

Perceived object trajectories during occlusion constrain visual statistical learning

JÓZSEF FISER

Brandeis University, Waltham, Massachusetts

BRIAN J. SCHOLL

Yale University, New Haven, Connecticut

AND

RICHARD N. ASLIN

University of Rochester, Rochester, New York

Visual statistical learning of shape sequences was examined in the context of occluded object trajectories. In a learning phase, participants viewed a sequence of moving shapes whose trajectories and speed profiles elicited either a bouncing or a streaming percept: The sequences consisted of a shape moving toward and then passing behind an occluder, after which two different shapes emerged from behind the occluder. At issue was whether statistical learning linked both object transitions equally, or whether the percept of either bouncing or streaming constrained the association between pre- and postocclusion objects. In familiarity judgments following the learning, participants reliably selected the shape pair that conformed to the bouncing or streaming bias that was present during the learning phase. A follow-up experiment demonstrated that differential eye movements could not account for this finding. These results suggest that sequential statistical learning is constrained by the spatiotemporal perceptual biases that bind two shapes moving through occlusion, and that this constraint thus reduces the computational complexity of visual statistical learning.

Human sensory systems handle enormous amounts of complex information during every waking moment. An increasing number of studies supports the idea that the processing of this information is implemented by an unsupervised form of continuous observational learning, which organizes both old and new sensory information into maximally efficient codes (Fiser & Aslin, 2005). This statistical learning mechanism has been demonstrated using speech streams in adults (Saffran, Newport, & Aslin, 1996), infants (Saffran, Aslin, & Newport, 1996), and nonhuman primates (Hauser, Newport, & Aslin, 2001), and using visual sequences in adults (Fiser & Aslin, 2002) and infants (Kirkham, Slemmer, & Johnson, 2002).

A key question is how such a powerful learning mechanism is sufficiently constrained to enable associative learning without suffering from intractable computational demands, given the infinite number of potentially relevant statistical structures in the input. We chose to investigate this question within the framework of visual statistical learning that occurs during brief periods of object occlusion. Occlusion is ubiquitous in the natural visual environment, since either the observer or the objects themselves undergo movement, and it is known that not all of the properties that define an object are retained in

working memory during occlusion (see, e.g., Simons & Rensink, 2005). In the present study, we ask which attributes of objects in pre- and postocclusion scenes influence how a statistical learning mechanism links those objects associatively.

Fiser and Aslin (2002) reported the first temporal statistical learning study involving visual shapes, and their paradigm employed occlusion to enable the successive presentation of semistructured shape sequences. A single object oscillated back and forth across a display; as it was occluded briefly in the center of the display, it underwent a transformation in shape according to particular rules, thereby creating a temporal stream of successive shapes moving uniformly and interrupted by brief occlusions. After a practice session, participants had to choose between two brief sequences of three consecutive shapes: The first sequence depicted shape transitions seen many times during practice, and the second depicted random transitions. Participants reliably favored the sequences seen during practice. In the Fiser and Aslin study, there was only one potential associative link, from shape n to shape $n+1$, and therefore only a single source for statistical learning: temporal succession. However, natural scenes contain many potential associative links between an object

that undergoes occlusion and the many objects available postocclusion. Are all such associative links considered equally by an unconstrained statistical learning mechanism, or do perceptual biases constrain inferences such that some links are preferred over others, and the computational demands on learning are thus reduced?

An ideal phenomenon for investigating this question is a bistable display in which two identical shapes appear in the upper corners of a square display and then move diagonally at the same speed until they simultaneously disappear behind a central occluder. The two shapes then reappear from behind the occluder and continue their straight motion trajectories along the two diagonal paths (Figure 1A). This is an ambiguous display that can be interpreted as either two shapes streaming past each other behind the occluder on a straight trajectory (a *streaming* display) or as two shapes colliding behind the occluder, abruptly changing their paths, and then reappearing on the same side of the occluder whence they originated (a *bouncing* display) (Feldman & Tremoulet, 2006; Sekuler & Sekuler, 1999).

In the present study, we asked whether statistical learning of shape pairs in multiobject occlusion displays would be influenced by the bias of perceiving a bouncing or a streaming motion pattern (Figure 1A). In theory, statistical learning could operate on the visual displays by assessing all of the available temporal co-occurrences of shape pairs equally, irrespective of the speed-defined bias to perceive the objects as bouncing or streaming. If so, then regardless of the speed change that shifts the percept from streaming to bouncing, both of the transitions from one shape to another during occlusion should be learned equally well. However, if visual statistical learning is constrained at the level at which representations of objects as spatiotemporal entities (individuals) are formed, then stimulus manipulations that result in changes in object speed should bias statistical learning during occlusion.

EXPERIMENT 1

Quantifying Bouncing Versus Streaming

The purpose of this experiment was to establish the optimal speed parameters for inducing either bouncing or streaming percepts with the same shapes and occluder to be used in Experiment 2.

Method

Participants. Sixteen undergraduates at the University of Rochester were each paid \$10.00 for their participation; all were naive to the experiment.

Stimuli. Sixteen arbitrary complex shapes were created, each with a unique color, and scaled to equal maximum extent of 2° in the Canvas drawing program (Figure 1B). Stimuli were presented on a 21-in. Sony Trinitron 500PS monitor at 1,024 × 728 resolution from a 1-m viewing distance using a Macintosh G4 computer and Macromedia Director software. On each trial, two shapes appeared in the two upper corners of the display 6° from the center of the screen, and then both shapes moved along their respective diagonal axes and disappeared behind a 2.4° × 11° black occluder in the middle of the display. While behind the occluder, both shapes changed randomly to different shapes in new colors (Figure 1C). These two new shapes emerged from behind the occluder and continued on diagonal trajec-

tories to the lower corners of the display. Total movement duration was fixed at 1.2 sec, but there were eight decelerating–accelerating speed profiles: 0% (uniform), 5%, 10%, 20%, 40%, 60%, 80%, and 100%, with 100% reflecting a peak change in speed from 15 cm/sec to 36 cm/sec. On each trial, the right and left shapes always had identical speed profiles in both the upper and lower halves of the display, and the degree of deceleration in the upper half was the same as the degree of acceleration in the lower half.

Design. There was an individual color–shape combination assigned to each participant. For each trial, 4 color–shape combinations from the set of 16, along with one of the eight possible speed profiles, were selected randomly. Each combination of color and shape and speed profile was used the same number of times during the experiment.

Procedure. Participants simply observed each animation and then used a keypress to report whether they had perceived streaming or bouncing. There were 12 repetitions of each speed profile randomly interleaved, resulting in a total of 96 trials.

Results and Discussion

Participants perceived approximately half of the 96 trials as streaming ($M = 47.12\%$, $SEM = \pm 1.82$) and half as bouncing [$M = 48.87\%$, $SEM = \pm 1.83$; $t(15) = 0.48$, $p = .639$]. These mean judgments showed a clear effect of speed profile [$F(7,105) = 14.32$, $p < .001$, $\eta^2 = .489$], shifting from robust streaming at uniform speeds (the 0% speed profile) to robust bouncing as the deceleration–acceleration profile reached 40% (see Figure 1D).

EXPERIMENT 2

Statistical Learning in Bouncing Versus Streaming Displays

In Experiment 2, we turned to the main question of this paper: Do the underlying perceptual mechanisms that result in bouncing versus streaming percepts during occlusion constrain visual statistical learning? The key test of this question is whether statistical learning of pairs of moving shapes that span occlusion is superior for pairs that are consistent with the bouncing or streaming perceptual bias compared with pairs that are inconsistent with the bias. It is important to note that both of these types of pairs were present in the familiarization displays an equal number of times, thus providing identical temporal order (associative) information for statistical learning.

Method

Participants. A group of 36 undergraduates, naive to the experiment, were each paid \$10.00 for their participation.

Stimuli. The same set of shapes from Experiment 1 was used, but with only two speed profiles: the uniform, or 0% speed profile (15 cm/sec), which induced a strong streaming percept, and the 40% deceleration–acceleration speed profile with a peak speed of 24 cm/sec, which induced a strong bouncing percept.

Design. The experiment consisted of two phases: familiarization and test. During the familiarization phase, the displays on individual trials differed from those in Experiment 1 in that only one shape started from either the left or right upper corner of the screen and moved diagonally behind the occluder. Then two shapes appeared from behind the occluder and continued moving to the two lower corners, just as in Experiment 1 (Figure 2A). Thus, due to the processes that underlie statistical learning, the two lower shapes were equally likely to be associated with the single upper shape that disappeared behind the occluder. In each display, all three shapes had either the uniform

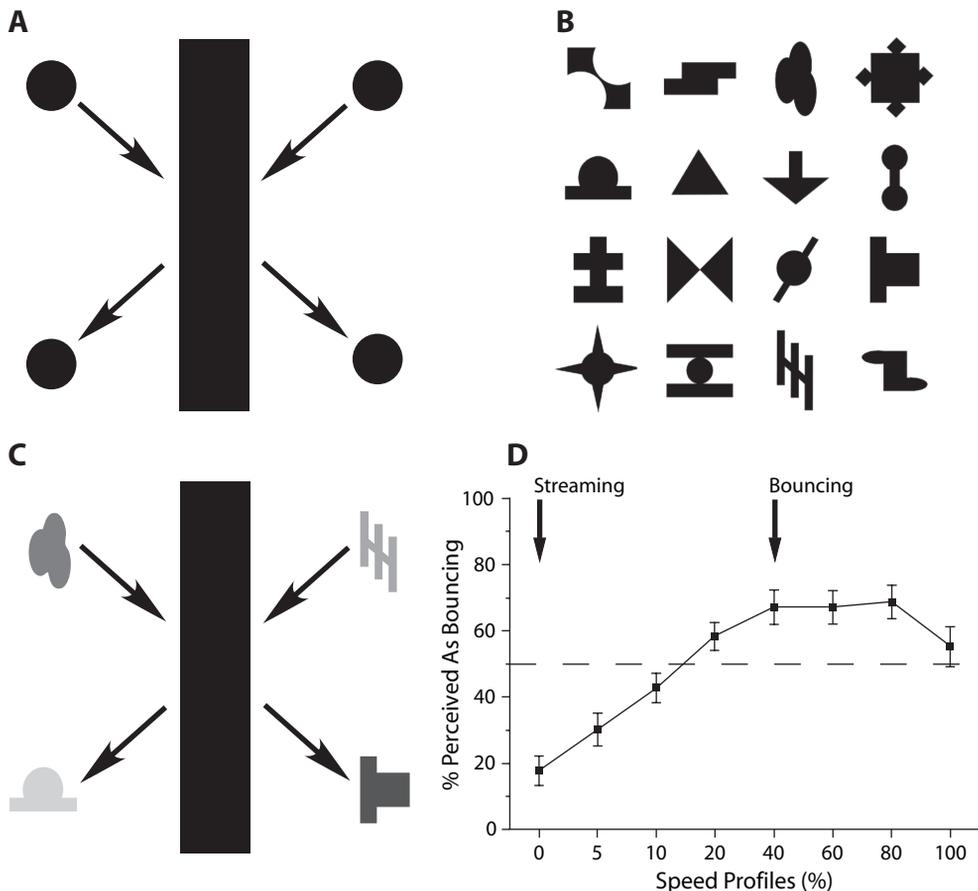


Figure 1. Variants of the streaming–bouncing paradigm. (A) The classic bistable streaming–bouncing display. Two identical shapes appear in the upper corners of a square display and then both move diagonally downward with equal speed and simultaneously disappear behind a central occluder. The two shapes then reappear from behind the occluder and continue their straight motion trajectories along the two diagonal paths. This is an ambiguous display that can be interpreted either as two shapes streaming past each other behind the occluder on a straight trajectory or as two shapes colliding behind the occluder, abruptly changing their paths, and then reappearing on the same side of the occluder whence they originated. Arrows are used to demonstrate the motion of the shapes; they were not present in the display. (B) Sixteen arbitrarily generated shapes used in the experiments (each with a unique color, but shown in panel C as shades of gray). (C) A typical display used in Experiment 1. (D) The results of Experiment 1, showing the percentage of trials that were perceived as bouncing for each speed profile. The dashed line at 50% represents equal likelihood of perceiving the speed pattern as streaming or bouncing. Arrows show the two speed profiles selected for the streaming and bouncing conditions in Experiment 2. Error bars indicate standard errors of the mean.

(0%) streaming speed profile or the nonuniform (40% deceleration–acceleration) bouncing speed profile. There were eight different displays constructed from the 16 shapes, four with uniform and four with nonuniform speed profiles (Figure 2B). As in Experiment 1, each participant was assigned an individual color–shape combination.

The test phase consisted of 24 trials in which two displays were presented in succession in a two-interval forced choice (2IFC) task; participants simply selected the first or the second display as more familiar according to what they had observed during the familiarization phase. In contrast to the 1-to-2 shape displays used during familiarization, both of the 2IFC displays in the test phase were 1-to-1 shape displays. That is, a single initial object moved behind the occluder (just as in the familiarization phase), after which only one object emerged, following either a streaming or a bouncing trajectory. Both of the 1-to-1 displays in a given 2IFC test trial had the same speed profile (0% or 40%).

Eight of the 24 test trials provided the critical comparison of interest; that is, they required the participant to choose between two 1-to-1 test displays that came from the same 1-to-2 familiarization display. Both of these test displays initially presented the same shape traveling diagonally downward and then moving behind the occluder. The two displays differed in the shape that emerged from behind the occluder and the direction in which this second shape moved. In one case the second shape moved on the same trajectory as that of the initial shape, and in the other case the second shape moved on a trajectory at a 90° angle from the trajectory of the initial shape. Importantly, although both of these postocclusion trajectories appeared the same number of times, only one was consistent with the perceptual bias created by the speed profile of the objects presented during familiarization. For example, if in the first test display, the initial shape had a 0% speed profile that supported a streaming percept, then the shape that emerged on the opposite side

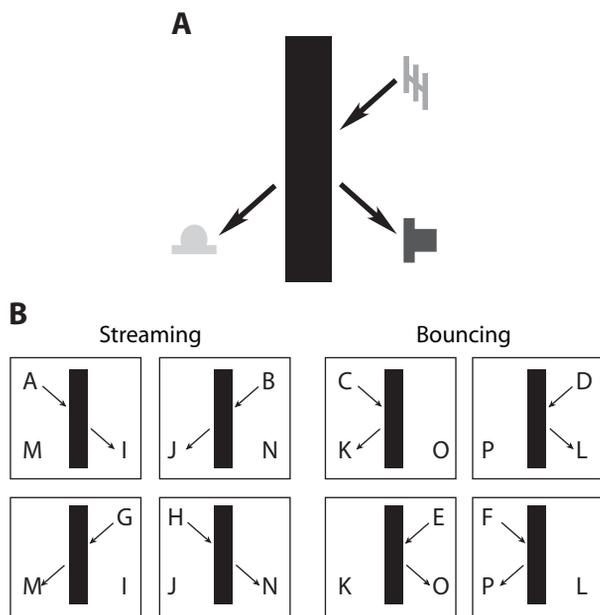


Figure 2. (A) A typical display of Experiment 2. The upper shape could start from either the right or the left upper corner. (B) The eight different displays repeated during the familiarization phase in Experiment 2. Letters stand for shapes that were individually selected for each participant. Two collinear arrows represent uniform speed, which creates a streaming percept, whereas two angled arrows on the same side of the occluder represent a decelerating–accelerating pattern, which induces a bouncing percept. Thus the two shapes connected by arrows in each display form the shape pairs linked perceptually by the motion bias.

of the occluder (along the same trajectory) was consistent with this perceptual bias (e.g., A to I or H to N in Figure 2B). In contrast, in the second test display, the shape that emerged on the same side of the occluder (following a 90° change in the trajectory) was inconsistent with this streaming perceptual bias (e.g., A to M or H to J in Figure 2B). For the 40% deceleration–acceleration bouncing speed profile, the relationship was reversed: The shape that emerged from behind the occluder on the same side on which the initial shape appeared was consistent with the bouncing perceptual bias, whereas the shape that emerged from behind the occluder on the side opposite that on which the initial shape appeared was inconsistent with this bias (e.g., in Figure 2B, C to K and F to P are consistent and C to O and F to L are inconsistent). Thus, a significant preference on these 2IFC trials for the test display that was consistent with the perceptual bias could emerge only if the interpretations induced by the speed profiles had biased the statistical learning process during familiarization. We will refer to these critical test trials as *bias-only* test trials because they equated the associative information between the two shape pairs, and only the perceptual bias information could enable the participants to perform above chance.

In addition to these 8 bias-only test trials, we also ran 16 2IFC test trials that differed from the bias-only trials on dimensions other than just the perceptual bias for bouncing or streaming. Like the bias-only test trials, all of these additional 16 test trials contrasted two 1-to-1 shape displays. However, in these trials, one display consisted of an upper and a lower object that had been paired during 1-to-2 familiarization trials (associated shapes; e.g., B to J or C to O), whereas another display consisted of objects that had never been paired in any trial during familiarization (unassociated shapes; e.g., E to M or H to P). Thus, performance on these test trials did not depend crucially on the perceptual bias of bouncing or streaming, but rather

on whether the temporal co-occurrence of shape pairs was learned across the occlusion event. Half of these 16 test trials contrasted unassociated shapes with associated shapes that were consistent with either the bouncing or streaming perceptual bias that was present during familiarization (e.g., B to J), and the other half contrasted unassociated shapes with associated shapes that were inconsistent with the perceptual bias (e.g., C to O). During each 2IFC test trial, the associated shapes in one test display moved in the same direction, with the same speed profile, as they had during the familiarization phase; this constituted a *consistent* trial. The unassociated shapes in the other test display had the same speed profile as the associated shapes, but the direction in which they moved could be either the same or different; this constituted an *inconsistent* trial. Thus, on consistent test trials both the temporal co-occurrence of the shapes and the bias due to the speed profile could be used to correctly select the associated shapes over the unassociated shapes (e.g., B to J vs. E to M). On inconsistent test trials, only the temporal co-occurrence of the shapes, not the speed profile, could be used to correctly select the associated shapes over the unassociated shapes (e.g., C to O vs. E to P).

Procedure. There were 24 repetitions of each display, presented in random order, for a total of 192 trials in the familiarization phase. Participants were asked to pay attention to the displays but otherwise had no particular task to perform. After the familiarization phase, participants completed a 2IFC test with 24 test trials (8 each of bias-only, consistent, and inconsistent) in an individually randomized order.

Results and Discussion

The results of Experiment 2 are shown in Figure 3. For consistent and inconsistent test trials, participants chose the associated pairs as familiar displays significantly more often than they chose the unassociated pairs [$t(35) = 10.58, p = .0001$, and $t(35) = 9.63, p = .0001$, for the consistent and the inconsistent trials, respectively]. These results support the previous finding that human adults are sensitive to the temporal co-occurrence of shapes, and they demonstrate that a passive inspection of such displays (at least when they are attended; see Turk-Browne, Jungé, & Scholl, 2005) is sufficient for statistical learning of shape associations (Fiser & Aslin, 2002). In addition, there was a marginally significant difference in performance between the consistent and the inconsistent trials [$t(35) = 1.96, p = .058$]. This suggests that altering the speed profile of the associated shapes in the test displays negatively affected recognition performance, presumably because the internal representations of the learned shape associations

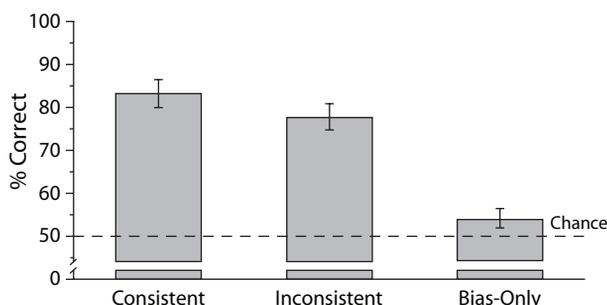


Figure 3. The results of Experiment 2. Participants preferred the perceptually linked pairs significantly more often in each of the trial types. Error bars indicate standard errors of the mean.

incorporated not only the temporal co-occurrences but also the speed profiles of the shapes.

More importantly, on the bias-only test trials, participants chose the shape pairs consistent with the perceptual bias of bouncing or streaming over the shape pairs inconsistent with the perceptual bias [$t(35) = 2.17, p = .037$]. The significant above-chance performance on these test trials cannot be attributed to the temporal succession of shapes or to familiarity with the speed profile, both of which were perfectly and equally consistent for both test displays. Rather, participants' preference for the perceptually biased shape pairs can be attributed only to the stronger link they perceived between those two shapes due to the speed profiles during the familiarization phase, in comparison with the link between the two shapes inconsistent with the bias.

Mitroff, Scholl, and Wynn (2005) also studied bouncing versus streaming displays, measuring both conscious percepts and lower level representations of object persistence by object-specific preview benefits (OSPBs) in Kahneman, Treisman, and Gibbs's (1992) object-reviewing RT paradigm. Mitroff et al. found that the two responses to the correspondence problem posed by such displays ("which went where?") could diverge: In some cases, object files bounced even while conscious percepts streamed. As shown in Experiment 2, visual statistical learning appears to employ the solution to the correspondence problem that accords with our conscious percepts, rather than the solution that exists on the level of object files. This pattern of results suggests that the input to the visual statistical learning mechanism is fairly high level and is not merely the product of simple exposure to low level feature sequences.

EXPERIMENT 3

Oculomotor Behavior During Learning of Bouncing Versus Streaming Displays

The results of Experiment 2 provide clear evidence that the effects of spatiotemporal biases in motion perception of an object undergoing occlusion can influence temporal statistical learning. What remains unclear is whether this bias in learning is related to specific eye movement patterns. For example, if the bouncing or streaming motion bias elicits tracking eye movements, they could potentially (but not necessarily) enable participants to more effectively extract the temporal co-occurrences of shapes that conform to this motion bias. In Experiment 3, we investigated eye movement patterns during the familiarization phase to determine whether gaze patterns could mediate the bias in statistical learning observed in Experiment 2.

Method

Participants, Stimuli, and Design. A new group of 36 undergraduates, naive to the experiment, participated. The experimental stimuli and design were identical to those used in Experiment 2, except that the stimuli were presented on a 42-in. plasma TV screen (Panasonic Model TH-42PHD5), and a remote eyetracking system (ASL Pan/Tilt Model 504) was used to assess the participants' eye movements. The diagonal axis of the effective display area on the

plasma screen was about 1.5 times larger than that of the 21-in. monitor. Thus the maximum width and height of the shapes were 3° , and they moved over greater distances to provide better resolution for detecting eye movements. The eyetracker consisted of a small infrared camera that captured an image of the eye and located the pupil and corneal reflection. This camera was mounted on servomotors controlled by both an optical tracking algorithm and an external magnetic headtracker (a Polhemus FASTRAK operating at 120 Hz), which located the eye when optical tracking failed.

Procedure. In addition to providing the same instructions as in Experiment 2, the experimenter manually calibrated the eyetracker camera and the headtracker at the beginning of the familiarization phase. The participants' eye movements were recorded only during the familiarization phase of the experiment; the test phase was exactly the same as in Experiment 2.

Results and Discussion

An analysis of the performance on the Experiment 3 test trials was completed using the same method as in Experiment 2, and it provided very similar results. The mean percent correct in selecting the associated shapes over the unassociated shapes was significantly above chance for both consistent trials [$M = 74.6\%$; $t(35) = 6.64, p = .0001$] and inconsistent trials [$M = 67.3\%$; $t(35) = 5.95, p = .0001$], and the difference between these groups was also reliable [$t(35) = 2.53, p = .016$]. More importantly, the key finding from Experiment 2 was replicated: On the bias-only test trials, participants chose the shape pairs consistent with the perceptual bias of bouncing or streaming over the shape pairs inconsistent with the perceptual bias [$M = 54.5\%$; $t(35) = 2.07, p = .046$].

To examine the patterns of eye movements during the familiarization phase, we selected only those participants who performed above chance on the bias-only test trials because they provided the strongest test of the hypothesis that the perceptual bias elicited visual tracking of the shape pairs. Fourteen participants performed above chance on the bias-only test trials in Experiment 3, but three of these did not provide appropriate eye movement data. The remaining 11 participants had a mean score of 68.2% correct in choosing the shape pairs consistent with the perceptual bias of bouncing or streaming [$t(10) = 12.00, p < .0001$].¹

The eye movements of these same 11 participants were analyzed individually for each trial to assess the first fixation that occurred in the lower half of the screen after the shape from the upper half of the screen disappeared behind the occluder. If the next fixation that fell in one of the two lower quadrants was consistent with the streaming or bouncing path suggested by the speed profile of the disappearing shape, the trial was coded as *congruent*. Otherwise, the trial was coded as *incongruent*. Since gaze did not always follow the reappearing shapes immediately, the coding of the first eye movement was applied regardless of whether the re-emerging shapes were still behind the occluder (anticipatory) or had already moved out from behind it (visual tracking). If the next fixation after occlusion, before or after the shapes in the lower half had appeared, was not directed to either of the lower quadrants, the trial was discounted. According to these criteria, about 40% of all trials could be considered as congruent or incongruent ($M = 77.4$ out of 192 trials; $SEM =$

1.79). However, the mean numbers of congruent and incongruent trials did not differ significantly [39.2 vs. 38.2; $t(10) = 0.446, p = .665$]. Thus the significant preference for choosing shape pairs consistent with the perceptual bias of bouncing or streaming is not simply a function of where participants are looking.

GENERAL DISCUSSION

In three experiments, we investigated the relationship between visual statistical learning and perceived object trajectories during occlusion and found a strong bias for perceived motion trajectories on the cross-occlusion associations that emerged from the learning process. The significance of these findings is that in the natural environment, an enormously large number of visual features is potentially available as input into a statistical learning mechanism. For this statistical learning mechanism to be computationally tractable, there must be powerful constraints on the input; one such constraint identified in the present series of experiments is the bias to perceive objects as moving along specific trajectories during an occlusion event. Our results complement results from researchers who to date have identified three other constraints on statistical learning: temporal and spatial adjacency (Baker, Olson, & Behrmann, 2004; Newport & Aslin, 2004), perceptual similarity (Creel, Newport, & Aslin, 2004; Turk-Browne, Jungé, & Scholl, 2005), and spatial embeddedness (Fiser & Aslin, 2005).

The present results not only expand this list to include dynamic constraints that govern which objects are likely to be linked over space and time in multiobject natural scenes, but also provide the first evidence for how such constraints can be manifested in a progressively more complex manner. Motion has been indicated previously as a dimension that helps to organize perception, even in infancy, by virtue of the common fate of elements in a scene (see Kellman, 1993, for a review). In the present study, we have demonstrated that more sophisticated anticipatory aspects of motion (such as where an object might move after being occluded) can influence statistical learning. Computationally, our findings fit naturally within a Bayesian framework (Feldman & Tremoulet, 2006; Kersten, Mamassian, & Yuille, 2004) in which perceptual biases of different complexity are incorporated as probabilistic priors constraining the statistical learning process (Fiser & Aslin, 2001, 2005). The integration of basic statistical learning abilities and perceptual constraints into this framework may provide a unified treatment of how learners develop sophisticated internal representations of their natural environments.

AUTHOR NOTE

This research was supported by NIH Research Grant HD037082 to R.N.A., NIH Training Grant T32 EY07125-13 to the University of Rochester, NSF Research Grant 0132444 to B.J.S., and research grant 2001-17783 from the Packard Foundation to R.N.A. We are grateful

to Julie Markant for testing the participants. Correspondence should be addressed to J. Fiser, Volen Center for Complex Systems, Volen 208 MS 013, Waltham, MA 02454 (e-mail: fiser@brandeis.edu).

REFERENCES

- BAKER, C. I., OLSON, C. R., & BEHRMANN, M. (2004). Role of attention and perceptual grouping in visual statistical learning. *Psychological Science*, *15*, 460-466.
- CREEL, S. C., NEWPORT, E. L., & ASLIN, R. N. (2004). Distant melodies: Statistical learning of nonadjacent dependencies in tone sequences. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *30*, 1119-1130.
- FELDMAN, J., & TREMOULET, P. D. (2006). Individuation of visual objects over time. *Cognition*, *99*, 131-165.
- FISER, J., & ASLIN, R. N. (2001). Unsupervised statistical learning of higher-order spatial structures from visual scenes. *Psychological Science*, *12*, 499-504.
- FISER, J., & ASLIN, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, & Cognition*, *28*, 458-467.
- FISER, J., & ASLIN, R. N. (2005). Encoding multielement scenes: Statistical learning of visual feature hierarchies. *Journal of Experimental Psychology: General*, *134*, 521-537.
- HAUSER, M. D., NEWPORT, E. L., & ASLIN, R. N. (2001). Segmentation of the speech stream in a non-human primate: Statistical learning in cotton-top tamarins. *Cognition*, *78*, B53-B64.
- KAHNEMAN, D., TREISMAN, A., & GIBBS, B. J. (1992). The reviewing of object files: Object-specific integration of information. *Cognitive Psychology*, *24*, 175-219.
- KELLMAN, P. J. (1993). Kinematic foundations of infant visual perception. In C. Granrud (Ed.), *Visual perception and cognition in infancy* (pp. 121-173). Hillsdale, NJ: Erlbaum.
- KERSTEN, D., MAMASSIAN, P., & YUILLE, A. (2004). Object perception as Bayesian inference. *Annual Review of Psychology*, *55*, 271-304.
- 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.
- MITROFF, S. R., SCHOLL, B. J., & WYNN, K. (2005). The relationship between object files and conscious perception. *Cognition*, *96*, 67-92.
- NEWPORT, E. L., & ASLIN, R. N. (2004). Learning at a distance I: Statistical learning of non-adjacent dependencies. *Cognitive Psychology*, *48*, 127-162.
- SAFFRAN, J. R., ASLIN, R. N., & NEWPORT, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926-1928.
- SAFFRAN, J. R., NEWPORT, E. L., & ASLIN, R. N. (1996). Word segmentation: The role of distributional cues. *Journal of Memory & Language*, *35*, 606-621.
- SEKULER, A. B., & SEKULER, R. (1999). Collisions between moving visual targets: What controls alternative ways of seeing an ambiguous display? *Perception*, *28*, 415-432.
- SIMONS, D. J., & RENSINK, R. A. (2005). Change blindness: Past, present, and future. *Trends in Cognitive Sciences*, *9*, 16-20.
- 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.

NOTE

1. Performance was 63.6% correct on the consistent test trials and 60.3% correct on the inconsistent test trials, suggesting that this subset of participants was about equally good at all three types of test trials, despite the presence of different cues.

(Manuscript received December 23, 2005;
revision accepted for publication May 25, 2006.)