipation accuracy during the SRT task were not related to individual differences in WM capacity.

Correlation Analysis – WM Capacity and Errors

The results of our previous analyses demonstrating a negative relation between WM capacity and anticipations, in conjunction with the lack of a relation between WM capacity and correct anticipations, suggest a negative relation between WM capacity and errors (or incorrect anticipations). To examine this relation, we computed the bivariate correlation between the overall number of errors and WM capacity. There was a negative correlation between the two measures, $r(33) = -.476, p = .004$. Thus, the number of errors decreased with increased WM capacity.

WM Capacity and the Number of Consecutive Correct Anticipations

³ Using Bayesian approach, the effect of Block was statistically significant (as indicated by the 95% credible interval), $estimate = .23, credible\ interval = [.001, .45]$, such that correct anticipations increased in Block 6.

In our final analysis, we investigated the relation between WM capacity and the number of consecutive correct anticipations (i.e., grouping sizes). The latter measure represents the size of chunks of correct anticipations across the SRT task. Here, we considered the chunks that contained at least two consecutive correct anticipations ($M = 2.50$, $SD = 1.03$). The number of consecutive correct anticipations increased with increased WM capacity, $b = .20$ [.08, .32], $SE = .06$, $p < .001$.

The finding that the grouping sizes (or chunks) of consecutive correct anticipations were related to WM capacity represents a strong indication that WM capacity imposes an upper bound of items considered during the SRT task.

Additional Experimental Control: Explicit Sequence Awareness

Out of the participants who filled out the questionnaire ($n = 30$) to assess awareness of the sequence, 31 % ($n = 11$) reported that they noticed something special about the experiment, 49 % ($n = 17$) reported that they noticed a pattern, while 23 % ($n = 8$) indicated that they could recall the pattern. Those participants who indicated that they could recall a pattern ($n = 8$) were asked to generate the sequence; they produced correct strings ranging from 2 to 7 ($M = 3.10$, $SD = 1.45$). These results suggest that although most participants reported that they detected some regularities in the task, few were able to reproduce any chunks from the sequence.

Discussion

The current study investigated the mechanisms underlying implicit learning as measured by a deterministic oculomotor SRT task. We used different anticipation measures as indicators of learning processes and investigated how these measures are affected by individual differences in WM capacity. Our results suggest that the two systems interact in intriguing ways. Specifically, WM capacity influences learning strategies (as measured by the overall number of anticipations),

but not learning outcomes (as measured by correct anticipations). More specifically, our results demonstrate that WM capacity is negatively related to the overall number of errors and positively related to the grouping sizes (or chunks) of consecutive correct anticipations. We will discuss each of our findings in turn, bearing in mind the limitations associated with any correlational approach in establishing a causal link between implicit sequence learning and the underlying cognitive processes.

Anticipation Measures and WM Capacity

Our results demonstrate a negative relation between WM capacity and the overall number of anticipations in the SRT task. Thus, WM capacity seems to influence learning processes by biasing individuals to engage in a more proactive (i.e., more anticipations - lower spans) or reactive (i.e., less anticipations - higher spans) mode during the SRT task. Moreover, the negative relation between WM capacity and anticipations varied in strength across different phases of the SRT task. Specifically, in the learning phase the effect of WM on anticipations decreased gradually across the four learning blocks, followed by a more pronounced effect in the interference block (Block 5), and another decrease in the recovery block (Block 6). Thus, the results of the current study suggest that implicit sequence learning relies on anticipatory processes. Moreover, we demonstrated that learners' anticipatory behavior in the SRT task changed as a combination of individual differences in WM capacity and the environmental cues. At this point, the results showing a negative relation between the predicted probabilities of anticipations and WM capacity could be interpreted from the competition theory point (Galea, Albert, Ditye, & Miall, 2010; Nemeth et al., 2013): weakening the reliance on executive processes (underlying attention-based learning) could have heightened the sensitivity to

statistical probabilities (critical for procedural learning). Thus by this account lower spans could outperform higher spans in this task.

On the other hand, consistent with a number of previous studies (e.g., Caljouw et al., 2016; Frensch & Miner, 1994; Guzmán, 2018; Kaufman et al, 2010; Unsworth & Engle, 2005), our results indicated that there was no relation between implicit sequence learning outcomes (in this case indexed by correct anticipations) and WM capacity. Thus, higher WM capacity did not lead to more overall accurate anticipations. Previously, these findings were interpreted to indicate the relative automaticity of implicit learning (Kaufman et al., 2010; Unsworth & Engle, 2005). Results from the explicit awareness questionnaire in the current study further support the notion that the majority of participants lacked awareness about the *sequentiality* of the stimuli presentation in the SRT task.

Finally, although the results of the current study indicated no relation between WM capacity and learning outcomes (indexed by correct anticipations) - a finding in line with previous research (e.g., Caljouw et al., 2016; Frensch & Miner, 1994; Guzmán, 2018; Kaufman et al, 2010; Unsworth & Engle, 2005), we demonstrated that higher WM capacity was related to larger groupings (or chunks) of correct anticipations. While previous studies have consistently reported a positive correlation between WM capacity and chunking during explicit sequence learning (Bo et al., 2009; Kennerley et al., 2004; Sakai et al., 2003; Shea et al., 2006), our study is the first to demonstrate the existence of this relation in implicit sequence learning. In previous studies where sequence learning occurred explicitly (i.e., with conscious intent), the relation between WM and chunking was interpreted to indicate the WM-dependent performance strategies during learning (Bo et al., 2009; Wymbs, Bassett, Mucha, Porter, & Grafton, 2012). Thus, the nuanced qualitative differences in implicit learning outcomes reported in the current

study would represent another indication of different efficient learning strategies in lower and higher WM spans during implicit learning. This is further supported by the negative relation between WM capacity and errors.

The results of the current study are theoretically important because they demonstrate that individual differences in WM capacity could account for differences in learning processes, and ultimately change individuals' anticipatory behavior, even when learning is implicit, without intention and awareness. Specifically, the negative relation between WM capacity and anticipations, together with the demonstrated positive relation between WM capacity and the number of consecutive correct anticipations, indicate that WM capacity imposes an upper bound on the number of sequential items that can be considered during the task. Anticipating at a higher rate in lower spans could be conceptualized as "starting small", a notion that was the focus of several studies across relatively independent domains (cf., Newport 1990; Elman, 1993; Medimorec, Mander, & Risko, 2018). "Starting small" reveals economies in learning by leveraging an individual's processing capacity against the size of their ideal learning unit or chunk. Furthermore, and in the same vein, higher anticipations can be argued to represent a straightforward strategy to reduce memory load with the aim of optimizing learning. On the other hand, higher spans are less constrained, and do not necessarily need to reduce the information load as much as lower spans. This allows higher spans to chunk probable patterns into larger units. Hence, the results of the current study indicate that individuals' anticipatory reactions to environmental cues change as a function of WM capacity. The results also suggest that WM processes, such as context-relevant updating of information and the formation of anticipations, can take place implicitly.

Implications for Future Research

The results of the current study clearly support the notion that individual differences in WM capacity are, at least partly, involved in implicit sequence learning. Deeper insights into the relations between different aspects of WM (i.e., storage and processing) and implicit sequence learning (using more complex probabilistic sequences), but also between different executive functions (e.g., relational integration, imagery, attention, updating, switching and inhibition; Janacek & Nemeth, 2015) and implicit sequence learning are much needed. Given increased interest in examining the links between language cognition and implicit learning (Arciuli, & Simpson, 2012; Kidd, 2012; Daltrozzo et al., 2017; Kidd & Arciuli, 2016; Milin, Divjak, & Baayen, 2017; Misyak & Christiansen, 2012; Shafto, Conway, Field, & Houston, 2012), determining how language learning and processing are supported by implicit learning remain important questions.

In addition, recent research across a number of domains has demonstrated that individuals use different strategies, such as explorative and exploitative actions when learning and adapting to environmental (conditional) probabilities (Dale et al., 2012; Fischer, & Holt, 2017, Milin et al., 2017; Stafford & Dewar 2014). Further investigating whether anticipatory behavior during implicit learning demonstrates similar explorative/exploitative dynamics should reveal hitherto un(der)explored dimensions of learning and advance our understanding of how individuals adapt to environmental statistics, for example during language learning.

Conclusion

Our results provide support for the existence of general cognitive strategies that are employed spontaneously during implicit sequence learning. Critically, such strategies seem to be efficient in detecting regularities in the environment and are modulated by individual differences in WM capacity. Specifically, WM capacity affects learning strategies, as demonstrated by the

negative relation between WM capacity and the rate of anticipation as well as between WM capacity and error rate, and the positive relation between WM capacity and the grouping size (or chunk size) of consecutive correct anticipations. On the other hand, WM capacity does not affect learning, as measured by the overall number of accurate anticipations.

Acknowledgments: We would like to thank Maciej Borowski and Dagmar Hanzlikova for their help with the data collection, and Adnane Ez-Zizi for statistical consultation.

Compliance with ethical standards

Funding: This research was supported by a Leverhulme Trust Research Leadership Award (RL-2016-001) to D. Divjak, which funded all authors.

Availability of data: The dataset supporting the conclusions of this article is available as an electronic supplementary material.

Conflict of interest: The authors declare that they have no conflict of interest.

Ethical approval: All procedures performed were in accordance with the ethical standards of the institutional research committee and with the 1964 Helsinki declaration and its later amendments or comparable ethical standards.

Informed consent: Informed consent was obtained from all individual participants included in the study.

References

- Arciuli, J., & Simpson, I. C. (2012). Statistical learning is related to reading ability in children and adults. *Cognitive Science*, *36*, 286–304
- Bates, D., Maechler, M., Bolker, B., & Walker, S., (2015). Fitting linear mixed-effects models using lme4. *Journal of Statistical Software*, *67*, 1–48.
- Bo, J., Borza, V., & Seidler, R. D. (2009). Age-related declines in visuospatial working memory correlate with deficits in explicit motor sequence learning. *Journal of Neurophysiology*, *102*, 2744-2754.
- Bo, J., & Seidler, R. D. (2009). Visuospatial working memory capacity predicts the organization of acquired explicit motor sequences. *Journal of neurophysiology*, *101*, 3116–3125.
- Bo, J., Jennett, S., & Seidler, R. (2012). Differential working memory correlates for implicit sequence performance in young and older adults. *Experimental Brain Research*, *221*, 467–477.
- Caljouw, S. R., Veldkamp, R., & Lamoth, C. J. C. (2016). Implicit and explicit learning of a sequential postural weight-shifting task in young and older adults. *Frontiers in Psychology*, *7*.
- Cleeremans, A., & McClelland, J. L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General*, *120*, 235–253.
- Conway, C. M., Bauernschmidt, A., Huang, S. S., & Pisoni, D. B. (2010). Implicit statistical learning in language processing: Word predictability is the key. *Cognition*, *114*, 356-371.
- Dale, R., Duran, N. D., & Morehead, J. R. (2012). Prediction during statistical learning, and implications for the implicit/explicit divide. *Advances in Cognitive Psychology*, *8*, 196–209.

- Daltrozzo, J., Emerson, S. N., Deocampo, J., Singh, S., Freggens, M., Branum-Martin, L., & Conway, C. M. (2017). Visual statistical learning is related to natural language ability in adults: an ERP study. *Brain and Language, 166*, 40–51.
- Elman, J. L. (1993). Learning and development in neural networks: the importance of starting small. *Cognition, 48*, 71-99.
- Erickson, L. C., & Thiessen, E. D. (2015). Statistical learning of language: Theory, validity, and predictions of a statistical learning account of language acquisition. *Developmental Review, 37*, 66-108.
- Feldman, J., Kerr, B., & Streissguth, A. P. (1995). Correlational analyses of procedural and declarative learning performance. *Intelligence, 20*, 87–114.
- Fischer, H., & Holt, D. V. (2017). When high working memory capacity is and is not beneficial for predicting nonlinear processes. *Memory & Cognition, 45*, 404-412.
- Fox, J., & Weisberg, S. (2011). *An {R} Companion to Applied Regression*, Second Edition. Thousand Oaks CA: Sage.
- Gabriel, A., Maillart, C., Stefaniak, N., Lejeune, C., Desmottes, L., & Meulemans, T. (2013). Procedural learning in specific language impairment: Effects of sequence complexity. *Journal of the International Neuropsychological Society, 19*, 264–271.
- Galea, J. M., Albert, N. B., Ditye, T., & Miall, R. C. (2010). Disruption of the dorsolateral prefrontal cortex facilitates the consolidation of procedural skills. *Journal of Cognitive Neuroscience, 22*, 1158-1164.
- Guzmán, F. J. (2018). The influence of personality and working memory capacity on implicit learning. *Quarterly Journal of Experimental Psychology, 71*, 2603–2614.

- Hasson, U., Cashdollar, N., Weisz, N., & Ruhau, P. (2016). The role of working memory in the probabilistic inference of future sensory events. *Cerebral Cortex*, *27*, 2955-2969.
- Harrison, T. L., Shipstead, Z., & Engle, R. W. (2015). Why is working memory capacity related to matrix reasoning tasks? *Memory & Cognition*, *43*, 389-396.
- Hunt, R. H., & Aslin, R. N. (2001). Statistical learning in a serial reaction time task: access to separable statistical cues by individual learners. *Journal of Experimental Psychology: General*, *130*, 658-680.
- Janacsek, K., & Nemeth, D. (2013). Implicit sequence learning and working memory: correlated or complicated? *Cortex*, *49*, 2001-2006.
- Janacsek, K., & Nemeth, D. (2015). The puzzle is complicated: when should working memory be related to implicit sequence learning, and when should it not? (Response to Martini et al.). *Cortex*, *64*, 411-412.
- Jimenez, L., & Vazquez, G. A. (2005). Sequence learning under dual-task conditions: alternatives to a resource-based account. *Psychological Research*, *69*, 352-368.
- Jongbloed-Pereboom, M., Nijhuis-van der Sanden, M. W. G., & Steenbergen, B. (2019). Explicit and implicit motor sequence learning in children and adults; the role of age and visual working memory. *Human Movement Science*, *64*, 1-11.
- Kane, M. J., Brown, L. H., McVay, J. C., Silvia, P. J., Myin-Germeys, I., & Kwapil, T. R. (2007). For whom the mind wanders, and when: an experience sampling study of working memory and executive control in everyday life. *Psychological Science*, *18*, 614-621.
- Kaufman, S. B., DeYoung, C. G., Gray, J. R., Jiménez, L., Brown, J., & Mackintosh, N. (2010). Implicit learning as an ability. *Cognition*, *116*, 321-340.

- Kennerley, S. W., Sakai, K., & Rushworth, M. F. (2004). Organization of action sequences and the role of the pre-SMA. *Journal of Neurophysiology* 91, 978–993.
- Kidd, E. (2012) Implicit statistical learning is directly associated with the acquisition of syntax. *Developmental Psychology* 48, 171–184.
- Kidd, E., & Arciuli, J. (2016). Individual differences in statistical learning predict children's comprehension of syntax. *Child Development*, 87, 184–193.
- Kinder, A., Rolfs, M., & Kliegl, R. (2008). Sequence learning at optimal stimulus–response mapping: evidence from a serial reaction time task. *The Quarterly Journal of Experimental Psychology*, 61, 203–209.
- Long, J. A. (2018). jtools: Analysis and presentation of social scientific data. R package version 1.0.0, <URL: <https://cran.r-project.org/package=jtools>>.
- Lüdecke, D. (2018). sjPlot: data visualization for statistics in social science. R package version 2.6.0, <<https://CRAN.R-project.org/package=sjPlot>>.
- Marcus, D. J., Karatekin, C., & Markiewicz, S. (2006). Oculomotor evidence of sequence learning on the serial reaction time task. *Memory & Cognition*, 34, 420–432.
- Martini, M., Furtner, M. R., & Sachse, P. (2013). Working Memory and Its Relation to Deterministic Sequence Learning. *PLOS ONE* 8: e56166.
- Masters, R. S. (1992). Knowledge, knerves and know-how. The role of explicit versus implicit knowledge in the breakdown of complex motor skill under pressure. *British Journal of Psychology*, 83, 343–358.
- Mathôt, S., Schreij, D., & Theeuwes, J. (2012). OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods*, 44, 314–324.

- Medimorec, S., Mander, C., & Risko, E. F. (2018). Memory demands in linguistic compensation. *The Quarterly Journal of Experimental Psychology*, *71*, 1234-1239.
- Meissner, S. N., Keitel, A., Südmeyer, M., & Pollok, B. (2016). Implicit motor sequence learning and working memory performance changes across the adult life span. *Frontiers in Psychology*, *8*.
- Milin, P., Divjak, D., & Baayen, R. H. (2017). A learning perspective on individual differences in skilled reading: Exploring and exploiting orthographic and semantic discrimination cues. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *43*, 1730–1751.
- Misyak, J. B., & Christiansen, M. H. (2012). Statistical learning and language: an individual differences study. *Language Learning*, *62*, 302–331.
- Nemeth, D., Janacsek, K., Polner, B., & Kovacs, Z. A. (2013). Boosting human learning by hypnosis. *Cerebral Cortex*, *23*, 801-805.
- Newport, E. L. (1990). Maturation constraints on language learning. *Cognitive Science*, *14*, 11-28.
- Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: evidence from performance measures. *Cognitive Psychology*, *19*, 1–32.
- R Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>.
- Sakai, K., Kitaguchi, K., & Hikosaka, O. (2003). Chunking during human visuomotor sequence learning. *Experimental Brain Research*, *152*, 229-242.
- Schvaneveldt, R. R., & Gomez, R. L. (1998). Attention and probabilistic sequence learning. *Psychological Research*, *61*, 175–190.

- Schwarb, H., & Schumacher, E. H. (2012). Generalized lessons about sequence learning from the study of the serial reaction time task. *Advances in Cognitive Psychology*, *8*, 165–178.
- Shafto, C. L., Conway, C. M., Field, S. L., & Houston, D. M. (2012). Visual sequence learning in infancy: domain-general and domain-specific associations with language. *Infancy*, *17*, 247-271.
- Shea, C. H., Park, J. H., & Braden, H. W. (2006). Age-related effects in sequential motor learning. *Physical Therapy*, *86*, 478 – 488.
- Stadler, M. A. (1989). On learning complex procedural knowledge. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *15*, 1061–1069.
- Stadler, M. A., & Frensch, P. A. (1998). *Handbook of Implicit Learning*. Thousand Oaks, CA: Sage Publications.
- Stafford, T., & Dewar, M. (2014). Tracing the trajectory of skill learning with a very large sample of online game players. *Psychological Science*, *25*, 511-518.
- Stark-Inbar, A., Raza, M., Taylor, J. A., & Ivry, R. B. (2016). Individual differences in implicit motor learning: task specificity in sensorimotor adaptation and sequence learning. *Journal of Neurophysiology*, *117*, 412-428.
- Torriero, S., Oliveri, M., Koch, G., Gerfo, E. L., Salerno, S., Petrosini, L., & Caltagirone, C. (2007). Cortical networks of procedural learning: evidence from cerebellar damage. *Neuropsychologia*, *45*, 1208-1214.
- 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.

- Turk-Browne, N. B., Scholl, B. J., Chun, M. M., Johnson, M. K. (2008). Neural evidence of statistical learning: efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience, 21*, 1934–1945.
- Unsworth, N. (2017). Examining the dynamics of strategic search from long-term memory. *Journal of Memory and Language, 93*, 135–153.
- Unsworth, N., & Engle, R. W. (2005). Individual differences in working memory capacity and learning: Evidence from the serial reaction time task. *Memory and Cognition, 33*, 213–220.
- Vakil, E., Bloch, A., & Cohen, H. (2017). Anticipation measures of sequence learning: manual versus coulometer versions of the serial reaction time task. *Quarterly Journal of Experimental Psychology, 70*, 579–589.
- Virag, M., Janacek, K., Horvath, A., Bujdosó, 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*, 2081–2089.
- von Bastian, C. C., Locher, A., & Rufin, M. (2013). Tatool: a java-based open-source programming framework for psychological studies. *Behavior Research Methods, 45*, 108–115.
- Weitz D., O’Shea G., Zook N., & Needham W. (2011). Working memory and sequence learning in the Hebb digits task: awareness is predicted by individual differences in operation span. *The American Journal of Psychology, 124*, 49–62.
- West, G., Vadillo, M. A., Shanks, D. R., & Hulme, C. (2018). The procedural learning deficit hypothesis of language learning disorders: we see some problems. *Developmental Science, 21*, e12552.

Wilkinson, L., & Shanks, D. R. (2004). Intentional control and implicit sequence learning.

Journal of Experimental Psychology: Learning, Memory, & Cognition, *30*, 354–369.

Wymbs, N. F., Bassett, D. S., Mucha, P. J., Porter, M.A., & Grafton, S. T. (2012). Differential

recruitment of the sensorimotor putamen and frontoparietal cortex during motor chunking

in humans. *Neuron* *74*, 936–946.