Cognition 201 (2020) 104291

Contents lists available at ScienceDirect

Cognition

journal homepage: www.elsevier.com/locate/cognit

How sequence learning unfolds: Insights from anticipatory eye movements

Amir Tal[a,*], Eli Vakil[a,b]

[a] Gonda Multidisciplinary Brain Research Center, Bar-Ilan University, Ramat-Gan 52900, Israel
[b] Department of Psychology, Bar-Ilan University, Ramat-Gan 52900, Israel

ARTICLE INFO

Keywords:
Sequence learning
SRT
Implicit learning
Statistical learning
Chunking
Eye tracking

ABSTRACT

The acquisition of sequential knowledge is pivotal in forming skilled behavior. Despite extensive research of sequence learning, much remains unknown regarding what knowledge participants learn in such studies, and how that knowledge takes form over time. By tracking eye-movements made before stimuli appear on screen during a serial reaction time (SRT) task, we devised a method for assessing learning at the individual participant level in an item-based resolution. Our method enables uncovering what participants actually learn about the sequence presented to them, and when. Results demonstrate that learning is more heterogeneous than previously thought, driven by learning both of chunks and of statistics embedded in the sequence. Also, learning develops rapidly, but in a fragmented and non-sequential manner, eventually encompassing only a subset of available regularities. The tools developed in this work may aid in further dissociating processes and mechanisms underlying sequence learning and its impairments, in normal and in clinical populations.

1. Introduction

A fundamental prerequisite for acquiring a skill, from brushing teeth to understanding speech, is the ability to chain together information and actions into an efficient, unitized procedure. This ability is called sequence learning. It has long been acknowledged in the field of psychology (Lashley, 1951), but the paradigm most research of sequence learning is based upon is the serial reaction time (SRT) task, introduced in Nissen and Bullemer (1987).

In SRT, participants are given a mapping between stimuli and responses, and are required to follow it when presented with a stream of stimuli. Unbeknownst to them, stimuli are presented in a repeating sequential order. The typical finding of SRT is that response time (RT) grows quicker with exposure to this order. Accordingly, RT sharply slows down if the ordering is changed (interference effect) and returns to its original speed if it is restored (recovery effect; Fig. 1A – fixed-sequence group). These effects are taken as evidence that with exposure to the stimuli stream participants learn the underlying sequential order, thereby anticipating upcoming targets and acting upon them with enhanced speed and accuracy (Marcus, Karatekin, & Markiewicz, 2006).

A striking finding in SRT is that learning does not depend on explicit awareness of the sequence (Nissen & Bullemer, 1987). Participants are typically left with partial to no explicit knowledge of what they have learned, and SRT had subsequently become one of the principle paradigms for studying implicit learning (Destrebecqz & Cleeremans, 2001). Because knowledge may be at least partially implicit, understanding what is learned during SRT has been a considerable challenge in the field. Significant progress has been made in uncovering the type of knowledge acquired – whether it is the order of response selection or the order of response execution that participants learn (for review see Schwarb & Schumacher, 2012). Progress has also been made regarding the mechanism driving this learning – whether participants form representations of chunks within the stimuli stream (Koch & Hoffmann, 2000), acquire statistical knowledge pertaining to that stream (Jiménez, 2008), or both (Kirsch, Sebald, & Hoffmann, 2010). However, what actual knowledge participants have acquired by the end of the task remains unknown (Reed & Johnson, 1994).

To answer this question, a meticulous item-based analysis is needed at the level of the individual participant. In this work we offer a new method for dynamic item- and individual-based analysis of sequence learning during an SRT task. It is based on a new metric of performance in SRT, that of oculomotor anticipation.

In spatial SRT tasks, participants shift their gaze towards the location of upcoming targets in direct proportion to their familiarity with the sequence governing the task (Marcus et al., 2006; Vakil, Bloch, & Cohen, 2017). The major effects of SRT, including interference and recovery effects, can be found in such anticipatory gaze shifts just as in RT (Fig. 1B – fixed-sequence group). Oculomotor anticipation, however, offers additional new possibilities of analysis in comparison to RT. Mainly, the oculomotor signal informs on mistakes participants make in

---

[*] Corresponding author at: Department of Psychology, Columbia University, New York, NY 10027, USA.
E-mail addresses: at3405@columbia.edu (A. Tal), Eli.Vakil@biu.ac.il (E. Vakil).

https://doi.org/10.1016/j.cognition.2020.104291
Received 10 June 2019; Received in revised form 31 March 2020; Accepted 3 April 2020

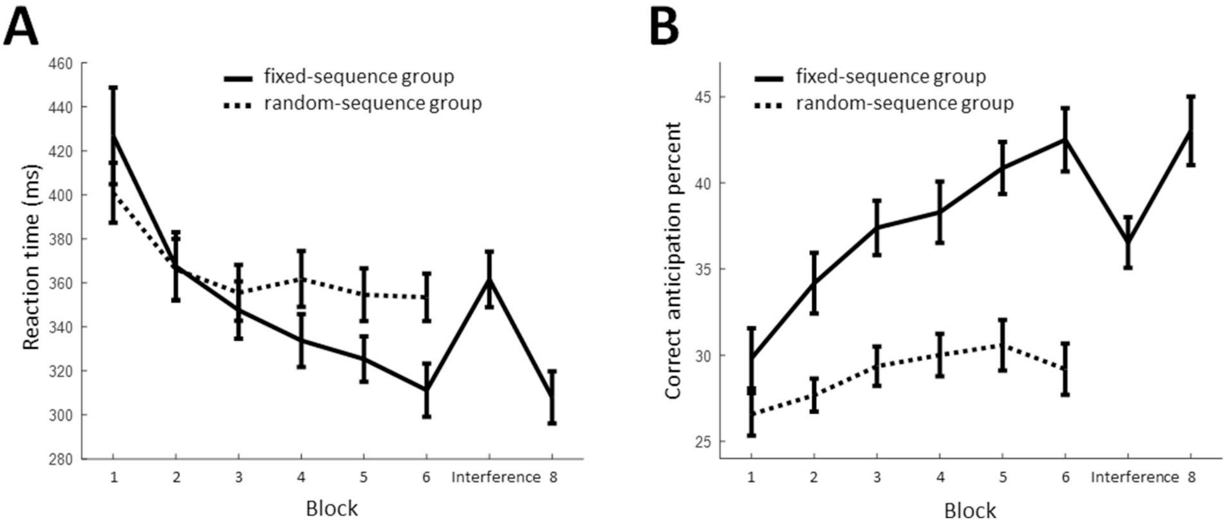

Fig. 1. A) In standard SRT (fixed-sequence group), RT drops throughout the learning phase (learning effect), but abruptly rises when the sequence is switched (interference effect). When the sequence is reverted back to the original in the final block, RT is restored to its previous level (recovery effect). When there is no fixed sequence governing the task (random-sequence group), a smaller learning effect is found in blocks 1–6. B) Percentage of correct oculomotor anticipations mirrors learning, interference and recovery effects found in RT. Again, robust effects exist in the fixed-sequence group, while anticipations in the random-sequence group reflects a smaller learning effect. Error bars represent SEM.

their anticipation of the stimuli sequence, information that does not exist in RT. The rates and types of these mistakes provide insight on the learning processes taking place while the task unfolds in time. Using probabilistic analysis of correct and incorrect anticipation, we manage to pinpoint which items are learned within an SRT sequence, and the time in which this learning occurs.

In previous work we have demonstrated that two major types of learning processes take place during SRT: statistical learning and sequence learning (Tal, Bloch, Cohen-Dallal, Aviv, Ashkenazi, Bar, & Vakil, submitted). Statistical learning provides participants with understanding of the task “grammar”; that is, understanding what type of target transitions could be expected and what type of transitions could not take place in the current task. Sequence learning, on the other hand, provides participants with knowledge of the specific sequence governing the task. The focus of the current work is on the evolvement of sequence learning throughout SRT.

Naturally, when a single sequence drives the stimuli stream, as in standard SRT, that sequence both dictates and conforms to the statistical grammar of the task, and so statistical learning and sequence learning facilitate each other and are difficult to tease apart. To isolate effects of sequence learning from those of statistical learning, a control group was therefore included in this study. In this group, the task stimuli stream was governed by the same statistical grammar as the standard SRT task, but contained no fixed-sequence that could be learned. This paper begins with an item-based analysis of these two groups, highlighting the different outcomes afforded by sequence learning. The remainder of the paper focuses on the standard SRT group and the insights that can be learned on how sequence learning evolved over the course of the experiment.

2. Materials and methods

2.1. Participants

Fifty-nine undergraduate students (37 females, mean age = 24.7, range: 18–37) participated in the experiment for course credit or 30 NIS (~$8.5 US). Participants encompassed two experimental groups. The fixed-sequence group (n = 29) correspond to the “MA group” whose data was originally collected in the experiment published in Vakil et al., (2017). The random-sequence group (n = 30) were recruited for this study.

The study was approved by the ethics committee of the Psychology Department in Bar-Ilan University, and each participant gave written informed consent.

2.2. Stimuli

Four white squares, in diamond formation, were presented against grey background on an LCD computer screen of 1680 × 1050 pixel resolution (size 47 × 29 cm). A black dot indicating the target could appear in the center of any one of the squares, or in neither. Squares were of size 6 × 6 cm, and dots of 1.5 cm diameter.

2.3. Procedure

Participants were seated in front of a computer screen and were asked that when targets (dots) appear on screen they locate them in their vision as soon as possible and press a corresponding keyboard key. Keys were keypad arrow keys that corresponded to target location: up arrow for the top square, left arrow for the left square, etc. Each target was presented for 3000 ms or until a key was pressed. A 500 ms screen with no target acted as an inter-stimulus interval (ISI), separating the disappearance of one target from the appearance of the next (Fig. 2A).

For participants in the fixed-sequence group (n = 29), target locations followed a fixed sequential order of length twelve, in which location frequency and first order transition probabilities are counterbalanced. Such an order is termed second order conditional (SOC) because the location of the next target is completely determined by the last two locations in the stream, while only one or zero previous locations are not enough to predict the next target location (Reed & Johnson, 1994). Two such sequences were used, *sequence A*: *3-4-2-3-1-2-1-4-3-2-4-1* and *sequence B*: *3-4-1-2-4-3-1-4-2-1-3-2* (numbers corresponding to location: *1*-down, *2*-left, *3*-right, *4*-up). Nine concatenated sequences constituted one block (108 stimuli). The experiment consisted of eight blocks in total, each starting from a different position within the sequence: locations 1, 5, 10, 8, 4, 12, 1, 2 for blocks one to eight respectively. Blocks one to six and eight were constructed from one sequence, while the seventh block, called the *interference block*, was constructed from the other. Participants were randomly assigned into those whose main sequence was *sequence A* and their interference sequence was *sequence B*, and those who had it the other way around (n = 15 & n = 14, respectively).

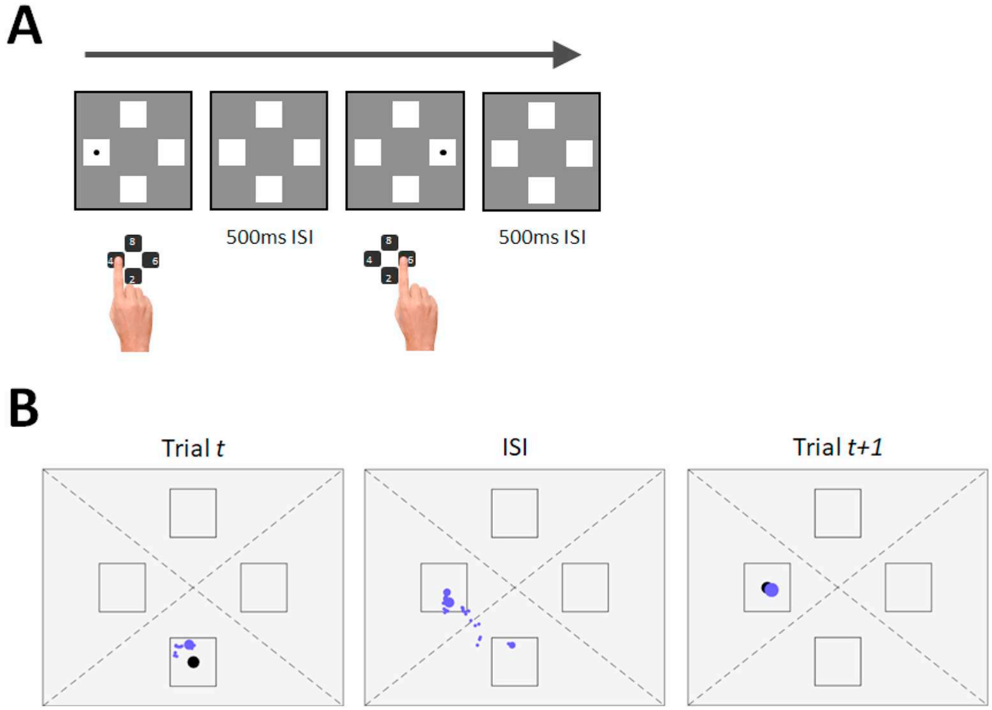

Fig. 2. A) Experimental paradigm. Participants are asked to press the arrow key that corresponds to the location a target dot occupies on screen. Target appearances are separated by an ISI of 500 ms in which there is no target. B) Fixation analysis. Gaze location is tracked throughout the experiment. Only fixations during ISIs are analyzed, but all eye movements, including saccades, of a single participant throughout two trials and the ISI in between them are plotted here for illustration. The screen is divided into four equally-sized triangle-shaped AOIs based on proximity to target location. In this example, a target appeared at the bottom location on trial *t*, and indeed the participant fixated around that location. During the following ISI, fixations (marked as larger blue circles, size reflecting fixation duration) were detected in the bottom AOI, and then in the left AOI. In the subsequent trial *t* + *1*, the next target appeared in the left location, meaning that the fixation during the ISI is regarded a correct anticipation in analysis. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

For participants in the random-sequence group ($n = 30$), an order of 648 target locations was designed so that it maximally mimics the location statistics used in the fixed-sequence group, but without containing a fixed sequential order. The order was obtained by computationally searching for an arrangement that maximally complies with all desired statistical constraints, as are elaborated next. Resulting frequencies are therefore close but not precisely uniform. First, all locations appeared at near uniform frequency (*M* = 25% *SD* = 0.3%). Second, locations did not repeat back-to-back but all other first order transitions (e.g. *1*, *3* or *4* after *2*) appeared at near uniform frequency (*M* = 33.3% *SD* = 0.9%). Third, twelve triplets that constitute a reversal (e.g. *2-1-2*, *2-3-2*, *2-4-2*, …) all occurred at a similar frequency (*M* = 1% *SD* = 0.1%), comprising together 11.7% of the stimuli stream, which is slightly higher than the 8.8% they constitute in the fixed-sequence stimuli (due to one reversal that inhabits that order). Lastly, all other second order transitions (e.g. *3* or *4* after *2-1*) appeared at near uniform frequency (*M* = 44.1% *SD* = 1.4%). The complete stimuli order can be found in Supplementary Fig. 1. The goal of this procedure was to isolate learning of stimuli statistics from learning of a particular sequence. An order of only six blocks was used because the seventh interference and eighth recovery blocks are irrelevant when no fixed sequence is used.

2.4. Data acquisition

Participants' keypress times and eye movements were collected throughout the experiment. Eye movements were captured using SMI iView 250 RED Eye Tracker. Calibration was done at experiment onset using a standard nine-point grid for both eyes.

2.5. Data analysis

The experiment is broken down into three phases: learning (blocks 1–6), interference (block 7) and recovery (block 8). Participants of the fixed-sequence group underwent all eight blocks, while participants of the random-sequence group underwent only the first six blocks. RT corresponds to the time since stimuli appeared and until the correct key was pressed. Trials of incorrect or no response were removed from analysis (0.8% of the data).

Eye movements captured during ISIs underwent fixation analysis. Fixations were detected using the SMI built-in velocity based algorithm. They were found in 99% of ISIs. Fixations were considered anticipation of an upcoming target in the target location that they were closest to, effectively dividing the screen into four distinct areas of interest (AOI; Fig. 2B). When fixations were detected in more than one AOI during a single ISI (56% of ISIs) the first fixated AOI was predominantly that of the previous target location (91%). Fixations were found in more than one novel AOI only in 5.1% of ISIs. Therefore, when more than one AOI was fixated during an ISI, AOIs fixated last were considered the participant's anticipation for analysis.

Throughout the paper, in cases in which sphericity was violated according to Mauchly's test ($p < .05$), Greenhouse–Geisser corrected values are reported.

2.6. Item-based analysis of sequence learning

In SOC sequences, the task stimuli stream is completely predictable based on second order transition probabilities. Each pair of neighboring targets is a “predictor” of a specific target that follows them. For complete learning of a sequence of length twelve, participants therefore need only learn where to orient their gaze after each of twelve predictors, i.e. twelve responses to twelve different stimuli. Each such triplet is one item of the sequence. In this analysis we therefore focus on where participants shifted their gaze following encounter with each stimuli pair, and treat this as their “response” to the pair (Marcus et al., 2006). We consider a participant to have learned a response if, whenever a pair of stimuli is seen, the participant consistently performs that response. Also, to be regarded learned, this response should persist until the end of the learning phase.

Therefore, per each pair of stimuli in the stimuli stream, each participant's responses are examined in all learning phase ISIs following its appearance (“relevant responses” in Fig. 3A). There are four possible target locations on screen, and so chance probability of producing any response is 25%. We examine all responses in this analysis, both correct and incorrect ones. The amount of each response is checked for surpassing the amount expected by chance, within two separate windows.

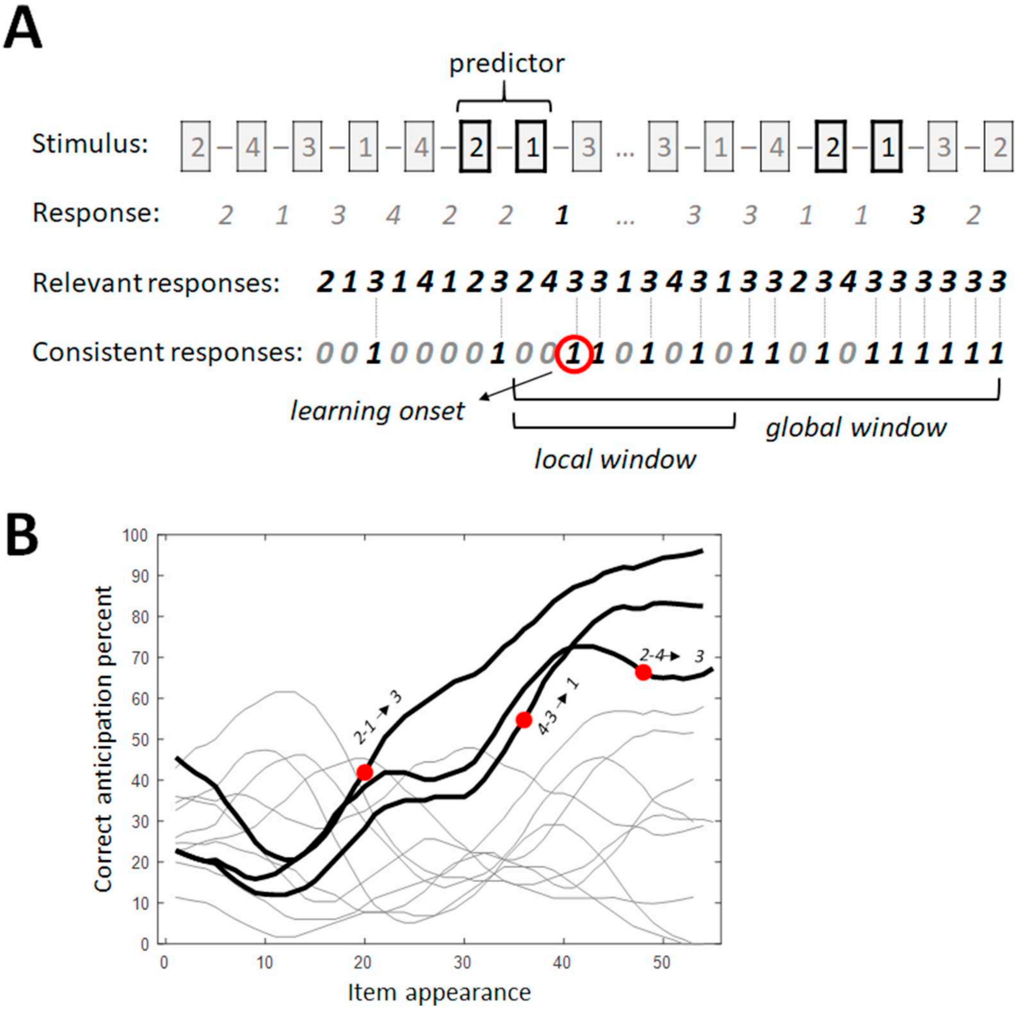

Fig. 3. A) Item-based analysis. The neighboring targets *2-1* are a predictor of an upcoming target *3* in the stimuli stream (upper row; numbers corresponding to location). A participant's fixation location in between targets is regarded her "response" to the targets that preceded it (2nd row). In this example the response *3*, which happens to be the correct response, is tested for learning. All responses following the predictor (relevant responses, 3rd row) are examined together, as a binary vector equal or not to the response under examination (4th row). Response consistency is examined in a local and a global window for passing binomial significance, with chance level of 25% and *p*-values of 0.05 and 0.01, respectively. Starting point moves back in time until performance in at least one of the windows is not significant. In this case, the local window fails at the ninth response. The first response from which both windows are significant is therefore the eleventh response, established as the learning onset of the response *3* to the stimuli pair *2-1*. B) Outcome of analysis of correct responses in one participant's data. Each line represents correct anticipation of a single item of the sequence (twelve overall) throughout the entire learning phase (averaged and smoothed over a sliding window). Bold lines indicate correct responses that have been identified as learned (their identity is given in the inset), and red circles mark learning onset. Non-learned items fluctuate around 25% success rate, as would be expected by chance. Only the three learned responses break away from chance in a consistent manner until the end of learning phase. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

Responses are examined under a global and under a local window, starting from the same point in time. The global window encompasses all ISIs from that point until the end of the learning phase, while the local window covers only the nine ISIs coming after that same point (nine appearances of each pair of stimuli extend the length of a single experiment block). Significance is tested according to a right-tailed binomial test, with chance level of 25% and *p*-values of $p \le 0.05$ and $p \le 0.01$ for local and global windows, respectively. Starting from the last nine ISIs, if both windows are significant, the response is considered to have been learned (the two windows completely overlap in this initial test, and so the stricter condition of the global window determines de facto whether the response is classified as learned or not). If learning is established, a search for learning onset is performed by shifting both windows back in time (the global window growing, the local window moving) until the earliest same response that satisfied the significance criteria of both windows is found.

This method locates the first moment from which a participant consistently responded in a certain way to a stimuli pair, i.e. in a significantly above-chance rate. While the global window is seemingly sufficient to test that, the local window ensures that significance is not due to any particular segment of time, but rather that the behavior was consistent in every step of the way. Otherwise, abundance of a single response confined to one segment in time would bestow significance upon large timescales, including remote chance responses. This method is based on a simplified model of learning and so undoubtedly misses some information, mainly in its binary classification of continuous learning. However, we wish to show that it is nevertheless useful in understanding how sequence learning unfolds over time. See Fig. 3 for an illustration of the method and example results.

3. Results

We have analyzed learning of all possible responses (correct and incorrect) for all stimuli pairs appearing in the stream. A breakdown of this analysis is given in Fig. 4A.

3.1. *Grammatical vs. ungrammatical learning*

First, we have examined the grammaticality of learned responses. Ungrammatical responses are those that do not conform to the statistical rules governing the task (rules that are shared both by the fixed-sequence and the random-sequence tasks). These responses are either remaining in the location of the last target ("stuck") or reverting the location of the target before last ("reversal"). In contrast to the former, reversals do appear in the task (there is one reversal in each sequence). But, because transitions do not revert to the location before last in the vast majority of the task (in over 90% of trials), having no reversals can be considered part of the task grammar, and the odd reversal an irregularity (Cleeremans & McClelland, 1991).

All participants have acquired substantially more grammatical

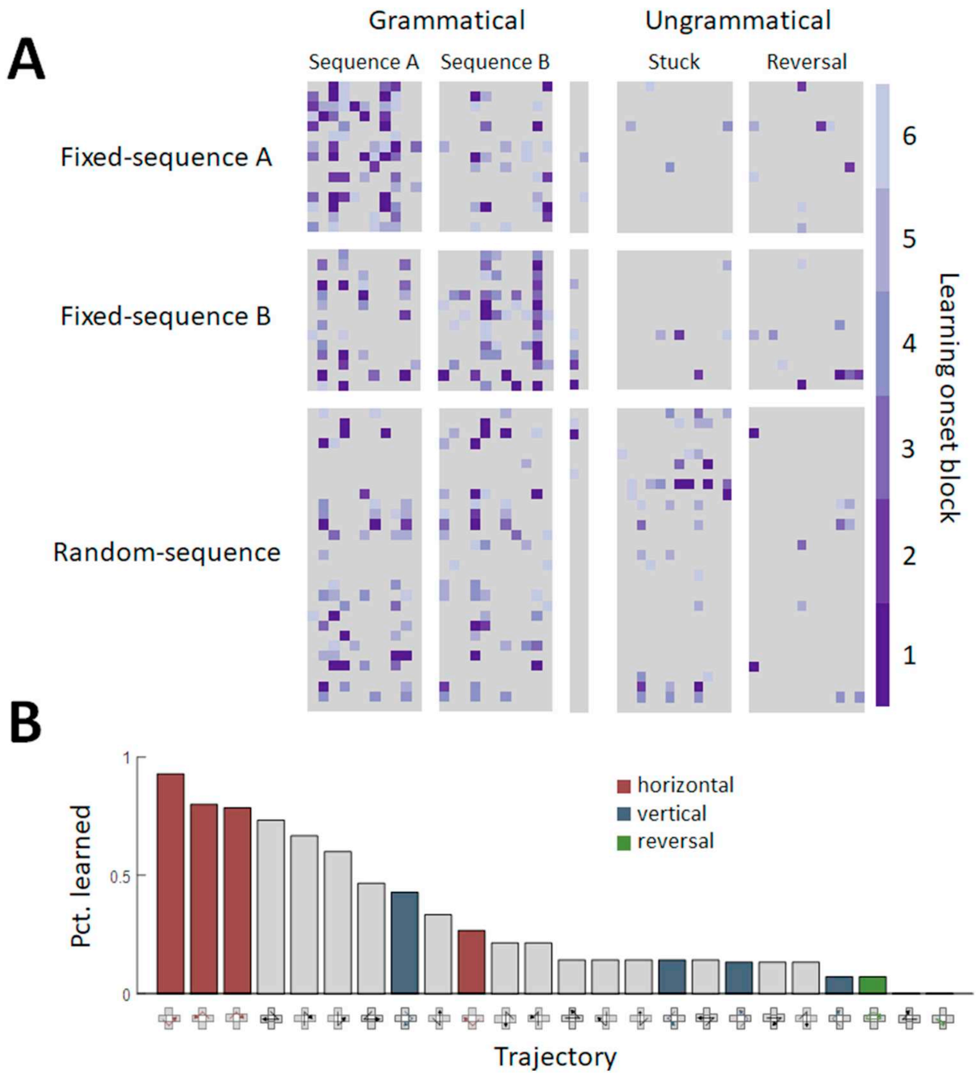

Fig. 4. A) Heat-map of learned responses. Rows represent participants, columns represent different responses, and color saturation represents learning onset. The upper group of rows belong to the fixed-sequence participants that had *sequence A* as their main sequence ($n = 15$). The middle group of rows are fixed-sequence participants that had *sequence B* as their main sequence ($n = 14$). The bottom group of rows are the random-sequence group ($n = 30$). Vertically, columns are divided according to their type. The three leftmost groups of columns are grammatical responses, corresponding to responses of *sequence A* (leftmost), of *sequence B* (middle) and two additional grammatical responses that do not appear in neither sequence. The two rightmost groups of columns are ungrammatical responses, divided into stuck responses (remaining in the location of the last target location) and reversal responses (reverting to the location of the target before last). B) Percentage of fixed-sequence participants that had learned each of the correct responses in their stimuli stream. Each response is given as a prefix-response trajectory and responses are sorted from most to least learned. The three most learned responses produce a natural horizontal "motion" from one side to the other.

responses than ungrammatical ones (4.76 ± 2.56 vs. 1.31 ± 1.76, $t(58) = 7.72, p < 0.0001$). Participants of the fixed-sequence group had learned more grammatical responses than their counterparts (6 ± 2.10 vs. 3.57 ± 2.4, $t(57) = 4.13, p < 0.001$). This is because the fixed-sequence group had learned more grammatical responses belonging to the sequence driving their stimuli stream (either *sequence A* or *sequence B*) than other grammatical responses (3.83 ± 1.69 vs. 2.17 ± 1.34, $t(28) = 4.04, p < 0.001$). Participants of the random-sequence group, on the other hand, had no preference for learning grammatical responses of *sequence A* or of *sequence B* (1.77 ± 1.45 vs. 1.7 ± 1.32, $t(29) = 0.27, p = 0.79$).

An additional difference between the groups is found in their acquisition of ungrammatical responses. While fixed-sequence participants and random-sequence participants learned reversals at the same rate (0.62 ± 0.9 vs. 0.33 ± 0.66, $t(57) = 1.4, p = 0.17$), random-sequence participants acquired more stuck responses than fixed-sequence participants (1.27 ± 1.64 vs. 0.38 ± 0.82, $t(57) = 2.62, p = 0.01$). We believe this difference reflects the fact that stuck responses are not only an anticipation of upcoming targets, but also an essentially different response in which no attempt of anticipation is made. In other work of ours we show that wrong anticipation attempts are more costly in RT than no anticipation attempts (Tal et al., 2020). We believe that waiting for the next target without trying to anticipate its location is a rational strategy when the correct location is yet to be learned. Indeed, results from that work find that correct anticipations increase at the expense of stuck anticipations as the SRT task progresses. This could be a possible explanation of why consistent stuck behaviors are more prevalently found in the random-sequence group than in the fixed-sequence group.

Taken together, these findings highlight the difference between the learning processes that took place in the fixed-sequence task and the random-sequence task. Both groups are afforded a grammar they can learn (via statistical learning), but only the fixed-sequence participants have a particular sequence they can learn (via sequence learning). In the following sections we focus on sequence learning and how that had evolved in the fixed-sequence group throughout the task, by focusing on the responses these participants had learned that correspond to the sequence driving their stimuli stream.

3.2. *Anticipation of learned vs. non-learned items*

Participants of the fixed-sequence group had learned to anticipate 3.86 ± 1.66 out of the twelve items comprising the SRT pattern (32.2%). Learning occurred during the third block on average, (block number $M = 3.6$, $SD = 1.9$).

To test our classification of learning, we compare anticipation due to correctly-learned responses (learned items) to other anticipation occurring in the task, due to non-learned or to incorrectly-learned responses (non-learned items; Fig. 5A). Note that correctly-learned responses do not always yield successful anticipation. They yield

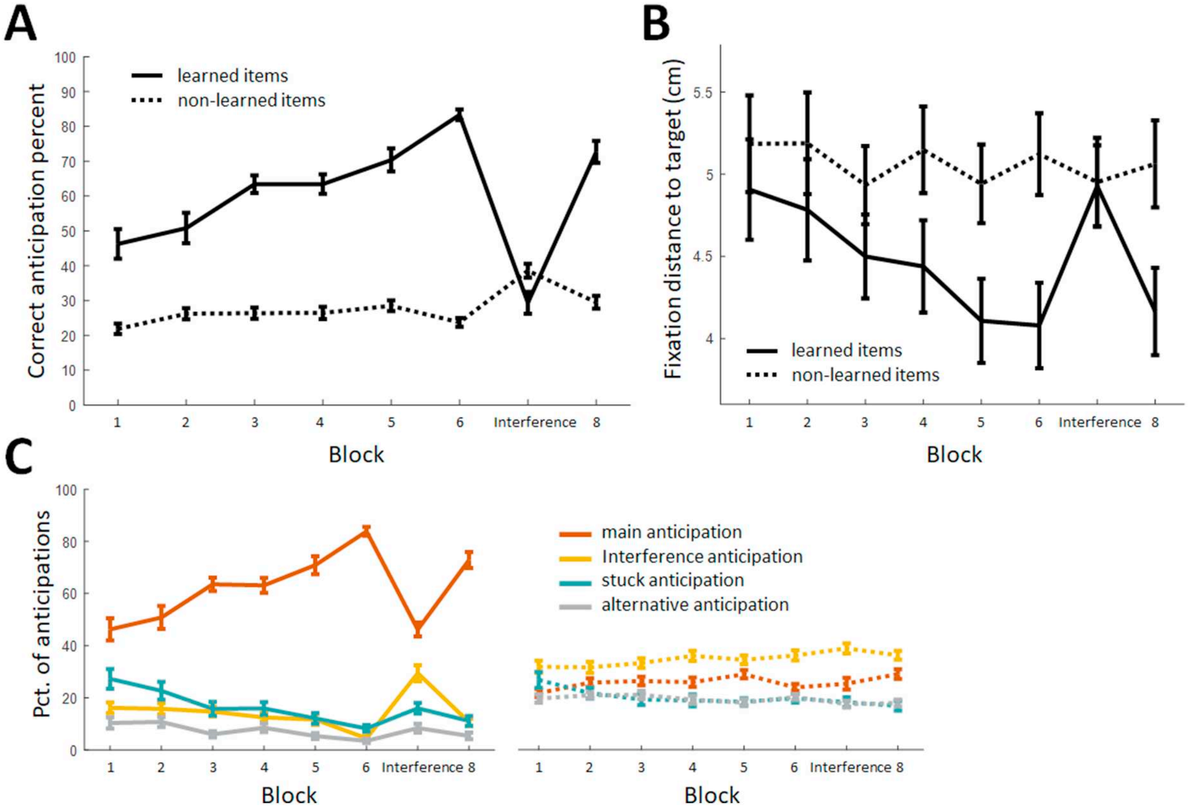

Fig. 5. A) Breakdown of correct anticipations into those belonging to learned items and those that do not. B) Average distance between ISI fixations and target locations, in learned and in non-learned items. C) Anticipatory behavior semantics. Following any two stimuli in the task, a fixation can be made towards four different locations: the location that should follow according to the main sequence ("main anticipation"), the location that should follow according to the interference sequence ("interference anticipation"), the location of the last stimulus ("stuck anticipation") and the remainder fourth location ("alternative anticipation"). This breakdown of anticipatory behavior is plotted in the left panel for learned items and in the right panel for non-learned items. Error bars represent SEM.

incorrect anticipation during the interference block and occasionally during transitions between blocks. Also, note that all anticipation of an item is taken into account, including trials before learning onset, in order to observe how learning evolved from the very beginning of the experiment. So, for example, if a participant had learned the item *2-1-3* during block #4, all ISIs in the experiment that follow the appearance of *2-1* will be considered learned item ISIs, regardless of whether *3* indeed followed them and whether they occurred before block #4.

Anticipations of learned and non-learned items were correct at different rates during the learning phase (main effect of type: $F(1, 28) = 209.81$, $p < 0.001$, $\eta_p^2 = 0.88$). The different rate at experiment onset is partially due to an a-priori bias in gaze that some participants had (Shimojo, Simion, Shimojo, & Scheier, 2003), enabling them correct anticipation of some items from the very beginning of the experiment (see Temporal evolvement of learning for more details). Anticipations developed differently over time (interaction between type and practice: $F(2.9, 82.4) = 18.03$, $p < 0.001$, $\eta_p^2 = 0.39$). Correct anticipation of the subset of learned items rose from 46.2% to 83.3% on average throughout blocks one to six ($F(2.8, 77.4) = 24.63$, $p < 0.001$, $\eta_p^2 = 0.47$). Correct anticipation of non-learned items also rose with practice ($F(5, 140) = 4.25$, $p = 0.001$, $\eta_p^2 = 0.13$), but more moderately, from 21.8% to 23.7% on average.

Sequence interference affected behavior ($F(1, 28) = 108.1$, $p < 0.001$, $\eta_p^2 = 0.79$) differently for learned and non-learned items (main effect of type: $F(1, 28) = 127.8$, $p < 0.001$, $\eta_p^2 = 0.82$; interaction of type and block: $F(1, 28) = 212.9$, $p < 0.001$, $\eta_p^2 = 0.88$), and similar effects were found when the original sequence was restored in the recovery phase. Behavior in non-learned items was modulated by sequence interference ($F(1, 28) = 46.45$, $p < 0.001$, $\eta_p^2 = 0.62$) and recovery ($F(1, 28) = 13.20$, $p = 0.001$, $\eta_p^2 = 0.32$), but, importantly, in the opposite direction of typical SRT findings. Only anticipations of learned items had become less correct when the practiced sequence changed, and recovered when it was restored, indicating learning of the sequence.

Notably, the high rate of correct anticipations that participants reached for learned items indicates that results are not a mere artifact of separating items with coincidental high performance. Surrounding learning onset, participants performed the learned response during 52.8% of ISIs ($SD = 17.7\%$). By the end of the experiment, the learned response was performed during 83.3% of ISIs ($SD = 13\%$). This continued steady incline in performing the learned response since learning onset ($t(111) = -14.59$, $p < 0.0001$) suggests that these items were indeed learned, and that classification is robust to our particular choice of statistical threshold.

Additional indication that anticipation attempts of learned items are different from those of non-learned items can be found in the proximity these fixations have to the upcoming target locations (Fig. 5B). While fixations of non-learned items tended to land at a stable distance of 5.07 ± 0.11 cm from the target throughout the experiment ($F(3.3, 91.1) = 0.67$, $p = 0.58$, $\eta_p^2 = 0.02$), fixations of learned items grew closer to the target over the learning phase ($F(2.9, 80.6) = 4.27$, $p < 0.01$, $\eta_p^2 = 0.13$). Interestingly, fixation distance to target undergoes interference and recovery ($F(1, 28) = 23.41$, $p < 0.001$, $\eta_p^2 = 0.46$ & $F(1, 28) = 18.72$, $p < 0.001$, $\eta_p^2 = 0.40$, respectively). This suggests that when the sequence changes, these fixations no longer reflect knowledge of the target but rather an information seeking behavior, similar to anticipation attempts of non-learned items ($t(28) = -0.18$, $p = 0.86$).

3.3. Semantics of anticipatory behavior

The rate of correct anticipations cannot indicate whether sequence shift was accompanied by any behavioral change, or rather whether participants maintained their previous behavior despite different outcomes. A deeper understanding of oculomotor behavior is afforded by examining anticipated locations with relation to the context in which they occur, instead of merely considering them a correct or incorrect response. We therefore categorized ISI fixations based on the two stimuli that preceded the ISI. Fixations on the location that would follow the previous two targets according to the main sequence were termed "main anticipations", on the location that would come next according to the interference sequence "interference anticipations", on the location that the last target had occupied "stuck anticipation", and on the last fourth location "alternative anticipation" (Tal et al., 2020; Fig. 5C).

Main anticipations reflect learning of the main sequence (sequence learning). Their production grew with practice within both learned and non-learned items, but significantly more within learned items (main effect of block: $F(2.9, 81.7) = 26.36, p < 0.001, \eta_p^2 = 0.49$; main effect of type: $F(1, 28) = 197.27, p < 0.001, \eta_p^2 = 0.88$; interaction effect: $F(2.9, 81.6) = 17.77, p < 0.001, \eta_p^2 = 0.39$). Moreover, production of main anticipations dropped with sequence interference and then recovered with sequence reinstatement only within learned items ($F(1, 28) = 235.00, p < 0.001, \eta_p^2 = 0.89$, $F(1, 28) = 40.17, p < 0.001, \eta_p^2 = 0.59$ for interference and recovery within learned items, respectively; $F(1, 28) = 0.31, p = 0.58, \eta_p^2 = 0.01$, $F(1, 28) = 2.08, p = 0.16, \eta_p^2 = 0.07$, for interference and recovery within non-learned items, respectively). This means that the signals indicating sequence learning in SRT were almost entirely captured within the subset of ISIs (32.2%) classified as belonging to correctly learned items. Importantly, classification depended on data from the learning phase only, and so interference and recovery effects (from blocks seven and eight) are independent of the training data and provide sound validation of the classification.

Interestingly, during sequence interference, main anticipations of learned items were not replaced by other responses equally, but were predominately replaced with the new correct interference anticipation (interaction between blocks six and seven and behavior type, considering interference, stuck and alternative types only: $F(1.4, 40.5) = 16.04, p < 0.001, \eta_p^2 = 0.36$). This is an additional indication of robust statistical learning achieved by this point of the experiment, as the interference sequence is the only viable option according to the task grammar once the main anticipation has been violated. We believe this is not yet learning of the new interference sequence (and thank both anonymous reviewers for their helpful insight on this matter), as fixation distance to target seems to indicate search during this block and not knowledge (Fig. 5B). Nevertheless, it is remarkable that this behavioral adaptation was exclusive to correctly learned items (Fig. 5C, left panel), as behavior following non-learned items was indifferent to sequence changes (main effect of blocks six and seven considering interference, stuck and alternative types only: $F(1, 28) = 0.31, p = 0.58, \eta_p^2 = 0.01$; interaction between blocks and behavior type: $F(1.2, 34.7) = 1.83, p = 0.19, \eta_p^2 = 0.06$; Fig. 5C, right panel). This means that stimuli pairs that had a correct learned response associated with them were nevertheless more amenable to new learning when the sequence changed than stimuli pairs that were not learned to begin with.

3.4. RT to learned vs. non-learned items

The ability to tease apart trials corresponding to learned items from those corresponding to non-learned items enables a re-examination of the standard RT findings of SRT in finer resolution. A breakdown of RT according to the type of trial it occurred in, as detected by our eye-tracking method, is offered in Fig. 6.

RT declines with practice, regardless of whether it occurred in learned or non-learned items ($F(2.3, 65) = 23.05, p < 0.001, \eta_p^2 = 0.45$). However, RTs differed between both types ($F(1, 28) = 46.15, p < 0.001, \eta_p^2 = 0.62$). Specifically, there is a difference in the effect practice had on them (interaction between block and trial type: $F(5, 140) = 6.84, p < 0.001, \eta_p^2 = 0.20$), as RT of learned items dropped at a steeper rate as the experiment progressed. This difference can be taken to represent the benefit that sequence learning grants quickness of response. To the extent that the model used in this work captures all learning in the task, the decline in RT of non-learned items reflects all improvement in the task that is not a result of sequence learning (Tal et al., 2020; Schwarb & Schumacher, 2012).

Corroborating this interpretation of the RT curves, RT of learned and non-learned items is equivalent during the interference block, ($t(28) = 1.45, p = 0.16$). Also, an interference effect in block seven is evident only in RT of learned items ($t(28) = -8.28, p < 0.001$), and not in RT of non-learned items ($t(28) = -0.99, p = 0.33$). However, see Sequence learning distilled for a possible account of the small interference trend existing also in non-learned RT. Therefore, the difference between the two curves enables an approximation of the distilled effect sequence learning has on RT (Fig. 6B). After three blocks of practice already, a steady benefit of roughly 70 ms is reached.

3.5. Temporal evolvement of learning

Our analysis allows examining what has been learned and when on an individual and item-based resolution. A complete heat-map of these results is given in Fig. 4A.

22.3% of sequence items were learned during the first block, 24.1% during the sixth, and the remainder 53.6% during the blocks in between quite uniformly ($M = 13.4\%$, $SD = 2.8\%$ per block). While our method may slightly bias late learning, as binomial significance is more easily met in smaller windows than in larger ones, the high rate of learning during the first block could not be explained as an artifact of the method. Rather, it indicates either early acquisition of items in the sequence, or a predisposition (a-priori to the experiment) to perform the correct response in those items. We find evidence of both. In fourteen cases, items were identified as learned from their very first occurrence. This means that performance was above chance in those items throughout the entire experiment. It is hard to consider these items learned, therefore, but rather maintained, as correct anticipation of them was high from the very beginning ($M = 86.5\%$, $SD = 13.2\%$), partially accounting for the difference between learned and non-learned item anticipation from the very onset of the experiment. When disregarding these fourteen items, learning onset of learned items occurred at the 30.6 ± 15.0 encounter with them, and the first learning each participant had made occurred after 18.4 ± 14.1 encounters with the item.

With regards to the chronological evolvement of learning the sequence, we find no indication of concatenation. That is, we find no tendency to learn responses adjacent to already learned ones. Instead, participants appear to have learned items within the sequence in a fragmented manner. This is particularly pronounced in early stages of learning. Only one participant in our study had learned a second item concatenated to the first ($p < 0.05$ under a binomial test, given the chance probability of learning an adjacent item is $\frac{2}{11}$). Overall, we find that the extent of concatenation throughout the task does not exceed what would be expected by chance (Supplementary Fig. 2).

3.6. Individual-based and item-based differences

As mentioned before, participants have learned 3.86 ± 1.66 of the twelve items presented to them ($M = 32.2\%$, $SD = 13.9\%$ of the sequence). However, substantial individual differences are found. Performance ranged from two participants that have learned only one item, to one participant that has learned eight. Individual differences are known to exist in statistical learning tasks and may underlie these differences as well (see Clinical potential).

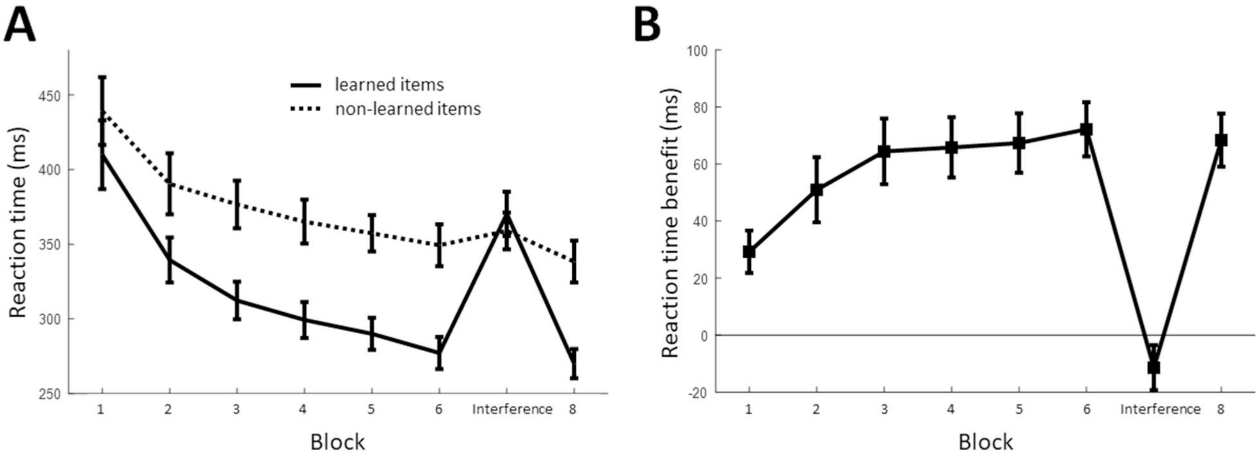

Fig. 6. A) RT to learned and to non-learned items separately. RT improves at a steeper rate following correctly learned in comparison with other stimuli pairs. Also, interference and recovery effects driven by sequence shift (blocks 6 vs. 7, and 7 vs. 8 respectively) exist only within responses following correctly learned stimuli pairs. Sequence interference brings RT to learned items to the speed of RT to non-learned items, while the latter are almost not affected by sequence shift. This suggests that RT improvement in non-learned items is not a product of sequence learning. B) Distilled sequence learning effect. The difference between RT to learned and to non-learned items (RT benefit) isolates the contribution of sequence learning to RT. Error bars represent SEM.

In addition to differences across participants, a striking difference is found between items. In both groups, certain items were learned by the vast majority of participants, while others were learned by almost or entirely none. An aggregated histogram can be seen in Fig. 4B. The three most learned items consist of a horizontal trajectory of targets, from one side to the other via the top or bottom positions. As had been reported in Reed and Johnson (1994) chunks within the sequence that correspond to movements that are likely to be familiar to participants are more salient, and may be learned more quickly. Indeed, horizontal trajectories had a 69.5% chance of being learned, while other items a mere 24.6% ($t(22) = 3.57$, $p < 0.01$). Vertical trajectories did not, however, muster a similar effect ($t(22) = -0.98, p = 0.34$). Lastly, two reversals (items in which the third location is the same as the first) inhabited the fixed-sequence stimuli stream, one in each sequence. Only one participant learned a reversal item of the sequence. See Chunking or statistical learning section for further discussion on the effects item structure had on “learnability”.

4. Discussion

Eye tracking during an SRT task provided an ongoing indication of anticipation towards upcoming targets in sequentially ordered stimuli. Anticipation reflected an information gathering process at early stages of learning, and then reflected knowledge when learning has been accomplished. Probabilistic analysis of this anticipation at an item-based resolution uncovered what parts of the sequence each participant had learned, and when that learning had taken place. According to this analysis, participants that had a fixed sequential order in stimuli had converged onto more consistent gaze responses than participants that could rely only on learning the task statistical grammar. The responses they have converged to corresponded to the sequence driving their stimuli stream. Participants who were exposed to a fixed sequence have learned, on average, a little under a third of the regularities making that sequence. Also, learning of the sequence did not progress via clustering. These findings suggest that the typical sequence learning effect found in SRT is due to partial and fragmented learning of the target sequence. In addition, the actual benefit of sequence learning to RT can be isolated from other factors influencing it for the first time, and is found to asymptote quite early in the experiment.

4.1. What is learned of the task sequence?

The analysis carried out in this work examined learning dependent upon the minimally sufficient sub-sequence that can support it, which in the current paradigm is two preceding stimuli. Learning of stimulus transition probabilities of the first and second order (i.e. dependent upon one or two previous stimuli) had already been demonstrated in probabilistic versions of SRT, in which these probabilities were the only rule governing the stimuli (Cleeremans & McClelland, 1991; Hunt & Aslin, 2001). It had also been found to arise even when there was no order governing the stimuli (Maheu, Dehaene, & Meyniel, 2019). However, in deterministic SRT as in the fixed-sequence group, learning can rely on longer regularities in the data. Our results show that participants nevertheless rely on second order transitional probabilities for learning, which is the most efficient option available to them.

The strongest evidence suggesting that second order transition probabilities are learned comes from the interference phase, when the sequence is suddenly replaced. Pairs of stimuli during this phase predict to a high degree that participants carry out the response that fits the main sequence. Because pairs are the largest segment shared between the main and the interference sequence, what participants rely on to make these responses cannot be more than two preceding stimuli. It should be noted that this does not rule out the possibility that other transitional probabilities were learned as well. We find evidence of first order learning, but to a much smaller degree than second order, and higher order transitional probabilities may have also been acquired but that is impossible to detect in an SOC sequence. Also, statistics other than transition probabilities are extracted and learned during the task (Tal et al., 2020; Maheu et al., 2019; Reed & Johnson, 1994). However, the current analysis reveals that learning derived from second order transitional probabilities accounts for the lion's share of the sequence learning effect observed in SRT.

4.2. Learning speed

Learning in this task can evolve quite quickly. Two presentations of sequence items were sufficient for participants to begin responding to some with above chance precision, mirroring similar findings in brain imaging (Turk-Browne, Scholl, Chun, & Johnson, 2009).

This speed allows participants to adapt to a sequence and acquire it to a certain extent within one block already. Hence, during the interference phase, participants not only abandon the practiced sequence, but manage to respond correctly to the new sequence as well. Little attention has been given in the literature to how participants process the interference sequence. Our results suggest that replacement of the old sequence is accompanied with high correct response to the new one, and so a portion of the effect sought after in RT is counteracted. The RT interference effect therefore underestimates the actual contribution that learning of the main sequence had in RT.

We were surprised to find that the swift adaptation to the new sequence during the interference phase was completely contained within the same subset of items that were learned during the learning phase. Responses that had been associated with certain stimuli pairs were replaced with new responses, while responses to other stimuli pairs remained quite indifferent to the change in sequence even though they were "available" for a response to be associated with them. We believe that the high associability of learned items exists in block seven because learning is still in its initial stages. Wrong anticipation attempts induce an error that facilitates error-driven learning, while non-learned items have no such drive. However, we expect that longer practice would cause entrenchment of the learned regularities, resulting in reduced sensitivity to change – an intriguing pursuit for future study.

4.3. Fragmented learning

A novel outcome of our work, which had not been examined so far to the best of our knowledge, is the order in which items in the sequence are learned. Sequence learning is believed to be the process facilitating skill acquisition. A major characteristic of skills, such as playing a musical piece or performing a somersault, is that producing a single element from within them (positioning hands in a certain way during a musical piece or bending the torso with a certain direction and strength during a somersault) requires reproducing several, and sometimes even all, preceding elements. Therefore, it may be appealing to imagine that sequence learning progresses via incremental concatenation, forming a larger and larger inter-dependent learned pattern. However, participants in our study learned the sequence in a fragmented manner. Items disjoint from one another were picked up as practice progressed, with no apparent preference for neighbors.

Also, participants in our task had learned only a handful of the items comprising the sequence, even though each was presented to them fifty-four times during the learning phase. This finding supports the indications found in Moisello et al. (2009) that only partial knowledge of the sequence order is acquired during SRT. It also fits implicit learning literature, in which it was shown that only a fraction of learnable items are learned in any given experiment (Schlagbauer, Muller, Zehetleitner, & Geyer, 2012; Smyth & Shanks, 2008).

Taken together, our finding suggests that learning during the SRT task evolves in a quick and fragmented manner. Second order transition probabilities are learned in isolation, not unlike what we would expect if items were independently located in a stimuli stream in random order. It is possible, however, that only initial stages of learning during deterministic SRT are fragmented. Salient items are picked up very early on, but then knowledge could be chunked together if sufficient practice is provided (see next section). In either case, this finding, while less intuitive, abides well with our understanding of attentional processing. Attention is not homogenously distributed across time, as is most clearly demonstrated in the attentional blink phenomenon (Shapiro, Arnell, & Raymond, 1997). Allocation of attention at certain moments would facilitate learning, but would also predict no learning within several hundreds of milliseconds following these moments. When exposed to a novel sequence of stimuli, all items are unfamiliar and require attention to be learned. It is therefore reasonable that preliminary stages of learning a sequence will be characterized by isolated learned items, scattered across the sequence.

4.4. Chunking or statistical learning?

Our findings reveal that learning did not distribute evenly across all items within the SRT task. Upon examination, one structure tying targets and the two stimuli preceding them quickly stands out as meaningful in predicting learnability. Targets that were learned the most were those that formed, with the two stimuli preceding them, a triplet beginning on one side, finishing at the other, and appearing at the top or bottom position in between. These trajectories are special in three ways that may have facilitated learning.

Firstly, trajectories as described above evoke perception of horizontal motion. Due to this motion illusion, participants have an a-priori disposition to correctly anticipate their third location given the first two (e.g. a target appearing first on the left and then on the top may naturally trigger orientation of attention to the right). Thus, search time for the correct response following this stimuli pair will be shortened. This point underscores again that participants are not "tabula rasa" when entering an experiment, and it is important to take into account a-priori tendencies regarding natural stimuli when measuring performance in laboratory settings (Siegelman, Bogaerts, Elazar, Arciuli, & Frost, 2018). Secondly, we expect that horizontal trajectories stand out in the stimuli stream, increasing the probability they will in fact be learned. As found in Koch and Hoffmann (2000) and in Kirsch et al. (2010), salient relational structures are exploited by participants during the SRT task to form chunks of predictable responses. The horizontal trajectories embedded in our task are likely to form such perceptual structures and therefore get chunked together. Lastly, and not unrelated, horizontal trajectories should be more easily encoded in memory than other "meaningless" trajectories.

These findings support the case of chunk learning (CL) in SRT over statistical learning (SL; Du & Clark, 2017; Jiménez, 2008). However, two major caveats to this account can be raised. Firstly, vertical trajectories do not evoke the same effect that horizontal ones do. It seems reasonable that this is due to an asymmetry in visual processing, namely the Horizontal-Vertical Anisotropy (Corbett & Carrasco, 2011; Feng, Jiang, & He, 2007). Performance in visual tasks is better along the horizontal axis compared to the vertical one, supposedly granting prevalence to horizontal movements over vertical ones in our experiment. Also, this predisposition could be enhanced by reading patterns (Feng et al., 2007). This study was conducted in Israel and therefore participants' mother tongue is typically read horizontally (Hebrew, Arabic or English). It could be that reproducing the study with participants with a vertically read language (e.g. Japanese) would yield different results. Proponents of the CL account can build on both these hypotheses to claim that dominance of the horizontal over vertical trajectories would cause the horizontal to be acquired first. Then, triplets that have overlap with horizontal trajectories will in fact be more difficult to learn due to the former's encapsulation into a single indivisible unit. Vertical trajectories all overlapped with horizontal ones in this study, so the fact that they were learned so poorly can be supporting, not conflicting evidence, for the CL account.

The second and more problematic caveat for the CL account, is the fact that two back-and-forth patterns that exist in the current task (e.g. bottom-left-bottom), elicited extremely poor learning. While it is reasonable to expect that such reversals would constitute salient relational structures and promote learning like motion trajectories do (Reed & Johnson, 1994), in the current task they seem to have brought about the opposite effect. In our opinion, reversals were particularly difficult to learn because they are "ungrammatical". Participants had learned that three consecutive targets appear in different locations of one another, and therefore orientation of gaze towards any of the last two locations will lead to a mistake. Knowledge derived from SL prevented knowledge acquisition via CL.

A major critique against CL is the finding that some chunks are apparent from the very onset of the experiment, and hence do not represent learning (Du & Clark, 2017; Song & Cohen, 2014). The current study may account for this. While some learning does appear to rely on predisposed behaviors and not acquired ones, most learned items are acquired within the task. Because we find that a few repetitions are sufficient for learning, it is reasonable that chunks appear from the very first block. Moreover, even though chunks may not grow in time (i.e. they do not form longer learned sequences via concatenation), response to them does improve with practice, indicating that they are not a mere pre-disposition but are continuously learned in the task.

To sum, it will be hard to explain our results as driven solely by either CL or SL. Rather, they seem to point to mixed contribution of the two mechanisms (Kirsch et al., 2010). Our results suggest that simple surface statistics are picked up very early in the task along with salient chunks in the stimuli stream. With practice, both mechanisms contribute to acquisition of additional items in the stream, but not necessarily those adjacent to former chunks. Concatenation may be absent in our results simply due to insufficient practice, and so additional study of this task with extended practice is needed.

A methodological conclusion may be drawn from these findings. Future studies wishing to examine learning of homogenous neutral stimuli would be wise to control for ease of learning. Just as first and second order statistics are controlled for, sequences should be composed in such a way that structures that are a-priori beneficial, either in perception or in response, are controlled for as much as possible.

4.5. Sequence learning distilled

The analysis undertaken in this work is one of separating signal from noise. It reveals that the sequence learning signal sought after in SRT research is larger than was previously thought. When considering only correctly learned items, oculomotor performance rises and drops by 37.1% and 54% following sequence practice and interference, respectively. For comparison, the original effect reported in Vakil et al. (2017) rose by 12.7% and fell by 6% only. In the domain of RT, in which sequence learning is typically measured, our analysis shows that practice improved RT of learned items by 132.8 ms on average and subsequent interference made it 93.2 ms slower. In comparison, the same effects were of size 115.5 ms and 50.3 ms respectively when averaged across all trials.

The “noise” in the SRT task is no less valuable than the signal. Non-learned items enable for the first time to measure RT to stimuli that is predictable but nevertheless not learned. To asses learning of regularities, studies typically use one of two control conditions. Performance in practiced regular items is compared either to performance in unpracticed regular items presented late in the experiment (as with the interference phase in the current SRT paradigm), or to performance in random items presented alongside the regular ones (e.g. Cleeremans & McClelland, 1991). The former has the disadvantage of not comparing similar time bins, and the latter has the disadvantage of comparing items of a different nature. A control that is both regular and online was thus far unavailable. Non-learned items provide this control because they are of the exact same nature of the learned items and are encountered under the same levels of motor expertise and fatigue. They also do not interfere with learning by signaling that the environment is not regular or had changed, as random stimuli may. Moreover, non-learned items constitute a control measurement that is tailored per individual, and so does not rely on a-priori assumptions of what will be learned and what not. This enables us to estimate that the true impact of sequence learning on RT in our task is of roughly 70 ms. It is achieved quite early in practice, and then maintained until the end.

Modelling both learned and non-learned RT curves as a declining power law function (Logan, 1992) suggests that in our task, given additional practice, both types of responses would asymptote at roughly the same time, making only negligible improvement from the eighteenth block onward (<1 ms). Learning benefit would eventually stabilize at a little over 80 ms. These analyses, in addition to their theoretical value, can be instrumental in future studies aiming to isolate processes underlying SRT in brain or behavior.

Lastly, classification of learning in this work was based on proactive oculomotor behavior. Although it generalized well into the domain of reactive motor response, RT revealed at least one imperfection in the oculomotor-based classification, in a slight interference effect in non-learned items. It is not clear whether this is due to sequence learning that had escaped our analysis, or rather to an overall slowing down of RT resulting from the sudden change in the sequence governing the task (Tal et al., 2020). The former case, however, may indicate learning that was driven by covert shifts of attention (Posner & Petersen, 1990). Such learning would not be detectable in the oculomotor signal and would therefore be missed by our analysis method.

4.6. Clinical potential

An intriguing issue to study, following this work, is the consistency of individual performance. If the capacity of learning in SRT proves a reliable individual trait, i.e. if performance at one time predicts how well the same individual will do another time, this analysis could form a basis for a fine-resolution metric of sequence learning capability. Such individual reliability had been found in several statistical learning tasks, including probabilistic variants of SRT (Siegelman & Frost, 2015). Surprisingly, it was not found in standard deterministic SRT, even though other implicit motor learning tasks have proven stable (Stark-Inbar, Raza, Taylor, & Ivry, 2017). As Stark-Inbar and colleagues have suggested, that null finding could have been due to the inherent noisiness of RT. The method introduced in this work may be sensitive enough to capture individual capacity of learning and prove reliable. An individual score capturing fine-grained aspects of the sequence learning process could help uncover what processes underlie SRT, inform on which are dissociable from one another, and tie these processes to individual cognitive skills and brain structures.

A sensitive learning metric could also prove important on a group-level analysis, especially regarding clinical populations. The SRT task is modulated by numerous patient groups, such as Alzheimer and Parkinson's disease (Ferraro, Balota, & Connor, 1993; Jackson, Jackson, Harrison, Henderson, & Kennard, 1995), damage to basal ganglia (Vakil, Kahan, Huberman, & Osimani, 2000), traumatic brain injury (Vakil, Kraus, Bor, & Groswasser, 2002), developmental dyslexia (Gabay, Schiff, & Vakil, 2012), children with damage to the cerebellum (Berger et al., 2005), spinal cord injury (Bloch, Tamir, Vakil, & Zeilig, 2016) and more. The typical finding in all these cases is impaired learning expressed either by a moderate learning curve (blocks 1–6) or a modest cost with sequence shift (blocks 6–7). There is no differentiation, therefore, between all these pathological groups because there is only the single measure of RT to evaluate performance. The in-depth analysis offered in this and previous work of ours can transform SRT into a more sensitive diagnostical tool, as it provides several measures in addition to RT: correct oculomotor anticipations, amount of oculomotor anticipation attempts, number of items learned, learning rate, etc. Hopefully, these measures would allow to create a unique profile of each of the pathological groups, by characterizing their unique performance profile. Characterizing individual and groupwise differences will improve dissociation of the cognitive and neural mechanisms underlying learning, and thus advance our understanding of learning difficulties in clinical and in normal populations.

CRediT authorship contribution statement

Amir Tal: Conceptualization, Methodology, Investigation, Software, Formal analysis, Visualization, Writing - original draft. **Eli Vakil:** Conceptualization, Investigation, Supervision, Project administration, Funding acquisition, Writing - review & editing.

Acknowledgements

This work was supported by the Farber Alzheimer's Center Foundation (# 259147). We thank Moshe Bar and Liad Mudrik for very helpful comments on earlier versions of this manuscript, and Simone Schwizer Ashkenazi for data collection of the random-sequence group.

Appendix A. Supplementary figures

Supplementary figures of this article can be found online at https://doi.org/10.1016/j.cognition.2020.104291.

Appendix B. Supplementary data

The data used for this work are available at https://osf.io/rj692/.

References

Berger, A., Sadeh, M., Tzur, G., Shuper, I. A., Kornreich, L., Inbar, D., ... Vakil, E. (2005). Motor and non-motor sequence learning in children and adolescents with cerebellar damage. *Journal of the International Neuropsychological Society, 11*, 482–487.

Bloch, A., Tamir, D., Vakil, E., & Zeilig, G. (2016). Specific deficit in implicit motor sequence learning following spinal cord injury. *PLoS One, 11*(6), 1–13.

Cleeremans, A., & McClelland, J. L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General, 120*(3), 235–253.

Corbett, J. E., & Carrasco, M. (2011). Visual performance fields: Frames of reference. *PLoS One, 6*(9), 1–10.

Destrebecqz, A., & Cleeremans, A. (2001). Can sequence learning be implicit? New evidence with the process dissociation procedure. *Psychonomic Bulletin & Review, 8*(2), 343–350.

Du, Y., & Clark, J. E. (2017). New insights into statistical learning and chunk learning in implicit sequence acquisition. *Psychonomic Bulletin and Review, 24*(4), 1225–1233.

Feng, C., Jiang, Y., & He, S. (2007). Horizontal and vertical asymmetry in visual spatial crowding effects. *Journal of Vision, 7*(2), 1–10.

Ferraro, F. R., Balota, D. A., & Connor, L. T. (1993). Implicit memory and the formation of new associations in nondemented parkinson's disease individuals and individuals with senile dementia of the alzheimer type: A serial reaction time (SRT) investigation. *Brain and Cognition, 21*, 163–180.

Gabay, Y., Schiff, R., & Vakil, E. (2012). Attentional requirements during acquisition and consolidation of a skill in normal readers and developmental dyslexics. *Neuropsychology, 26*(6), 744–757.

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*(4), 658–680.

Jackson, G. M., Jackson, S. R., Harrison, J., Henderson, L., & Kennard, C. (1995). Serial reaction time learning and Parkinson's disease: Evidence for a procedural learning deficit. *Neuropsychologia, 33*(5), 577–593.

Jiménez, L. (2008). Taking patterns for chunks: Is there any evidence of chunk learning in continuous serial reaction-time tasks? *Psychological Research, 72*(4), 387–396.

Kirsch, W., Sebald, A., & Hoffmann, J. (2010). RT patterns and chunks in SRT tasks: A reply to Jiménez (2008). *Psychological Research, 74*(3), 352–358.

Koch, I., & Hoffmann, J. (2000). Patterns, chunks, and hierarchies in serial reaction-time tasks. *Psychological Research, 63*(1), 22–35.

Lashley, K. S. (1951). The problem of serial order in behavior. In L. A. Jeffress (Ed.). *Cerebral mechanisms in behavior: The Hixon symposium* (pp. 112–136). .

Logan, G. D. (1992). Shapes of reaction-time distributions and shapes of learning curves: A test of the instance theory of automaticity. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 18*(5), 883–914.

Maheu, M., Dehaene, S., & Meyniel, F. (2019). Brain signatures of a multiscale process of sequence learning in humans. *ELife, 8*, 1–24.

Marcus, D. J., Karatekin, C., & Markiewicz, S. (2006). Oculomotor evidence of sequence learning on the serial reaction time task. *Memory & Cognition, 34*(2), 420–432.

Moisello, C., Crupi, D., Tunik, E., Quartarone, A., Bove, M., Tononi, G., & Ghilardi, M. F. (2009). The serial reaction time task revisited: A study on motor sequence learning with an arm-reaching task. *Experimental Brain Research, 194*(1), 143–155.

Nissen, M. J., & Bullemer, P. (1987). Attentional requirements of learning: Evidence from performance measures. *Cognitive Psychology, 19*(1), 1–32.

Posner, M. I., & Petersen, S. E. (1990). The attention system of the human brain. *Annual Review of Neuroscience, 13*(1), 25–42.

Reed, J., & Johnson, P. (1994). Assessing implicit learning with indirect tests: Determining what is learned about sequence structure. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 20*(3), 585–594.

Schlagbauer, B., Muller, H. J., Zehetleitner, M., & Geyer, T. (2012). Awareness in contextual cueing of visual search as measured with concurrent access- and phenomenal-consciousness tasks. *Journal of Vision, 12*(11), 1–12.

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*(2), 165–178.

Shapiro, K., Arnell, K., & Raymond, J. (1997). The attentional blink. *Trends in Cognitive Sciences, 1*(8), 291–296.

Shimojo, S., Simion, C., Shimojo, E., & Scheier, C. (2003). Gaze bias both reflects and influences preference. *Nature Neuroscience, 6*(12), 1317–1322.

Siegelman, N., Bogaerts, L., Elazar, A., Arciuli, J., & Frost, R. (2018). Linguistic entrenchment: Prior knowledge impacts statistical learning performance. *Cognition, 177*, 198–213.

Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language, 81*, 105–120.

Smyth, A. C., & Shanks, D. R. (2008). Awareness in contextual cuing with extended and concurrent explicit tests. *Memory and Cognition, 36*(2), 403–415.

Song, S., & Cohen, L. (2014). Impact of conscious intent on chunking during motor learning. *Learning & Memory, 21*(9), 449–451.

Stark-Inbar, A., Raza, M., Taylor, J. A., & Ivry, R. B. (2017). Individual differences in implicit motor learning: Task specificity in sensorimotor adaptation and sequence learning. *Journal of Neurophysiology, 117*(1), 412–428.

Tal, A., Bloch, A., Cohen-Dallal, H., Aviv, O., Ashkenazi, S., Bar, M., & Vakil, E. (2020). *Oculomotor anticipation reveals a hierarchical nature of sequence learning.* Submitted for publication.

Turk-Browne, N. B., Scholl, B. J., Chun, M. M., & Johnson, M. K. (2009). Neural evidence of statistical learning: Efficient detection of visual regularities without awareness. *Journal of Cognitive Neuroscience, 21*(10), 1934–1945.

Vakil, E., Bloch, A., & Cohen, H. (2017). Anticipation measures of sequence learning: Manual versus oculomotor versions of the serial reaction time task. *Quarterly Journal of Experimental Psychology, 70*(3), 579–589.

Vakil, E., Kahan, S., Huberman, M., & Osimani, A. (2000). Motor and non-motor sequence learning in patients with basal ganglia lesions: The case of serial reaction time (SRT). *Neuropsychologia, 38*(1), 1–10.

Vakil, E., Kraus, A., Bor, B., & Groswasser, Z. (2002). Impaired skill learning in patients with severe closed-head injury as demonstrated by the serial reaction time (SRT) task. *Brain and Cognition, 50*(2), 304–315.