Global and Local Statistical Regularities Control Visual Attention to Object Sequences

Alexa Romberg[1, 2], Yayun Zhang[3], Benjamin Newman[4], Jochen Triesch[5], Chen Yu[3]

[1] Department of Human Development & Quantitative Methodology and [2] Center for Advanced Study of Language
University of Maryland, MD, USA
[3] Department of Psychology & Brain Science and [4] School of Informatics and Computing
Indiana University-Bloomington, IN, USA
[5] Frankfurt Institute for Advanced Studies, Frankfurt, Germany
[1, 2] aromberg@umd.edu, [3,4] {yayzhang, newmanba, chenyu}@indiana.edu, [5] triesch@fias.uni-frankfurt.de

Abstract—**Many previous studies have shown that both infants and adults are skilled statistical learners. Because statistical learning is affected by attention, learners' ability to manage their attention can play a large role in what they learn. However, it is still unclear how learners allocate their attention in order to gain information in a visual environment containing multiple objects, especially how prior visual experience (i.e., familiarly of objects) influences where people look. To answer these questions, we collected eye movement data from adults exploring multiple novel objects while manipulating object familiarity with global (frequencies) and local (repetitions) regularities. We found that participants are sensitive to both global and local statistics embedded in their visual environment and they dynamically shift their attention to prioritize some objects over others as they gain knowledge of the objects and their distributions within the task.**

I. INTRODUCTION

Humans are quite good at detecting patterns in their environment, a skill that has been termed statistical, or distributional, learning. Statistical learning has been demonstrated in the sensitivity of human infants, children and adults to the predictability of object features [1], object sequences [2], [3], sequences of linguistic elements like syllables or words [4], and co-occurrence of words with particular objects [5]. Generally, statistical learning is thought to be an automatic process, though one that requires at least some attention to the stimulus [6]. As learners gain information about the patterns in their environment, we expect this information to shape their attention and, consequently, their future learning.

What can the dynamics of a person's attention tell us about what they are learning in a particular environment? In general, humans are assumed to distribute their attention in order to gain information about the environment, such as building representations of objects. Attention to a stimulus is thought to decline as the learner reaches ceiling for information "stored" about the object. This general model has provided the framework for analysis of infant looking behavior in experimental psychology [7]. Inherent in the Hunter & Ames (1988) descriptive model of infants' attention is a tension between novel and familiar stimuli. Learners are thought to be motivated to attend to familiar stimuli until they are sufficiently "learned", and then learners are thought to switch to preferring a novel stimulus.

In support of this general model, 3.5- and 6.5-mo-old infants have been found to prefer a familiar stimulus after limited exposure to that stimulus but to prefer a novel stimulus after a longer exposure [8]. When individual infants make the shift from preferring the familiar to the novel is thought to depend on individual differences, such as processing speed, and on properties of the stimulus (i.e., stimulus complexity). Given the same stimulus complexity, older infants are expected to shift from a familiarity to a novelty preference more quickly [9].

While the limits of our visual processing abilities are well documented (approximately 4 objects can be tracked and held in memory), less clear are the specific factors that cause attentional shifts between objects, for both infant and adult learners. The goal of the current project was to investigate how the dynamics of visual attention are related to learning in a multi-object environment. In particular, how does the frequency of an objects' presentation (a key component of statistical learning) drive a learner's looking behavior?

In the current study, we focus on how characteristics of the learning environment influence adult learners' attention allocation. Participants are shown simple scenes composed of four objects (Figure 1). Participants freely explore these scenes while their eye movements are being tracked. We manipulated the frequency of presentation of each object in order to investigate two types of statistics that might influence learners' looking behavior. First, across all trials, some objects were presented twice as often as other objects, creating a *global* frequency distribution. Second, a subset of objects on each trial is repeated on the next trial, creating a *local* frequency (i.e., repetition) distribution. This design is illustrated in Table 1.

This paradigm allows us to test how global and local statistics may trade-off or work together in order to drive participants' attention. For any particular trial, each object represented one of the four cells in Table 1. The top-left and bottom-right cells are easy to interpret, as the global and local statistics reinforce one another. The Low Frequency New item (LF-New) will always be the most novel stimulus on the screen and the High-Frequency Repeated item (HF-Rep) will always be the most familiar stimulus on the screen. Thus, learners seeking new information should look more to the LF-New than the HF-Rep objects.

The other diagonal in Table 1 is less straightforward. For these items, the global and local statistics conflict. The High Frequency-New item (HF-New) should be highly familiar overall (global) but is relatively novel at that moment (local). In contrast, the Low Frequency-Repeated item (LF-Rep) should be relatively novel overall but is familiar at that moment. Our design allows us to assess the relative

importance of these conflicting influences over time. In addition, we include two different object set sizes. Including more objects in the set reduces the frequency of presentation for all objects (i.e., global frequency), allowing us to investigate how looking behavior is influenced by smaller vs. larger difference between global and local statistics.

Table 1. Dimensions for object statistics across trials.

	Local Distribution	
Global Distribution	High Frequency-Repeated	High Frequency-New
	Low Frequency-Repeated	Low Frequency-New

For the purposes of the current report, our analyses focus on two aspects of participants' looking behavior: First, we asked whether participants actively control their attention across trials in response to the global item statistics embedded in the visual presentations. HF items were presented twice as frequently as LF items. Learners who have the goal of collecting an equivalent amount of information about each object over the course of the study, would be expected to adjust looking times in order to accommodate this difference. A learner might accumulate more looking time per presentation to each LF item compared with HF items with the end result that the learner accumulates about the same amount of looking time to each object by the end of the task, regardless of object frequency. Such an information-balance approach to visual attention would be in keeping with the general principles of the trade-offs between novelty and familiarity preferences discussed above.

Second, we asked how global and local regularities influence which objects participants prioritize for processing (i.e., fixate within the first second of trial onset) as participants gain experience with the tasks. We expected that early in each task participants would select objects randomly and that their looking behavior should become more systematic as they learn more about the system. Further, we predicted that participants might shift from an initial familiarity preference (prioritizing processing HF and/or repeated objects) to a novelty preference (prioritizing processing of LF and/or new objects).

II. METHOD

A. Participants

Twenty-three volunteers (13 females, $M_{age} = 23.57$, $SD_{age} = 3.68$) at Indiana University participated in the study.

B. Materials

Visual stimuli consisted of 27 color images of novel objects that were not readily nameable (see examples in Figure 1). Images were equal in size (300x300 pixels). All images were displayed on a white background on a 17 inch monitor with a resolution of 1280 × 1024 pixels.

C. Apparatus

The participants' eye gaze was measured by an Eye Tribe tracker. The eye-tracking system recorded gaze data at 30 Hz (accuracy = 0.5°- 1°). A chinrest was used to minimize head movements.

D. Experimental Design

Each participant completed 2 free-viewing tasks within a 20-minute session. Task 1 consisted of 36 trials and used 9 novel objects. Three objects belonged to a high frequency (HF) set while the other 6 objects belonged to a low frequency (LF) set. Four unique objects were presented on each trial. Two of these objects were pulled from the HF set, while the other 2 were pulled from the LF set. On each trial, one HF and one LF object persisted through to the next trial, but were displayed in two new locations. The two remaining stimuli were chosen by selecting at random one each from the HF and LF sets with the constraint that the LF object could not have been presented in the past 2 trials. No objects persisted for more than 2 consecutive trials. Across the 36 trials, each HF object appeared 24 times, and each LF object appeared 12 times (see Figure 1 and Table 2).

Task 2 consisted of 72 trials (HF: 6 objects; LF: 12 objects) and used the same design as Task 1, but with different object sets and frequencies (see Table 2), adding the constraint that the new HF object on each trial was not presented on the past 2 trials. The novel objects used in these two tasks did not overlap.

There were 9 possible display positions, evenly spaced along a circle centered on the middle of the monitor. These 9 positions (300x300 pixels) defined the 9 areas of interest (AOIs) in which looks were analyzed.
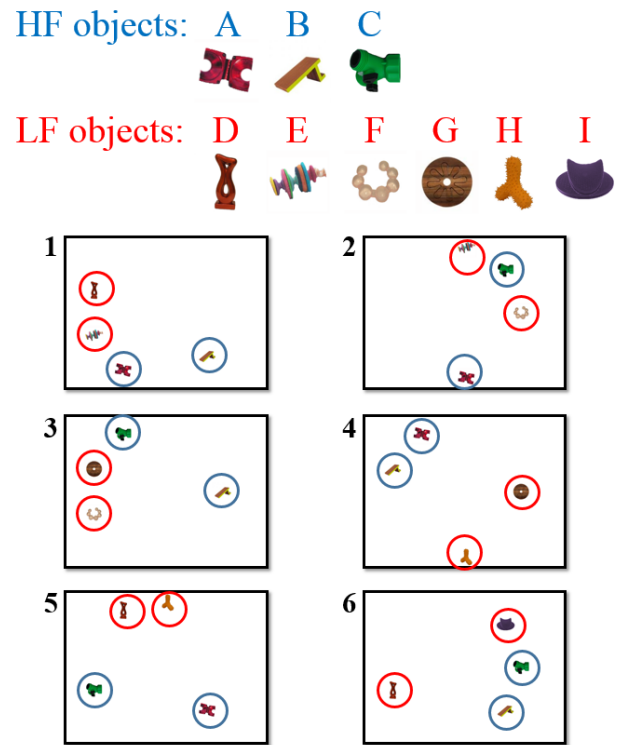

Figure 1. First 6 trials of Task 1. Blue circles/letters indicate HF objects and red circles/letters indicate LF objects. In trial 2, object A is HF-Rep, object C is HF-New, object E is LF-Rep, object F is LF-New. The colored circles are for illustration only and were not present during testing.

Participants were seated approximately 60cm from the monitor in a quiet room. They were told to watch and study the objects on the screen and that they would be asked questions about the objects at the end of the session.

The point of gaze was calibrated with a dot that appeared randomly at 9 locations across the screen. Each trial was preceded by a fixation cross in the center of the screen; the trial began (i.e., the objects were displayed) when participants' gaze was registered on the fixation cross. Trials were 5 seconds in duration. Trial order remained the same between participants but the order of tasks was randomized for each participant. Participants were allowed to take a short break between tasks and were required to redo calibration before each task. The entire testing session was about 20 minutes and no questions were given at the end of either task.

Table 2. Detailed design statistics for each Task.

		Task 1	Task 2
High Frequency	Number of objects	3	6
	Number of repetitions	24	24
	Frequency	66.6%	33.3%
Low Frequency	Number of objects	6	12
	Number of repetitions	12	12
	Frequency	33.3%	16.6%
Total	Number of objects	9	18
	Number of trials	36	72

III. RESULTS

Even though perfect tracking in a continuous mode was not possible due to technical limitations of the eye tracker or loss of attention, the overall quality of the tracking results was satisfactory. In both tasks, about 81% of trials reached 80% (roughly 2640 data points per trial). All trials were included in the following data analysis. For the purpose of analysis, a "look" was defined as a dwell time within an AOI lasting at least 0.1 seconds. When looks to the same AOI were separated by less than 0.5 seconds of missing data, the missing data was filled in as a look to the bracketing AOI.

Across both tasks, participants spent part of each trial looking at each of the 4 objects and generally made 6-8 total looks per trial (M_{task1}=6.70, SD_{task1}=1.80; M_{task2}=6.81, SD_{task2}=1.86). While the mean number of looks fluctuated from trial to trial, there was no clear trend across either task toward more or fewer looks over time (Figure 2 top panel).

Within a trial, individual look durations were unevenly distributed. As shown in Figure 3, look durations for both tasks were highly skewed. Most looks were about 0.4 seconds long, though looks ranged in duration from 0.1-4.5 seconds. The mean of the longest look made on each trial was between 1 and 2 seconds in duration (M_{task1}=1.39, SD_{task1}=0.35; M_{task2}=1.31, SD_{task2}=0.33, Figure 2 bottom panel). However, on average participants generated less than one look per trial that was longer than 1 second (M_{task1}=0.85, SD_{task1}=0.75; M_{task2}=0.80, SD_{task2}=0.86).
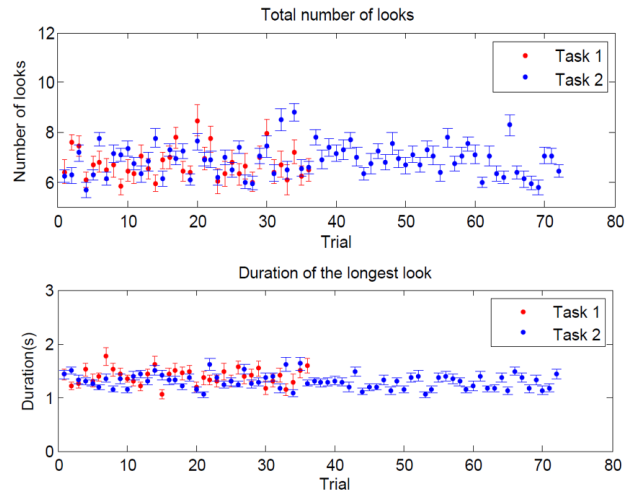

Figure 2. Top panel: total number of looks within each trial. Participants generally made 6-8 total looks per trial. Bottom panel: duration of the longest look within each trial. Mean duration of the longest look made on each trial was between 1 and 2 seconds.
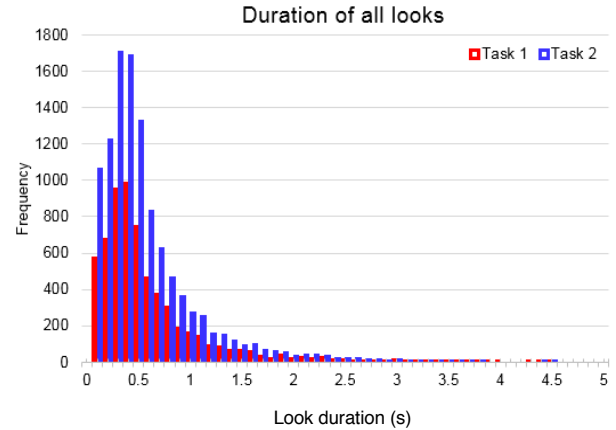

Figure 3. Distribution of look durations. Look durations were highly skewed with most looks last 0.4 seconds long.

A. Cumulative looking for HF and LF items

Our first question was the extent to which, across the whole of each task, participants actively controlled their attention in response to the global item statistics. Participants' mean cumulative looking to each object in Task 1 is plotted in Figure 4, with HF objects in the left panel and LF objects in the right panel. It is clear from the plots that presentation frequency is the primary driver of cumulative looking behavior. At the end of the task participants had accumulated twice as much looking time to the HF objects (*M*=21.6s) than the LF objects (*M*=11.9s). Further, the total cumulative looking time to the LF objects is approximately equal to the cumulative looking time to HF object after their first 12 presentations (*M*=11.1s). The same pattern was found in Task 2 (M_{HF}=20.5s, M_{LF}=11.8s).

The steady accumulation of looking time across trials suggests that participants allocated equal attention to each object on each trial rather than spending more time on LF than HF items.

Figure 4. Accumulated looking time of HF and LF objects in Task1. Participants allocated equal attention to objects on each trial despite objects' different presentation frequencies.

B. Item-level Analyses: Effects of global and local statistics

Our second question was how the global and local frequency statistics influenced looking behavior at different time points in each task. To investigate participants' sensitivity to the frequency statistics, we analyzed the probability that each of the 4 objects (HF-Rep, HF-New, LF-Rep and LF-New) would be fixated within the first second of the trial. We focused on the first second as a measure of which objects the participants placed the highest priority on processing. We examined looking behavior across the study by dividing each task into 12-trial blocks. Comparing patterns of looking across Task 1 and Task 2 allows us to see how set size (i.e., the total number of objects) influences looking behavior. In Task 1, the smaller set size allows faster accumulation of information about each object and quicker discrimination between High Frequency and Low Frequency objects.

The probability that participants would fixate each of the 4 objects during the first second of each trial is presented in Figure 5. Across all trials in all blocks, participants most frequently fixated between 1 and 2 objects during the first second of the trial, though sometimes participants fixated 3 or even 4 objects during the first second. The primary trend that stands out across the three blocks in both Tasks in Figure 5 is participants' increasing discrimination between the 4 objects during the first second. While the models detailed below include significant differences between objects even within the first block, those effects are much larger in later blocks than in Block 1. This general trend illustrates the development of participants' knowledge of the statistical structure of the learning environment and they are sensitive to the statistical regularities embedded in the environment.

In the following statistical analyses, the raw data for each object for each trial was fit with a logistic multilevel model using the lme4 package in R [10]. The model included fixed effects of Frequency (High vs. Low), Repetition (New vs. Repeated) and Task (1 or 2) and their interaction terms, as well as random intercepts for Subject and Trial. Additional random effect structures were tested, including a random intercept for Subject x Trial and Subject x Task interactions. However, because none of these random factors were estimated by the model to account for any variance, they were left out of the final models. High Frequency, New and Task 1 served as reference groups coded 0 for the models.

1) Block 1

Inspection of the top panel of Figure 5 reveals an immediate difference in how participants prioritized each of the four object types across the two tasks, with the small set size of Task 1 driving attention to the more novel objects (i.e., Low Frequency objects were more likely to be fixated than High Frequency objects). The opposite was true for Task 2, in which participants were least likely to fixate the most novel object (i.e., LF-New).

The logistic multilevel model described above was fit to the Block 1 data. The three-way interaction was significant (b=1.05, SE=0.37, z=2.81, p = 0.005), confirming different patterns of looking across the two tasks. Models were fit to each task individually.

The model fit for Task 1 revealed only a main effect of Frequency (b=0.27, SE=0.13, z=2.04, p = 0.041), indicating that participants were more likely to fixate the LF than HF objects. Neither the main effect of Repetition nor the interaction term was significant. The model fit to Task 2 revealed a significant Frequency x Repetition interaction (b=0.61, SE=0.26, z=2.30, p = 0.021), a marginal main effect of Frequency (b=-0.35, SE=0.19, z=1.87, p = 0.062) and no effect of Repetition. This pattern of results indicates that participants were more likely to fixate the HF-New than the LF-New object but were more likely to fixate the LF-Rep than HF-Rep objects.

These results can be interpreted within the global/local framework outlined in the Introduction (see Table 1). Within the first 12 trials of Task 1 participants were favoring the globally novel stimuli. The larger set size in Task 2, in contrast, appears to drive fixations to globally familiar stimuli, as the most novel stimulus (LF-New) was the least likely to be fixated in the first second of the trial and participants had a small preference for HF over LF items..

2) Block 2

In the second block, participants were less likely to fixate the most familiar item (HF-Rep) than the other items regardless of set size (Figure 5, middle panel). The logistic multilevel model described above was fit to the data for Block 2. The three-way interaction between Frequency, Repetition and Task was only marginally significant (b=-0.70, SE=0.39, z=-1.81, p = 0.070), reflecting the general similarity across the two Tasks. The Frequency x Repetition interaction was significant (b=1.90, SE=0.28, z=6.87, p < 0.001), as was the main effect of Repetition (b=-0.97, SE=0.19, z=-4.86, p < 0.001). While participants were more likely to fixate New rather than Repeated items, this was primarily true for the HF items. For LF items, this pattern was reversed, particularly in Task 1.

3) Block 3

Inspection of the lower panel of Figure 5 makes clear that in the third block of Task 1, participants prioritized processing of the Repeated items, showing a clear dominance of the local context. With the larger set size of Task 2 (and therefore less total familiarity with individual objects and their distributions), participants continued to show an interaction between global and local statistics in Block 3.

The same logistic mixed effect model as described above was fit to the data for Block 3. The three-way interaction between Frequency, Repetition and Task was significant (b=-2.46, SE=0.41, z=-6.07, p <0.001), reflecting the different patterns described above. Individual models were fit to the data from each Task. For Task 1, there was a significant main effect of Repetition (b=1.10, SE=0.12, z=-1.69, p <0.001) and no main effect of Frequency or interaction. For Task 2, both main effects and their interaction were significant: Frequency (b=1.02, SE=0.20, z=5.19, p <0.001), Repetition (b=0.53, SE=0.20, z=2.70, p=0.007), Frequency x Repetition (b=-2.14, SE=0.29, z=-7.46, p <0.001).
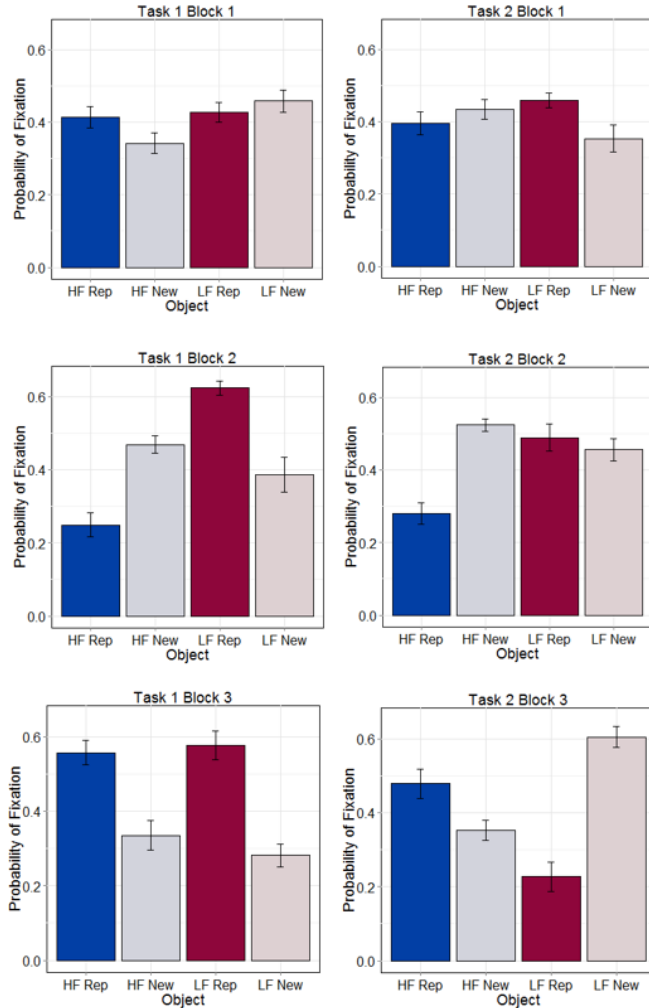

Figure 5. Mean probability of fixation within the first second for each of the 4 objects by Task (columns) and 12-trial Blocks (rows). Error bars represent standard error of the mean.

IV. DISCUSSION

The goal of the present study was to characterize how visual attention is allocated across objects varying in novelty. We manipulated both global frequency, by presenting some objects twice as frequently as others, and local frequency, by repeating some objects from one trial to the next. Additionally, the complexity of the task as a whole was manipulated by varying the total number of objects. Relative to Task 1, the larger set size of Task 2 effectively reduced the global frequency for all objects, creating a starker contrast between global and local statistics.

We found, somewhat surprisingly, that participants did not compensate for differences in item frequency by looking more to the less frequent items each time they were presented. While look dynamics early in the trial revealed sensitivity to global and local statistics, across the trial as a whole participants spent about equal time looking at each of the four objects. One possible explanation for this is that, given 4 objects and 5 seconds to look, participants had more time than required to process all the objects. With 4 objects displayed at a time, we would expect participants to be able to keep track of the objects in the display fairly easily [11] and perhaps the objects were simple enough that participants were able to reach ceiling for information gain within each trial and so did not need to strategize. If this is the case, then adjustments to the learning environment that increase the task demands would be expected to influence looking behavior. For example, one might increase the number of objects per trial, reduce the trial duration or increase the complexity of the objects. Each of these changes might drive participants to prioritize a subset of objects over others across the study as whole in addition to the initial fixations of the trial. Manipulating these task parameters would also allow a comparison between human data and that predicted by formal models [12].

We analyzed which objects participants fixated within the first second of each trial as a measure of which objects they were prioritizing for processing. In contrast to the total looking times, these item-level analyses demonstrated that participants' initial fixations within each trial were highly sensitive to global and local statistics, including set size.

Importantly, the different patterns of fixation across the two tasks during the first block of trials offer support for the idea that learners will show an initial familiarity preference until they have reached some processing threshold, at which point they show a novelty preference. During the first block, participants saw each HF object 8 times in Task 1 but only 4 times in Task 2. Initial looks for Task 1 show an overall (relatively weak) novelty preference while for Task 2 participants were actually least likely to look at the most novel object, focusing more of their initial processing on more familiar items. Thus, in the initial stages of learning, participants prioritized more familiar items, while they shifted to prioritizing more novel items as they became more familiar with both the items themselves and the statistical structure of the environment.

By the third block of trials, the differing set sizes were driving quite different looking behaviors during the first second of the trial. In Task 1, participants showed a marked effect of local statistics, prioritizing fixation on the repeated objects. By the end of the 3rd block in Task 1, participants had seen each HF object 24 times and each LF object 12 times. It seems that this amount of exposure was sufficient for participants to prioritize attending to the local context over the global frequency statistics. That is, participants may have learned all they could about the various objects themselves

and shifted their behavior to detecting the probability structure of the task as a whole.

It is possible, given the study design, that the looks to the repeating objects were evidence of statistical learning. Because no object persisted across more than 2 trials, the objects that were New on trial *n* would necessarily be the ones that were Repeated on trial *n+1*. With the smaller set size, this pattern may have become salient enough to drive participants to confirm which objects repeated before moving on to other objects.

Such a bias to confirm known patterns does not fit well within the novelty-familiarity framework described in the Introduction. However, it may be an important way that attention supports statistical learning. For example, more and less successful adult learners in a cross-situational word learning task show different patterns of gaze during learning [11]. In that task, participants view arrays of objects while listening to a list of words and the only cue to which word labels which object is the fact that they co-occur together across trials. Yu and colleagues (2012) found that while all participants began the study by fixating random objects during each word, by the end successful learners were consistently fixating the correct referent immediately after hearing the label, while less successful learners did not. Thus, it seems likely that confirmatory attention shifts play a role in developing and maintaining representations of statistical structure.

As the findings in the cross-situational word learning study discussed above illustrate, a characterization of real time attention dynamics based on object novelty could significantly influence our understanding in domains such as word learning. Many studies have shown that novel labels can drive infants' attention to novel objects, a behavior commonly known as the 'mutual exclusivity' response [13], [14], [15]. However, when information must be integrated across multiple contexts or situations (such as hearing the word *ball* on different days referring to different balls), successful learning requires balancing attention between the attraction of novelty and solidifying developing representations. Such trade-offs have been incorporated in recent successful models of cross-situational word learning [16].

The current work makes clear that real-time looking behavior is driven by complex interactions between different aspects of the environment. We did not measure individual differences, either with respect to what participants learned about the objects or to general participant characteristics such as processing speed, attention flexibility or working memory. Such invidiual characteristics are likely to be important to developing realistic models of attention dynamics. One recent model of multi-object visual attention focused specifically on how such characteristics might influence shifting between objects in an ideal learner [12]. The ideal leaner model was defined by parameters for individual differences relevant for object processing, including learning rate, memory decay, cost of switching between objects and prior knowledge of the objects. By simulating environments containing multiple objects, Pelz, Piantadosi and Kidd (2005) found that all these factors can influence how an ideal learner shifts attention between objects in order to maximize information gain.

Although the underlying dynamics of attention and learning are complicated, this model provides many testable hypotheses that can be applied to human attentional systems.

Development of such basic cognitive skills may also play an important role in the development of visual attention across infancy and childhood. Our study is a first step toward characterizing attention dynamics within a distributional learning framework in order to better understand how attention supports statistical learning.

REFERENCES

[1] J. Fiser and R. N. Aslin, "Unsupervised statistical learning of higher-order spatial structures from visual scenes," *Psychological Science*, vol. 12, no. 6, pp. 499–504, 2001.

[2] N. B. Turk-Browne, P. J. Isola, B. J. Scholl, and T. A. Treat, "Multidimensional visual statistical learning," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 34, no. 2, pp. 399–407, 2008.

[3] C. M. Conway and M. H. Christiansen, "Modality-constrained statistical learning of tactile, visual, and auditory sequences," *Journal of Experimental Psychology: Learning, Memory, and Cognition*, vol. 31, no. 1, pp. 24–39, 2005.

[4] J. R. Saffran, R. N. Aslin, and E. L. Newport, "Statistical learning by 8-Month-Old infants," *Science*, vol. 274, no. 5294, pp. 1926–1928, 1996.

[5] L. Smith and C. Yu, "Infants rapidly learn word-referent mappings via cross-situational statistics," *Cognition*, vol. 106, no. 3, pp. 1558–1568, 2008.

[6] J. M. Toro, S. Sinnett, and S. Soto-Faraco, "Speech segmentation by statistical learning depends on attention," *Cognition*, vol. 97, no. 2, pp. B25–B34, 2005.

[7] M. A. Hunter and E. W. Ames, "A multifactor model of infant preferences for novel and familiar stimuli," *Advance in Infancy Research*, vol. 5, no. 1, pp. 69–95, 1988.

[8] S. A. Rose, A. W. Gottfried, P. Melloy-Carminar, and W. H. Bridger, "Familiarity and novelty preferences in infant recognition memory: Implications for information processing," *Developmental Psychology*, vol. 18, no. 5, pp. 704–713, 1982.

[9] E. D. Thiessen, E. A. Hill, and J. R. Saffran, "Infant-directed speech facilitates word segmentation," *Infancy*, vol. 7, no. 1, pp. 53–71, 2005.

[10] D. Bates, M. Mächler, B. Bolker, and S. Walker, "Fitting linear mixed-effects models using lme4," *Journal of Statistical Software*, vol. 67, no. 1, 2015.

[11] C. Yu, Y. Zhong, and D. Fricker, "Selective attention in Cross-Situational statistical learning: Evidence from eye tracking," *Frontiers in Psychology*, vol. 3, 2012.

[12] M. Pelz, S. T. Piantadosi, and C. Kidd, "The dynamics of idealized attention in complex learning environments," in *Proc. IEEE-EPIROB*, 2015.

[13] E. Mather and K. Plunkett, "Novel labels support 10-month-olds' attention to novel objects," *Journal of Experimental Child Psychology*, vol. 105, no. 3, pp. 232–242, 2010.

[14] E. M. Markman, J. L. Wasow, and M. B. Hansen, "Use of the mutual exclusivity assumption by young word learners," *Cognitive Psychology*, vol. 47, no. 3, pp. 241–275, 2003.

[15] J. Halberda, "The development of a word-learning strategy," *Cognition*, vol. 87, no. 1, pp. B23–B34, 2003.

[16] G. Kachergis, C. Yu, and R. M. Shiffrin, "An associative model of adaptive inference for learning word–referent mappings," *Psychonomic Bulletin & Review*, vol. 19, no. 2, pp. 317–324, 2012.