

How a-priori biases affect sequence learning in a serial reaction time task

Amir Tal (at3405@columbia.edu)

Department of Psychology, Columbia University,
1190 Amsterdam Ave, New York, NY 10027 USA

Eli Vakil (Eli.Vakil@biu.ac.il)

Gonda Multidisciplinary Brain Research Center, Bar Ilan University,
Ramat Gan, 52900 Israel

Abstract

The ability to chain together sequences of information and action is pivotal to everyday acquisition of skills. Despite extensive research of sequence learning, little focus has been given to individual performance in standard tasks measuring this capability. As a result, little is known regarding what knowledge participants gain during such tasks. In the current work, an individual- and item-based analysis is performed of eye movements that occur during a spatial sequence learning task and reflect anticipation of upcoming target locations. We show that the knowledge participants acquire during the task is tightly linked to a-priori response biases they bring into the experiment. Results suggest that a-priori biases may be a sizeable influence on performance in learning experiments, that tends to be overlooked. Implications for designing and reading studies of sequence learning are discussed.

Keywords: sequence learning; SRT; biases; prior knowledge; individual differences

Introduction

Skillful behavior entails producing chains of discrete actions as a single unified procedure. Underlying the acquisition of skill is therefore an ability to sequence incoming information and ongoing action. This ability is termed sequence learning (Clegg, Digirolamo, & Keele, 1998; Lashley, 1951). In the lab, it is predominately studied using the serial reaction time (SRT) task (Nissen & Bullemer, 1987).

The typical finding of SRT is that participants grow quicker in responding to sequences they were previously exposed to compared to responding to unfamiliar sequences. Enhanced speed is taken to indicate improved anticipation of sequence elements ahead of time, even though knowledge of the sequence is not necessarily explicitly available to participants by the end of the task (Cleeremans & McClelland, 1991; Marcus, Karatekin, & Markiewicz, 2006). Due to the robustness and relative simplicity of the task, a huge literature has come to surround SRT. However, not much is known regarding the precise knowledge that participants gain during SRT. The reason for this is that most studies examine groupwise effects averaged across the entire target sequence. Higher-resolution knowledge of what portion of the sequence is learned by participants during the task and when it is learned would shed light on open questions in the field, such as what learning mechanism underlies sequence learning (Kirsch, Sebald, & Hoffmann, 2010). Such understanding would also be instrumental for

fine-grained assessment of individual learning capabilities, a feature that is particularly important in SRT due to its extensive use in the study of clinical populations (Ferraro, Balota, & Connor, 1993; Green, Kern, Williams, McGurk, & Kee, 1997).

Recently, a measure based on eye movements has been shown to capture the major learning effects of SRT similarly to reaction time (RT; Marcus et al., 2006; Vakil, Bloch, & Cohen, 2017). Locations that participants fixate their gaze at before sequence elements appear on screen seem to capture their anticipation of upcoming elements, and hence learning of the sequence. This novel metric of learning reveals mistakes that participants make in their anticipation, information that is unavailable in the standard RT measure. It is therefore possible using this metric to identify both correct and incorrect behavioral patterns during the task. In the current work we set out to perform such an analysis of the oculomotor anticipation signal during an SRT task, at the level of the individual participant. In particular, we aimed to examine what behavioral patterns participants have at the beginning and at the end of SRT, and how they relate to one another. Patterns of consistent behaviors at the end of the experiment supposedly reflect knowledge gained throughout the task and so should mirror the to-be-learned sequence that participants are exposed to. However, patterns existing at experiment onset may reflect predispositions held by participants a-priori to the study.

The topic of prior knowledge brought into the experiment by participants has long been acknowledged in the field of psychology (Harlow, 1949), but is mostly overlooked. An implicit assumption underlying most studies is that participants enter an experiment tabula rasa with regards to the experimental material. However, it has been shown that such an assumption is not always true. In an auditory statistical learning task, for example, participants' mother tongue had influenced whether experimental non-words were successfully acquired or not, depending on whether they aligned with phonetic transitions that are prevalent in the language (Siegelman, Bogaerts, Elazar, Arciuli, & Frost, 2018). Even short exposure to conflicting associations in an artificial language (Gebhart, Aslin, & Newport, 2009) or simply to the cue without the response in a contextual cueing task (Jungé, Scholl, & Chun, 2007) had hampered subsequent acquisition of such statistics.

In the current work we aimed to uncover biases that participants bring into a simple spatial target detection task

and examine how these relate to the eventual learning of a sequence that inhabits the task. To do so, we performed a probabilistic analysis of oculomotor anticipations participants had in a spatial SRT task. To enhance the power of our results we pooled together for this work participants from two different experiments: one standard SRT experiment and another SRT experiment of longer training. We analyze the first six blocks only, the length of the standard SRT, and so this data is equivalent in both studies.

In order to examine the homogeneity of biases found in the task, and to corroborate that response patterns at experiment onset do not reflect rapid learning occurring within the task but rather biases existing a-priori to it, three different stimuli streams were examined in this work. Participants of the fixed-sequence group were presented stimuli governed by a fixed repeating sequential order, as in standard SRT. There were two such sequences, and thus two subgroups with different stimuli streams. Participants in the random-sequence group were presented stimuli governed by the same statistical constraints as in the fixed-sequence group, but no fixed sequence inhabited the stream (see *Random-sequence group* section in *Methods*).

Methods

Participants

Sixty-eight undergraduate students (40 females, mean age = 24.4, range: 18-37) participated in the experiment for course credit or 30 NIS.

Procedure

Four white squares, in diamond formation, were presented against grey background on a computer monitor of 1680X1050 resolution. Participants were requested that when a black dot appears on screen, they locate it as quickly as possible and press a keyboard arrow key corresponding to the square that the dot appeared in (up arrow for top square, left arrow for left square, etc.). Each target was presented for 3s or until a button was pressed. A 500ms inter-stimuli interval (ISI) separated each target disappearance from the appearance of the next target (Figure 1). The experiment comprised 648 targets. Keypress times and eye movements were collected throughout the experiment (eye movements captured using SMI iView 250 RED).

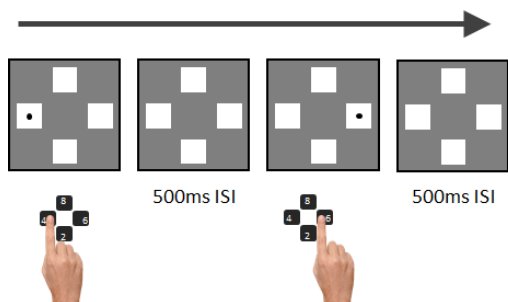


Figure 1: Experimental paradigm.

Fixed-sequence group. For participants in the fixed-sequence group ($n=38$), target locations followed a repeating sequential order of length twelve that is second-order conditional (SOC), meaning that location frequency and first-order transitions are counter-balanced (Reed & Johnson, 1994). Two such sequences were used. Target locations for twenty-four participants followed *sequence A*: 3-4-2-3-1-2-1-4-3-2-4-1, and target locations for fourteen other participants followed *sequence B*: 3-4-1-2-4-3-1-4-2-1-3-2 (numbers indicating location: 1 = down, 2 = left, 3 = right, 4 = up). Nine concatenated sequences constituted one block. Each block started from a different position within the sequence.

Random-sequence group. For participants in the random-sequence group ($n=30$), 648 targets were shown in an order with location statistics that maximally mimicked those of the SOC sequences, but without containing a repeating sequential order. Therefore, all locations appeared at similar frequencies without repeating the same location back-to-back, and first order transitions were counter-balanced, as in SOC sequences. In contrast to the SOC sequences, however, second order transitions were counter-balanced as well (and hence uninformative). Lastly, twelve second order transitions that constitute a reversal (e.g. 2-1-2, 3-4-3, etc.) occurred at near uniform frequency, embedding in the stimuli stream reversals at a ratio similar to that existing in the two SOC sequences.

Item-based analysis

Eye movements captured during ISIs underwent fixation analysis using SMI's built-in velocity-based algorithm for fixation detection. Fixations were found in 99% of ISIs. Fixations were considered anticipation of an upcoming target at the target location that they were closest to, effectively dividing the screen into four distinct areas of interest (AOIs). When fixations were detected in more than one AOI during an ISI, the last fixated AOI was considered the anticipation.

Under SOC sequences, the minimal learning required to solve the task is of the second order (e.g. after 2-1 comes 4) as frequency and first order transitions are uninformative. It has also been shown that participants preferentially acquire second order information in similar statistical learning tasks (Cleeremans & McClelland, 1991; Maheu, Dehaene, & Meyniel, 2019). Therefore, we analyzed second order learning, regarding triplets of target locations as separate items. Each pair of neighboring stimuli was regarded as a prefix, and the anticipation in the ISI that succeeded it was regarded the response to that prefix (hence reflecting learning of second order transitions). Twelve different prefixes appear in the stimuli stream (all possible pairs of 1, 2, 3 & 4, except for pairs in which both elements are the same). There are four possible responses to each prefix, and so forty-eight unique items exist in analysis.

Our analysis searched in each participant for anticipatory behaviors that were non-random. If a participant had

responded in a consistent manner to a prefix whenever that had appeared in the stimuli stream, we consider that a meaningful non-random behavior of the participant. If this is found at the end of the experiment, we consider it a “learned” behavior. If it is found at the very onset of the experiment, we consider it a “bias”.

Randomness was determined according to a right-tailed binomial test. Responses during the first or last block were examined per each prefix, corresponding to the first or last nine responses to that prefix (for analysis of learning or of bias, respectively). The amount of each response was tested for exceeding chance level, considering that the chance probability of producing each response is 25%. Response at a rate that could be attributed to chance with a p -value of 0.01 or under was considered non-random (corresponding to at least six identical responses out of the nine examined ones).

Throughout the paper, in cases in which sphericity is violated according to Mauchly’s test of sphericity ($p < 0.05$), Greenhouse-Geisser corrected values are reported.

Results

Participants grew quicker to respond to target locations throughout the task (main effect of block: $F(2.5, 164.9) = 38.28$, $p < 0.001$, $\eta_p^2 = 0.37$), but more so in the fixed-sequence group (interaction between group and block: $F(2.5, 164.9) = 4.28$, $p = 0.01$, $\eta_p^2 = 0.06$). Rate of correct oculomotor anticipation revealed similar effects (main effect of block: $F(3.4, 226.6) = 21.55$, $p < 0.001$, $\eta_p^2 = 0.25$; interaction between group and block: $F(3.4, 226.6) = 7.12$, $p < 0.001$, $\eta_p^2 = 0.1$; Figure 2).

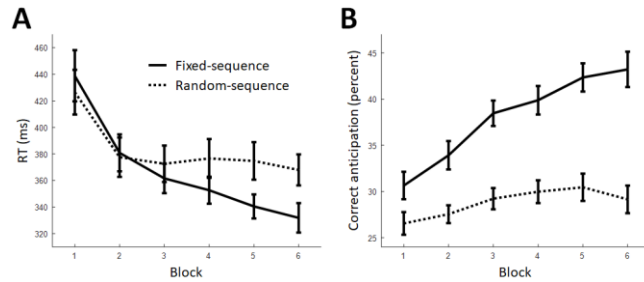


Figure 2: **A)** Average RT to target locations throughout the experiment. **B)** Average percentage of trials in which participants correctly anticipated upcoming target locations. Error bars represent SEM.

Item-based analysis

Our method of item-based analysis had detected an average of 5.9 biases ($SD = 2.6$) per each participant in the study, and an average of 6.2 learned items ($SD = 2.7$). In both cases, more non-random behaviors were found within the fixed-sequence group. Fixed-sequence group had an average of 6.6 biases ($SD = 2.6$) and random-sequence an average of 5.1 ($SD = 2.3$; $t(66) = 2.38$, $p < 0.05$). Fixed-sequence group had learned an average of 7 items ($SD = 2.4$) and random-

sequence and average of 5.2 ($SD = 2.8$; $t(66) = 2.99$, $p < 0.01$).

With regards to their identity, the twenty-four items that belonged to either *sequence A* or to *sequence B* were more prevalent as biases than the twenty-four items that belonged to neither of the sequences across all participants (main effect of item type: $F(1,65) = 7.18$, $p < 0.01$, $\eta_p^2 = 0.1$, main effect of participant group: $F(2,65) = 2.81$, $p = 0.07$, $\eta_p^2 = 0.08$, interaction between item type and participant group: $F(2,65) = 1.65$, $p = 0.2$, $\eta_p^2 = 0.05$). This is due to the fact that half of the latter type of items entail a reversal (a response in which gaze returns to the location of the target before last), which were less common as biases in all participants ($t(67) = 9.95$, $p < 0.001$).

Across all participants, items of *sequence A* were equally common as biases as items of *sequence B* ($t(67) = 1.23$, $p = 0.22$). Within participants of the fixed-sequence group only, items from the sequence governing the task were just as likely to be biases as items from the alternative sequence ($t(37) = 0.52$, $p = 0.61$).

Learning, however, was preferentially of the sequence driving the stimuli stream within the fixed-sequence group ($t(37) = 5$, $p < 0.001$). Participants of this group learned 4 items of their own sequence ($SD = 2$) and only 1.9 items of the other sequence ($SD = 1.2$). No preference for learning items of either sequence was found within the random-sequence group ($M = 1.5$, $SD = 1.3$ for *sequence A* items and $M = 1.4$, $SD = 1.2$ for *sequence B* items; $t(29) = 0.52$, $p = 0.61$). See Figure 3 for a heatmap of individual biases and learning as identified in analysis.



Figure 3: Heatmaps of non-random oculomotor behaviors – biases on the left and learning on the right. Rows indicate participants and columns indicate items. The top group of rows are participants of the fixed-sequence group whose target location order was determined by *sequence A*. The middle group of rows are fixed-sequence participants whose target location order was determined by *sequence B*. The bottom group of rows are participants of the random-sequence group. Items of *sequence A* and those of *sequence B* are grouped in the leftmost and middle group of columns, respectively. The other 24 items that are not part of either *sequence A* or *sequence B* are grouped on the right.

Effect of bias on performance and learning

In order to assess the influence of biases on sequence learning, we next turned to analyze participants of the fixed-sequence group only. Participants of the fixed-sequence group had 6.6 biases on average ($SD = 2.6$). Of these, 4.5 biases ($SD = 2.4$) conflicted with the correct response (e.g. for a participant of *sequence B*, in which 3-1 is followed by 4, a bias existed to respond 2 after 3-1). 2.1 ($SD = 1.4$) biases matched the correct response.

At the beginning of the experiment, biases dictated the rate of correct response to their prefix, depending on whether they matched the correct response or not. The effect of this different starting point diminished with practice ($F(1,453) = 27.11, p < 0.001, \eta_p^2 = 0.06$). But, even by the last block of the experiment, the rate of correct response differed according to what relevant bias had existed in the first block ($F(2,453) = 38.09, p < 0.001, \eta_p^2 = 0.14$). Correct response rate was highest in items whose starting point was highest (matching biases: $M = 67.4\%$, $SD = 30.1\%$), and lowest when the starting point was lowest (conflicting biases: $M = 31.2\%$, $SD = 29.2\%$; Figure 4A). As nine of our subjects underwent extended training, we were able to perform a preliminary examination of whether the effect of biases persisted over an experiment three times longer. Indeed, an effect of initial bias existed even eighteen blocks since experiment onset ($F(2,105) = 8.89, p < 0.001, \eta_p^2 = 0.15$), as correct response to items with matching biases reached 81.2%, ($SD = 20.2\%$), but only 48.7% ($SD = 35\%$) to items with conflicting biases (inset of Figure 4A).

The existence of a bias at experiment onset, and whether the bias matched or conflicted with the correct response, also predicted whether the correct response was eventually learned ($F(2,453) = 32.23, p < 0.001, \eta_p^2 = 0.13$). When there was no relevant bias related to a prefix, the correct response was learned in 32% of the cases. When a bias existed and conflicted with the correct response, the correct response was learned in only 19.4% of cases. When the bias and the correct response were the same, learning occurred in 67.5% of cases (Figure 4B).

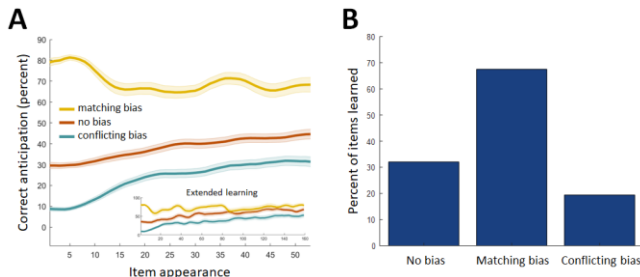


Figure 4: **A)** Percent of correct oculomotor anticipation of items according to what bias, if any, existed for their prefix. In inset is the same graph, but only for the nine participants that underwent extended learning over 18 blocks. Shaded area represents SEM. **B)** Percent of items correctly learned given the bias that existed for their prefix.

Agreement between bias and experiment group

To examine the effect of biases on SRT outcomes, we quantified per each participant the extent to which her biases fit the experimental group she was assigned to. Per each bias that matched the sequence governing the participant's stimuli stream we counted one fit point (+1), and per each bias that conflicted with the sequence we deducted one fit point (-1). This gave a score that reflected each individual's fit to the experimental group she was randomly assigned to, based on her a-priori biases. This score could theoretically vary between -12 (if the participant had a biased response to all prefixes and none of them were the correct response) and 12 (if all prefixes had a correct bias). But no such extreme case existed, and the average fit score was -2.4 ($SD = 2.9$).

There are more options and hence a higher probability of an a-priori bias to conflict with the correct response than to match it, as is evident in the amount of conflicting and matching biases detected in our analysis, and so most participants had a negative fit score (71.1%). To assess the effect of fit, we therefore compared participants that had a non-negative fit score (the best fitting participants), to an equal number of participants with the worst fit score (score of -4 or below; $n = 11$ for both groups). The remaining 16 participants that had intermediate fit scores were not included in this analysis.

In correct anticipation rate, both main effects of block and of fit were significant ($F(2.4, 47) = 18.16, p < 0.001, \eta_p^2 = 0.48$ and $F(1, 20) = 12.34, p < 0.01, \eta_p^2 = 0.38$ respectively), and there was no interaction between block and fit ($F(2.4, 47) = 2.1, p = 0.13, \eta_p^2 = 0.1$; Figure 5A). However, the lack of interaction could be due to the test being underpowered, as it seems that correct anticipation of participants with a bad fit catches up with correct anticipation of participants whose fit is not bad (correct anticipation in block 6: $t(20) = -1.41, p = 0.17$). More participants with extended training are needed to examine whether this is true.

In RT, ANOVA across trajectories of the two groups revealed, again, a main effect of block on RT ($F(2.4, 48.6) = 15.43, p < 0.001, \eta_p^2 = 0.44$), a main effect of fit ($F(1, 20) = 10.27, p < 0.01, \eta_p^2 = 0.34$), and no interaction between fit and block ($F(2.4, 48.6) = 1.49, p = 0.23, \eta_p^2 = 0.07$; Figure 5B). The effect of fit on RT persisted throughout all six blocks of this experiment and, in contrast to correct anticipation, showed no sign of diminishing. Moreover, a correlation existed between fit score and individual RT across all participants ($r(36) = -0.39, p < 0.05$), such that higher fit scores predicted faster RT during the last block of the experiment (Figure 5C).

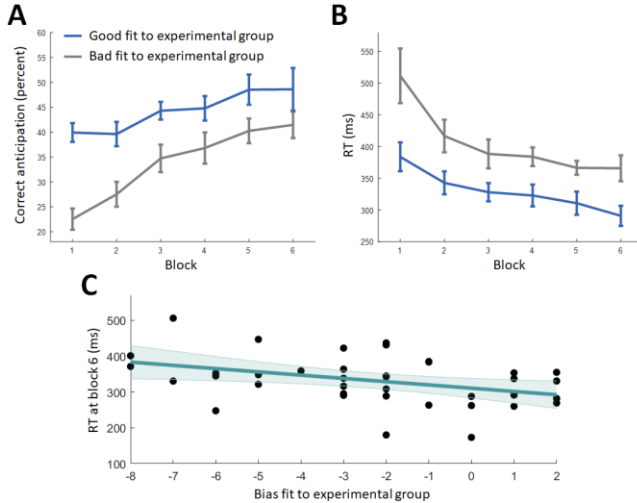


Figure 5: **A)** Average correct anticipation rate of participants whose biases were either a reasonable or a poor fit with the sequence driving their stimuli stream. **B)** Average RT according to the same partition. Error bars represent SEM. **C)** Relationship between fit score and RT by the end of the experiment. Dots reflect individual participant data, solid line reflects best linear fit and shaded area reflects 95% confidence interval.

Discussion

The current work examines consistent patterns of anticipation at the onset and conclusion of an SRT task, on an individual- and item-based level. Results indicate that there is a strong relationship between behavioral biases participants enter an experiment with, and the knowledge they eventually gather during that experiment.

Prevalence and identity of biases

Across participants, we found that half of the stimuli pairs ($M = 5.9$) elicited a consistent oculomotor response at experiment onset. The prevalence of this finding highlights the size of the tabula rasa assumption we tend to make in laboratory studies of learning, usually implicitly.

Biases were slightly more prevalent in fixed-sequence participants than in random-sequence participants ($M = 6.6$ vs. $M = 5.1$, respectively). We suggest that this may be due to weak biases that could be lost within the first block already, and so were dependent on the stimuli stream for reinforcement. Weak biases could be preserved if they matched the stimuli stream in fixed-sequence participants. However, they would always be conflicting in the random-sequence case (as any bias) because any consistent response in that stream would produce more mistakes than correct responses, and so they would be lost in this group. Similarly, items corresponding to reversals were less prevalent as biases than other items, and we suggest again that this is due to their elimination very early in the experiment due to negative reinforcement. “Stuck” items, in which gaze remains in the same location as the last target in

the prefix, are even more erroneous in the current task than reversals, but biases of this sort did exist to a larger extent. We suggest that this is because leaving gaze in place indicates not only anticipation of an upcoming target, but also lack of any anticipation attempt (passive waiting for the next target). Future study of biases should perhaps include a test phase prior to the main task in which biases are tested in isolation before the experimental material is presented. This may prevent contamination by rapid learning, but on the other hand may introduce other confounds stemming from exposure to the experimental material in configurations that differ from those of the task.

Biases distributed similarly between *sequence A* items and *sequence B* items in fixed-sequence and in random-sequence participants. The fact that these biases appeared similarly regardless of whether they were correct or incorrect for the participant affirms that the biases we detect derive mostly from processes that are a-priori to the experiment material.

We found no principle predicting what makes certain items more prone to become a bias than others. We hypothesized that salient structures, either in stimuli or in response, such as shifting eyes to the left following targets at the right and then at the bottom (forming a clockwise motion) would tend to appear as a bias more often than others. Such structures have been shown to be easier to learn in SRT (Reed & Johnson, 1994). However, we found no evidence to support this hypothesis. Additional study with larger samples is required to examine whether universal principles exist in determining common biases.

Item-based learning

Participants in our study that had a repeating sequence underlying their stimuli stream had learned, according to our analysis, only 4 of the possible 12 items comprising the sequence (33%). This sparsity resonates with the finding that only a subset of learnable items in statistical learning tasks are eventually learned (Schlagbauer, Muller, Zehetleitner, & Geyer, 2012). Participants have learned an additional 3 items, on average, that did not correspond to the stimuli sequence and were hence incorrect. Participants of the random-sequence group have learned 5.2 items during the experiment, all of which are incorrect.

Learning in our analysis reflects a convergence towards a certain behavior at the last block of the experiment. We find that this convergence is meaningful, as participants predominantly converge towards the correct behavior when a correct behavior exists (Tal & Vakil, in press). Convergence onto incorrect responses, both in the fixed-sequence group and in the random-sequence group, could reflect either incorrect learning (e.g. learning based on first order transitions, that will be correct only in a third of trials), or acquisition of stereotypical responses that are unrelated to the prefix that preceded them (e.g. due to fatigue). We find slight support for both explanations, but on a small scale. Importantly, we find similar entropy in learned responses of both groups during the final block of

the study, ruling out the possibility that learning in the random-sequence group reflects uniform responses independent of the their prefix more so than in the fixed-sequence group. Additional study is needed to understand the nature of incorrect learning, whether it is resilient and enhanced over time, and whether a relationship exists between correct and incorrect learning occurring in the task. Both with regards to correct and to incorrect learning, a certain amount of false positive findings can be expected to exist in our analysis. Specifically, tests of item learning rely on an alpha level of 0.01, simplistically suggesting a false positive rate of 1%. This would correspond to 8.2 items incorrectly classified as learned, constituting 1.9% of the 422 items that we have detected as learned. This approximation is probably inflated, as 66% of the detected learned items were at a response rate corresponding to alpha levels of 0.001 or lower, but in either case it should have only a negligible effect on results.

Influence of biases on learning

The particular biases that participants expressed at the very first trials of the experiment had a lasting effect on performance. Correct anticipation of all items improved with practice, but the gap produced by biases at experiment onset still existed in performance when the experiment ended. Preliminary evidence suggests that this gap in performance prevails even after extended training, three times longer than the current experiment (whose length is quite typical of SRT studies). Our findings therefore suggest that biases affect eventual outcomes of SRT studies, both in the oculomotor measure of performance and in the more standard measure of RT learning effects.

Moreover, initial biases greatly influenced the amount, identity and extent of the items that participants had eventually learned during the experiment. Items presented to a participant were over three times more likely to be learned if the participant had biases that happened to match them, than if she had biases that happened to conflict with them.

Taking learning potential into consideration in addition to learning endpoint is especially warranted in domains in which learning is considered a reliable individual capacity such as in statistical learning (Siegelman & Frost, 2015). This is because overriding a conflicting bias with a new response entails more learning than simply learning the same response when no predisposition needs to be overcome, let alone simply maintaining that predisposition if it happens to agree with the to-be-learned response to begin with. Worse endpoint performance could therefore nevertheless conceal a better capacity for learning. Of our thirty-eight participants in the fixed-sequence group, thirteen (34.2%) would be in a better position to learn the sequence presented to them if they had been randomly assigned to the other sequence rather than to the one they were assigned to. In the field of sequence learning, such information is of added value because SRT is used in clinical populations to assess learning dysfunctions. We hope similar work in this vein would enhance the resolution

of investigation into individual capabilities and biases, to form a measure of sequence learning that could inform on finer aspects of learning and on the various reasons that may underlie differences in performance.

Acknowledgements

This work was supported by the Farber Alzheimer's Center Foundation (# 259147). We thank Simone Schwizer Ashkenazi for data collection of the random-sequence group, and Ariel Roitman and Michal Roth for data collection of the extended-training group.

References

- Cleeremans, A., & McClelland, J. L. (1991). Learning the structure of event sequences. *Journal of Experimental Psychology: General*, 120(3), 235–253.
- Clegg, B. A., Digirolamo, G. J., & Keele, S. W. (1998). Sequence learning. *Trends in Cognitive Sciences*, 2(8), 275–81. Retrieved from
- 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*.
- Gebhart, A. L., Aslin, R. N., & Newport, E. L. (2009). Changing Structures in Midstream: Learning Along the Statistical Garden Path. *Cognitive Science*, 33(6), 1087–1116.
- Green, M. F., Kern, R. S., Williams, O., McGurk, S., & Kee, K. (1997). Procedural Learning in Schizophrenia: Evidence from Serial Reaction Time. *Cognitive Neuropsychiatry*, 2(2), 123–134.
- Harlow, H. F. (1949). The formation of learning sets. *Psychological Review*, 56(1), 51–65.
- Jungé, J. a, Scholl, B. J., & Chun, M. M. (2007). How is spatial context learning integrated over signal versus noise? A primacy effect in contextual cueing. *Visual Cognition*, 15(1), 1–11.
- 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.
- 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).
- 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), 420–32.
- Nissen, M. J., & Bullemer, P. (1987). Attentional Requirements of Learning: Evidence from Performance Measures. *Cognitive Psychology*, 19(1), 1–32.

- 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.
- 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.
- Tal, A. & Vakil, E. (in press). How sequence learning unflods: Insights from anticipatory eye movements. *Cognition*.
- 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.