

# Splitting the variance of statistical learning performance: A parametric investigation of exposure duration and transitional probabilities

Louisa Bogaerts<sup>1</sup> · Noam Siegelman<sup>2</sup> · Ram Frost<sup>2,3,4</sup>

Published online: 7 January 2016  
© Psychonomic Society, Inc. 2016

**Abstract** What determines individuals' efficacy in detecting regularities in visual statistical learning? Our theoretical starting point assumes that the variance in performance of statistical learning (SL) can be split into the variance related to efficiency in encoding representations within a modality and the variance related to the relative computational efficiency of detecting the distributional properties of the encoded representations. Using a novel methodology, we dissociated encoding from higher-order learning factors, by independently manipulating exposure duration and transitional probabilities in a stream of visual shapes. Our results show that the encoding of shapes and the retrieving of their transitional probabilities are not independent and additive processes, but interact to jointly determine SL performance. The theoretical implications of these findings for a mechanistic explanation of SL are discussed.

**Keywords** Visual statistical learning · Sequence learning · Individual differences

Statistical learning (SL), or learning of the distributional properties of sensory input across time and space, is the mechanism by which cognitive systems discover the underlying

regularities in the environment. As such, SL plays a key role in the segmentation, discrimination, and categorization of input, shaping the basic representations for a wide range of sensory, motor, and cognitive abilities (see Frost, Armstrong, Siegelman, & Christiansen, 2015, for a discussion). The term “SL” thus refers to the ability to learn and assimilate an array of possible statistical properties of sensory events. These include their aggregated relative frequency, their variance, and mostly, the extent of their co-occurrence (see Thiessen, Kronstein, & Hufnagle, 2013, for a review). The present article is concerned with the latter form of computation.

Starting from Saffran's original work (Saffran, Aslin, & Newport, 1996), which revealed that infants are able to segment speech on the basis transitional probabilities, a large number of studies have demonstrated that people often display remarkable sensitivity to the co-occurrence of items embedded in a continuous stream. This has been shown across ages from adults to newborns (e.g., Bulf, Johnson, & Valenza, 2011), across sensory modalities (*Visual*: e.g., Fiser & Aslin, 2001; Kirkham, Slemmer, & Johnson, 2002; Turk-Browne, Jungé, & Scholl, 2005; *Auditory*: e.g., Gebhart, Newport, & Aslin, 2009; Saffran, Newport, Aslin, Tunick, & Barrueco, 1997; *Tactile*: e.g., Conway & Christiansen, 2005), with both adjacent (e.g., Endress & Mehler, 2009) and nonadjacent contingencies (e.g., Gomez, 2002; Newport & Aslin, 2004). Interestingly, although in all of these studies the tested sample as a group showed clear evidence of learning, not all individuals were shown to perform better than chance (see Siegelman & Frost, 2015, for a discussion). What determines the efficacy of detecting the co-occurrence of events in a stream? Why do some individuals show clear evidence of learning in typical SL tasks, whereas others seem to perform at chance? What are the cognitive operations underlying this capacity? These complex questions hold the promise of revealing critical insights

---

✉ Louisa Bogaerts  
bog.louisa@gmail.com

<sup>1</sup> CNRS & Aix-Marseille University, Marseille, France

<sup>2</sup> Hebrew University of Jerusalem, Jerusalem, Israel

<sup>3</sup> Haskins Laboratories, New Haven, CT, USA

<sup>4</sup> Basque Center of Cognition, Brain and Language, San Sebastian, Spain

regarding the mechanisms driving SL, leading to deeper comprehension of what SL abilities could predict and why.

In a recent theoretical discussion of the factors influencing the variance of SL performance, Frost and his colleagues (2015) suggested that this variance should be split into two main sources: (1) variance related to the *efficiency in encoding* the individual elements in the stream within the modality of their presentation—that is, the ability to create internal representations of each element of the continuous perceptual input—and (2) variance related to the relative computational *efficiency of detecting* the distributional properties of the encoded representations, registering their transitional probabilities. Whereas the efficacy of creating detailed and reliable internal representations of individual elements appearing in a fast sequential input could be traced to the neuronal mechanisms that determine the effective resolution of one's sensory system, the computational efficiency of detecting the transitional probabilities of these elements could be traced to capacities for binding temporal and spatial contingencies by the medial-temporal lobe memory system (Karuza et al., 2013; Kim, Lewis-Peacock, Norman, & Turk-Browne, 2014; Schapiro, Gregory, Landau, McCloskey, & Turk-Browne, 2014). This view suggests that both encoding and binding abilities constrain the learning of regularities, and they *jointly* determine the actual performance of an individual in a given SL task. Moreover, it presupposes some form of temporal processing modularity, in which the internal representations computed from the inputs are subject to higher-level computations that bind them to register their distributional properties. Here we explored for the first time the possible predictions of this theoretical framework. We orthogonally manipulated factors related to encoding and binding constraints, and measured their relative contributions to SL performance on the group and individual levels.

In the study we report here, we focused on performance in the extensively used, visual statistical-learning (VSL) task (Arciuli & Simpson, 2011, 2012; Frost, Siegelman, Narkiss, & Afek, 2013; Siegelman & Frost, 2015; Turk-Browne et al., 2005). In the VSL task, participants are presented with a stream of complex visual shapes, organized in pairs or triplets whose constituent shapes follow each other in a predictable sequence (typically, transitional probability = 1). Following a familiarization phase, participants are tested to assess their ability to report which shapes appeared in the stream in the original order. The VSL task allows a unique opportunity to experimentally address our theoretical question by disentangling the (1) encoding and (2) learning components of statistical dependencies in SL. In any continuous stream of shapes, experimenters can independently manipulate (1) shape exposure duration (ED)—that is, the amount of time that the stimulus is physically available for processing—and (2) the transitional probabilities (TPs) within the shapes. Whereas ED is a parameter affecting the efficacy of

processing the visual stimuli for encoding them into internal representations (e.g., Loftus & Kallman, 1979; Potter & Levy, 1969), TP is a parameter related to the efficiency of registering their distributional properties. Jointly manipulating these parameters within subjects could thus provide important information regarding individual susceptibility to encoding constraints versus individual sensitivity to correlational transparency (see Frost et al., 2015, for discussion).

In the present study, we did exactly this. Our participants, university students, participated in a series of VSL tasks, in all of which they watched evenly paced streams of complex visual shapes. However, rather than using a fixed ED or a fixed TP, as in most current SL studies (but see Hunt & Aslin, 2001), in each session we manipulated the EDs and TPs of shapes in the stream in a within-subjects factorial design. The ED was set to 200, 600, or 1,000 ms per shape, and the TP between shapes could be quasi-regular (.6, .8) or fully regular (1.0). Following the familiarization phase, participants were tested to assess how well they had learned the statistical contingencies of the shapes in each of the streams, given the different presentation constraints. Thus, by looking at the change in performance across the tasks, we could examine the independent influences of the EDs and TPs within the stream on SL performance, as well as their possible interaction.

This design allowed us to address, in parallel, critical theoretical questions that have not been addressed so far: What are the impacts of incremental ED and TP changes on SL? Do they impact SL independently of (and additively to) each other, as would be predicted by a temporal-processing modularity assumption, or do they show substantial interaction? If they do interact, what is the nature of this interaction? Finally, what does the distribution of individual sensitivities to both factors look like in the population? More generally, we asked, how is human performance in extracting regularities from the input affected by different constraints (linearly? logarithmically? inverted-U shaped?) when constraints are imposed on the time allocated for encoding events and when the extent of event predictability is varied.

## Method

### Participants

Fifty adults (38 females, 12 males), all students at the Hebrew University, participated in the study for course credit or payment. Their ages ranged from 21 to 27 years (mean = 23.4). The participants were all native Hebrew speakers.

### Design and materials

The experiment required each participant to perform nine VSL tasks. The VSL tasks included 22 complex visual shapes

(Turk-Browne et al., 2005). In each condition and for each participant, 16 of the 22 shapes were randomly chosen and randomly organized to create eight ordered pairs (the remaining six shapes were used for the screening items; see below). The eight pairs were presented continuously, one after the other, in a random order, to create a familiarization stream in which each pair appeared 24 times, with the constraint that the same pair could not be repeated twice in a row.

ED (the time that the shape appeared on the screen) and TP (the conditional probability of the second shape of each pair appearing after the first) were manipulated so that each factor included three levels: ED of 200, 600, and 1,000 ms, and TPs of .6, .8, and 1. Combining the three levels of each factor created nine tasks overall (see Fig. 1). The interval between shapes was always fixed to 100 ms, to avoid introducing any chunking bias. The manipulation of TPs was done by including random noise in the .6 and .8 conditions: For example, for each pair AB during familiarization in the TPs = .8 subtest, shape B appeared after shape A 80 % of the time, but 20 % of the time, shape B was randomly replaced by another shape X, while avoiding immediate repetition of shapes. Shape X was never the start of a new pair (i.e., only after its presentation was a new pair presented).

Depending on the ED used, keeping the number of repetitions and the shape interval constant, the familiarization phase of each of the tasks lasted from 2 to 7 min. Participants were asked to attend to the stream and were not told that the stream was constructed of pairs. Following the familiarization stream, participants were instructed that they would now see two pairs of shapes on the screen (see Fig. 1) and that their task would be to report which pair was more familiar to them. They were then tested with 38 two-alternative forced choice trials. Thirty-two of the test trials contrasted (1) “true pairs”—two shapes that appeared as a pair during the familiarization phase (the TP between the shapes being .6, .8 or 1.0, depending on the condition)—and (2) “foils”—two shapes that did not appear as a pair during familiarization. Foils were constructed without violating the position of the shapes within the original pairs (e.g., for two true pairs AB and CD, the possible foils could be AD or CB, but not AC or DB). Scores in the SL task then ranged from 0 to 32, calculated as the number of correct identifications of pairs during the test phase. The remaining six test trials aimed to identify and screen participants who did not attend the familiarization stream. These trials contrasted “true pairs” with a pair containing a novel shape that had not appeared at all during familiarization (see Romberg & Saffran, 2013, for a similar procedure). Participants who missed 18 or more of the 54 screening items (six in each of the nine tasks) were excluded from the analyses. Following this screening, eight participants were excluded from the analysis. Due to a technical problem, the data of two participants in the ED = 200, TP = 1 condition were not saved. All subsequent analyses are based on the remaining 42 participants.

## General procedure

The nine SL subtests were initiated by the participants from home, through an online platform. All nine subtests had to be completed in a period of 30 days, with no less than 24 h between sessions. The mean time interval between sessions was 2.3 days ( $SD = 1.2$ ). Participants were instructed to do the task alone in a quiet room and to avoid external distractions (i.e., to turn off their cell phone and music), and they were asked to have only the experiment window open. The order of the tasks was random.

## Results

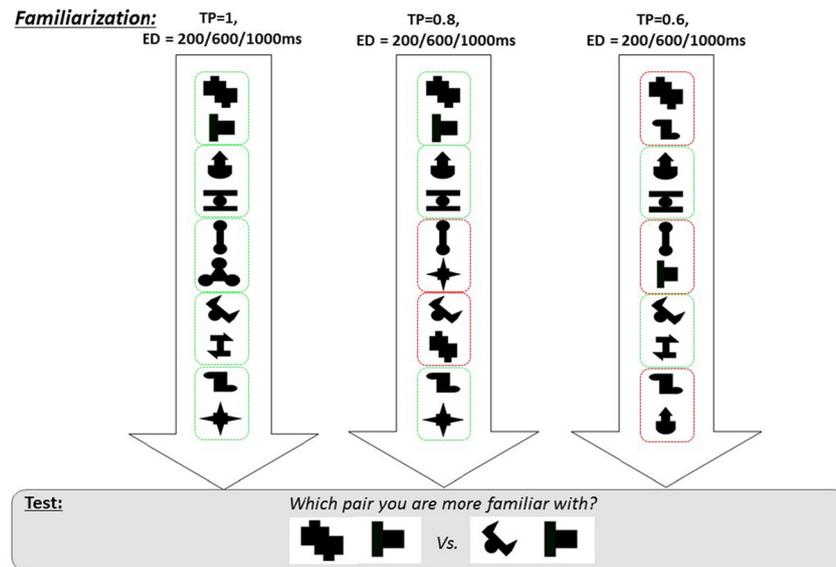
### Group level

Table 1 summarizes performance in the nine tasks. Performance in all nine conditions was significantly better than chance (in one-sample *t* tests comparing mean performance to 50 % chance, all  $ps < .01$ ). This suggests that successful learning was present even with the very fast presentation rates and lower TPs. A one-way repeated measures analysis of variance (ANOVA), with session number (1–9) as a predictor, revealed that the level of performance did not change significantly across the nine sessions [ $F(8, 312) = 1.49, p = .16$ ; see also Siegelman & Frost, 2015, for similar findings].

In order to examine the influences of TP, ED, and their interaction on SL performance, we conducted a logistic mixed-effect analysis using the lme4 package in R (Bates, Maechler, Bolker, & Walker, 2015). The dependent variable was accuracy in the forced choice test (excluding the screening items). The model included the fixed effects of TP, ED, their interaction, the position of the target pair within the forced choice question (i.e., whether the target was first or second), and trial number in the test.<sup>1</sup> The random-effect structure included a by-subjects random intercept and random slopes for TP, ED, and their interaction. The predictors TP and ED were both centered and standardized, trial number was centered, and the target position variable was dummy-coded (target in first position = 0; second position = 1). The model included  $N = 12,032$  observations and had a log-likelihood of  $-7,664.7$ .

We found significant main effects of TPs ( $\beta = .18, SE = 0.04, p < .001$ ) and of EDs ( $\beta = .23, SE = 0.04, p < .001$ ). These effects are presented in Fig. 2: As can be seen, the effect of ED was found to be linear (with an improvement of 6.0 % from 200 to 600 and a similar improvement of 4.5 % from 600

<sup>1</sup> Trial number was not a significant predictor of performance,  $\beta = -.002, SE = 0.002, p = .18$  suggesting that the repeating the same target pairs and foils in the test phase did not alter performance.



**Fig. 1** The nine visual statistical-learning tasks: combining three transitional probability (TP) levels (.6, .8, and 1) and three exposure duration (ED) levels (200, 600, and 1,000 ms)

to 1,000). A paired *t* test on the individual-differences scores between the adjacent levels of ED confirmed that the extent of improvement between the lower two and the top two ED levels did not differ significantly [ $t(41) = 0.61, p = .54$ ].<sup>2</sup>

In contrast, the effect of TPs seemed to deviate from linearity (with a small difference of 1.9 % between the lower TPs of .6 and .8, which implicated both quasi-regularities, and a larger difference of 5.9 % between .8 and full regularity, TP = 1). Indeed, a paired *t* test revealed a marginally significant difference between the TP = .6 to .8 and the TP = .8 to 1 differences [ $t(41) = 1.78, p = .08$ ], suggesting a possible non-linearity in the influence of TP on SL performance. This suggests a possibly qualitative difference between the full regularity and quasi-regularity of shapes in the stream.

In addition to the two significant main effects of TP and ED, we found a significant interaction between TP and ED ( $\beta = .09, SE = 0.03, p < .01$ ), which is depicted in Fig. 3. At the fast ED of 200 ms, performance was low even with high TPs (although above chance in all three conditions), suggesting that the fast presentation rate impaired the detection of co-occurrence of shapes. For the intermediate and slow ED levels (600 and 1,000 ms), the lines reflecting the three levels of TP diverge; whereas performance was found to be similar in the two conditions that implicated quasi-regularities (TP = .6 and .8), a different trend was revealed for full regularity (TP = 1), with high performance already at ED = 600 ms. These data suggest an interesting interplay between encoding constraints

and the extent of regularity in determining SL performance. We will discuss this further in the General Discussion.

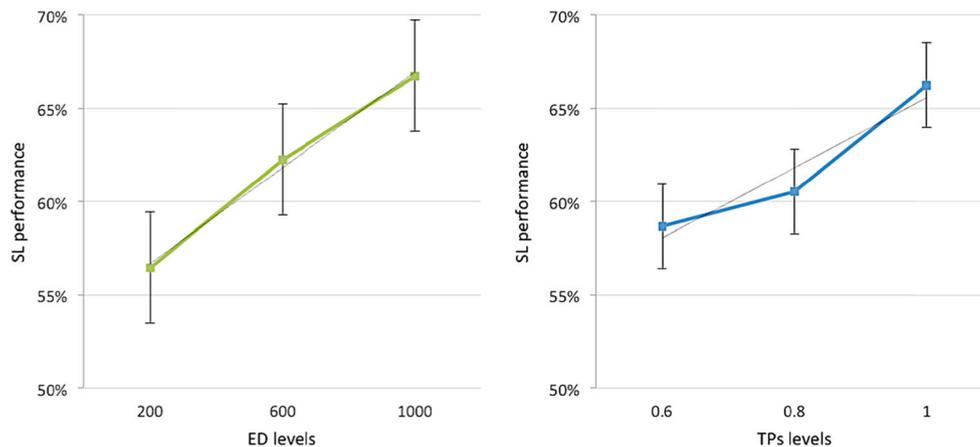
**Individual level**

We now turn to the insights revealed by considering patterns of individual performance. For each individual we extracted his/her slope for the effect of ED and TPs from the mixed logit model by looking at the by-subject random slopes. Note, that given the small number of manipulated levels of ED and TP per individual (three levels in each factor only), these slopes inevitably incur substantial noise. Nevertheless, in spite of this noise, some patterns stand out, pointing to interesting directions for future research. Figure 4a presents the scatterplot of slopes for ED and TPs of all participants in our study. The striking result is the high correlation between the slopes within subjects ( $r = .55, p < .001$ ). Two outlier observations stand out, and if they are removed the correlation between slopes increases to  $r = .75 (p < .001)$ . Thus, it seems that participants who showed greater sensitivity to changes of ED tended to show greater sensitivity to changes of TP, and vice versa. Note that some of this observed correlation could be driven by participants who did not show any evidence of learning: Because these participants failed to exhibit any learning, they

**Table 1** Mean performance rate in each of the nine tasks (standard deviations are in parentheses)

	TP = .6	TP = .8	TP = 1
ED = 200 ms	54.8 % (11 %)	56.3 % (14 %)	58.4 % (14 %)
ED = 600 ms	59.5 % (13 %)	59.7 % (15 %)	67.6 % (17 %)
ED = 1,000 ms	61.8 % (12 %)	65.6 % (17 %)	72.8 % (18 %)

<sup>2</sup> Because in our procedure the interstimulus interval between shapes remained constant (see the Design and Materials section), the possibility that our results reflected the rate (or length) of presentation, rather than ED per se, cannot be overruled and should be acknowledged.



**Fig. 2** Mean scores as a function of the ED (left) and TP (right) manipulations. Error bars denote standard errors. The thin gray lines represent the best linear fits

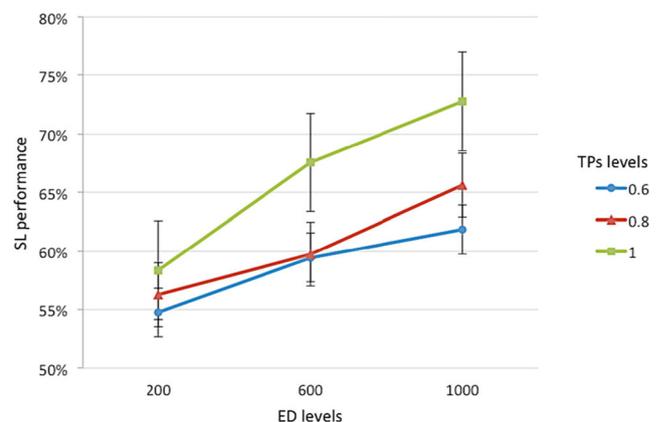
produced similar flat slopes for ED and TP, affecting the size of the correlation. However, even if individuals who did not exhibit significant learning at the individual level across the tasks ( $n = 9$  excluded participants) were removed from the analysis,<sup>3</sup> a substantial correlation between susceptibility to ED and TP still remained ( $r = .48$ ,  $p < .01$ ; see Fig. 4b). These intriguing findings suggest that from an individual-differences perspective, the ability to benefit from extended event duration and greater event regularity seems to be a unified individual capacity.

## Discussion

In the present study, we independently manipulated TP and ED in a visual SL task to dissociate factors related to the encoding of visual shapes and the higher-order process of learning their distributional properties. We asked how each type of constraint affects SL and how their interaction determines performance in the task. Our results provide a set of critical findings. First, we found that, at least within the range of 200 to 1,000 ms, ED impacts SL performance in a linear way, so that longer exposure of shapes results in better learning of their conditional probabilities. This converges with the earlier evidence provided by Turk-Browne et al. (2005) and Arciuli and Simpson (2011), who manipulated ED between subjects and reported improved SL performance at slower presentation rates (see also Emberson, Conway, & Christiansen, 2011). Second, we found that introducing quasi-regularity in the stream impacts learning along a trajectory that seems to deviate from linearity. Although this deviation was marginally significant, our results showed relatively

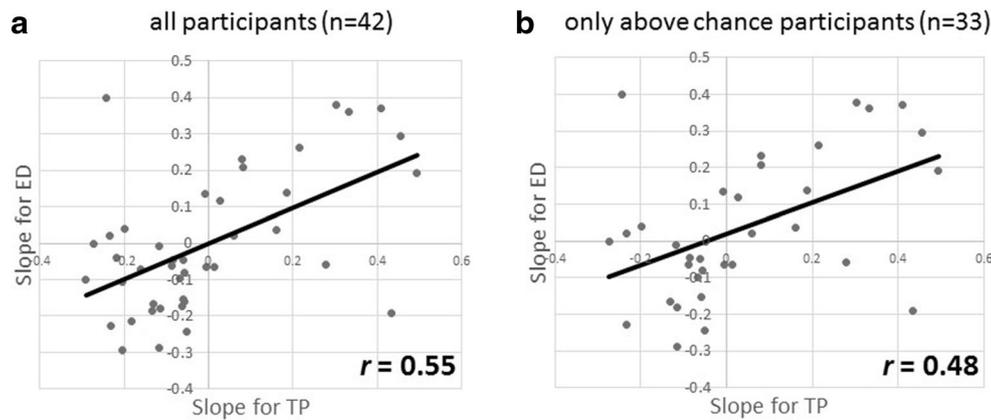
small changes in SL performance when TP increased from .6 to .8, but a substantially large improvement when the TP implicated full regularity. This pattern of performance suggests that, at least at the group level, full regularity of shapes in the streams may be qualitatively different from any quasi-regularity, in terms of improving SL performance.

However, importantly, our experimental design allowed us to go beyond the independent influences of ED and TP within the stream on SL performance, to examine their interaction. The striking finding of our study was the interplay between encoding constraints and the extent of regularity in determining the learning outcomes. Overall, our results suggest that sensitivity to the extent of TP in the stream was modulated by ED, and vice versa. Although our findings show that even very short EDs (of 200 ms) were sufficient to encode the visual shapes, resulting in above-chance learning, the extent of regularity of the shapes in the stream had a relatively smaller impact on learning, so that SL performance was relatively low even with full regularity. With additional exposure (ED of 600 ms), a large difference in performance between full regularity (TPs = 1) and quasi-regularity (TPs = .6, .8) emerged, but no difference between the two levels of quasi-regularity.



**Fig. 3** Interaction between TP and ED. Error bars denote standard errors

<sup>3</sup> The exclusion criterion for this analysis was set to success in 159 trials out of the 288 across the nine tasks (i.e., a mean success rate of 55.2%). According to the binomial distribution, this is the minimal number of successful trials needed to present significantly above-chance learning at the individual level.



**Fig. 4** Correlations between the slopes for ED and TP, over the whole sample (Fig. 4a, left) and only for participants exhibiting learning at the individual level (Fig. 4b, right)

Full sensitivity to the full range of TPs was found only with the longest exposure of the shapes. This pattern of findings suggests that encoding shapes and retrieving their TPs are not independent and additive processes. Rather, the distributional properties of shapes in the stream and their predictability may serve to *facilitate* their encoding, in the case of suboptimal, shorter EDs, and conversely, an increase in the exposure time enhances sensitivity to fine differences in TP. These findings have implications for a mechanistic description of the cognitive events occurring in the typical VSL task. Rather than considering temporal-processing modularity, in which the encoding of shapes into internal representations feeds into the subsequent phase of extracting their distributional properties, the encoding and extraction of TPs seems to be a two-way street, with each dimension affecting the other. Whether this bidirectional dependency is causal in one direction or the other requires further investigation.

The present findings are relevant to current debates regarding the extent of modality specificity in SL (see Frost et al., 2015, for a review and discussion), and the relations between the subprocesses involved in SL (e.g., Thiessen et al., 2013). In the context of visual shapes, recent imaging studies have implicated, on the one hand, higher-level visual networks (Nastase, Iacovella, & Hasson, 2014), and on the other hand, the domain-general hippocampus and medial-temporal lobe memory system (Schapiro et al., 2014; Turk-Browne, Scholl, Chun, & Johnson, 2009). Our findings thus offer possible constraints for understanding how both modality-specific (encoding of visual shapes) and modality-general (extracting distributional properties) computations result in the extent of learning regularities in the visual modality. These processes do not seem to be independent and sequential, so that the completion of one would initiate the launching of the other.

Our discussion so far has focused on the group-level performance, yet from an individual-differences perspective, another striking result is the high correlation between the ED and

TP trajectories within participants. This high correlation suggests that individuals who showed greater sensitivity to changes of ED tended to also show greater sensitivity to changes of TP, and vice versa (note that this correlation is based on a relatively large number of participants and held even after we removed individuals who did not exhibit significant learning). This is an intriguing finding, since it suggests that individual abilities to overcome both encoding constraints (here operationalized as limitations of event duration) and learning constraints (here operationalized as noise related to event regularity) are interrelated. A possible interpretation of this finding is that, perhaps, the high correlation between sensitivities to ED and TP was driven by peripheral factors, such as a general state of attentiveness to the task. To investigate this hypothesis, we calculated the partial correlations between the individual ED and TP slopes, controlling for average performance on the screening items, a proxy for attentiveness. The partial correlation between slopes within participants was still large and significant (full sample,  $r_{\text{partial}} = .43$ ,  $p < .01$ ; after removing outliers,  $r_{\text{partial}} = .68$ ,  $p < .001$ ; with only above-chance participants,  $r_{\text{partial}} = .40$ ,  $p < .05$ ).

A point that deserves some attention is the presence of a number of negative individual slopes (see Fig. 4). Whereas a negative slope associated with the ED manipulation intuitively makes sense (some individuals who are fast encoders might fail to allocate attention when shapes in the stream are presented at a slow rate), the negative slopes for TP are harder to explain. Although it is possible that the negative slopes for TP represent simple noise, which is inevitable in such an experimental design, possible insight for this phenomenon can be drawn from recent studies suggesting that different populations of neurons encode full regularity and quasi-regularity (Nastase et al., 2014). From the perspective of information theory, quasi-regularity is more informative than full regularity. Indeed, Kidd, Piantadosi, and Aslin (2012) have recently shown that infants maximally attend to stimuli that are neither too predictable nor too unpredictable. The negative slopes of

some of the participants in our sample may hint toward individual differences in the point of optimal degree of the extent of regularity for learning. This, however, will require further investigation, aiming to establish whether individual slopes for ED and TP are indeed stable characteristics of an individual (see Siegelman & Frost, 2015, for measures of reliability in SL tasks).

In conclusion, the present study suggests that manipulating task parameters in a within-subjects parametric design provides considerable insight regarding the cognitive operations underlying visual SL. Research using a similar methodology has the promise of establishing how encoding and higher-order learning factors account for the variance in performance in other modalities, leading to a better understanding of the mechanisms of SL.

**Author note** This article was supported by the Israel Science Foundation (Grant No. 217/14, awarded to R.F.), and by the National Institute of Child Health and Human Development (Grant Nos. RO1 HD 067364, awarded to Ken Pugh and R.F., and PO1-HD 01994, awarded to Haskins Laboratories). L.B. is a research fellow of the Fyssen Foundation. We are indebted to Steve Frost for his valuable comments.

## References

- Arciuli, J., & Simpson, I. C. (2011). Statistical learning in typically developing children: The role of age and speed of stimulus presentation. *Developmental Science, 14*, 464–473.
- Arciuli, J., & Simpson, I. C. (2012). Statistical learning is related to reading ability in children and adults. *Cognitive Science, 36*, 286–304.
- Bates, D., Maechler, M., Bolker, B., & Walker, S. (2015). lme4: Linear mixed-effects models using Eigen and S4 (R package version 1.1-8). Retrieved from <http://cran.r-project.org/package=lme4>
- Bulf, H., Johnson, S. P., & Valenza, E. (2011). Visual statistical learning in the newborn infant. *Cognition, 121*, 127–132.
- Conway, C. M., & Christiansen, M. H. (2005). Modality-constrained statistical learning of tactile, visual, and auditory sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 31*, 24–39. doi:10.1037/0278-7393.31.1.24
- Emberson, L. L., Conway, C. M., & Christiansen, M. H. (2011). Timing is everything: Changes in presentation rate have opposite effects on auditory and visual implicit statistical learning. *Quarterly Journal of Experimental Psychology, 64*, 1021–1040.
- Endress, A. D., & Mehler, J. (2009). The surprising power of statistical learning: When fragment knowledge leads to false memories of unheard words. *Journal of Memory and Language, 60*, 351–367.
- Fiser, J., & Aslin, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science, 12*, 499–504.
- Frost, R., Armstrong, B. C., Siegelman, N., & Christiansen, M. H. (2015). Domain generality versus modality specificity: The paradox of statistical learning. *Trends in Cognitive Sciences, 19*, 117–125. doi:10.1016/j.tics.2014.12.010
- Frost, R., Siegelman, N., Narkiss, A., & Afek, L. (2013). What predicts successful literacy acquisition in a second language? *Psychological Science, 24*, 1243–1252.
- Gebhart, A. L., Newport, E. L., & Aslin, R. N. (2009). Statistical learning of adjacent and nonadjacent dependencies among nonlinguistic sounds. *Psychonomic Bulletin & Review, 16*, 486–490.
- Gomez, R. L. (2002). Variability and detection of invariant structure. *Psychological Science, 13*, 431–436.
- Hunt, R. H., & Aslin, R. N. (2001). Statistical learning in a serial reaction time task: Access to separable statistical cues by individual learners. *Journal of Experimental Psychology: General, 130*, 658–680. doi:10.1037/0096-3445.130.4.658
- Karuz, E. A., Newport, E. L., Aslin, R. N., Starling, S. J., Tivarus, M. E., & Bavelier, D. (2013). The neural correlates of statistical learning in a word segmentation task: An fMRI study. *Brain and Language, 127*, 46–54. doi:10.1016/j.bandl.2012.11.007
- Kidd, C., Piantadosi, S. T., & Aslin, R. N. (2012). The Goldilocks effect: Human infants allocate attention to visual sequences that are neither too simple nor too complex. *PLoS One, 7*, e36399. doi:10.1371/journal.pone.0036399
- Kim, G., Lewis-Peacock, J. A., Norman, K. A., & Turk-Browne, N. B. (2014). Pruning of memories by context-based prediction error. *Proceedings of the National Academy of Sciences, 111*, 8997–9002.
- Kirkham, N. Z., Slemmer, J. A., & Johnson, S. P. (2002). Visual statistical learning in infancy: Evidence for a domain general learning mechanism. *Cognition, 83*, B35–B42.
- Loftus, G. R., & Kallman, H. J. (1979). Encoding and use of detail information in picture recognition. *Journal of Experimental Psychology: Human Learning and Memory, 5*, 197–211.
- Nastase, S., Iacovella, V., & Hasson, U. (2014). Uncertainty in visual and auditory series is coded by modality-general and modality-specific neural systems. *Human Brain Mapping, 35*, 1111–1128.
- Newport, E. L., & Aslin, R. N. (2004). Learning at a distance I: Statistical learning of non-adjacent dependencies. *Cognitive Psychology, 48*, 127–162.
- Potter, M. C., & Levy, E. I. (1969). Recognition memory for a rapid sequence of pictures. *Journal of Experimental Psychology, 81*, 10–15. doi:10.1037/h0027470
- Romberg, A. R., & Saffran, J. R. (2013). All together now: Concurrent learning of multiple structures in an artificial language. *Cognitive Science, 37*, 1290–1320.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science, 274*, 1926–1928. doi:10.1126/science.274.5294.1926
- Saffran, J. R., Newport, E. L., Aslin, R. N., Tunick, R. A., & Barrueco, S. (1997). Incidental language learning: Listening (and learning) out of the corner of your ear. *Psychological Science, 8*, 101–105.
- Schapiro, A. C., Gregory, E., Landau, B., McCloskey, M., & Turk-Browne, N. B. (2014). The necessity of the medial temporal lobe for statistical learning. *Journal of Cognitive Neuroscience, 26*, 1736–1747. doi:10.1162/jocn\_a\_00578
- Siegelman, N., & Frost, R. (2015). Statistical learning as an individual ability: Theoretical perspectives and empirical evidence. *Journal of Memory and Language, 81*, 105–120.
- Thiessen, E. D., Kronstein, A. T., & Hufnagle, D. G. (2013). The extraction and integration framework: A two-process account of statistical learning. *Psychological Bulletin, 139*, 792–814. doi:10.1037/a0030801
- Turk-Browne, N. B., Jungé, J. A., & Scholl, B. J. (2005). The automaticity of visual statistical learning. *Journal of Experimental Psychology: General, 134*, 552–564. doi:10.1037/0096-3445.134.4.552
- 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*, 1934–1945. doi:10.1162/jocn.2009.21131