

# Statistical learning of movement

Joan Danielle Khonghun Ongchoco<sup>1</sup> · Stefan Uddenberg<sup>2</sup> · Marvin M. Chun<sup>2</sup>

© Psychonomic Society, Inc. 2016

**Abstract** The environment is dynamic, but objects move in predictable and characteristic ways, whether they are a dancer in motion, or a bee buzzing around in flight. Sequences of movement are comprised of simpler motion trajectory elements chained together. But how do we know where one trajectory element ends and another begins, much like we parse words from continuous streams of speech? As a novel test of statistical learning, we explored the ability to parse continuous movement sequences into simpler element trajectories. Across four experiments, we showed that people can robustly parse such sequences from a continuous stream of trajectories under increasingly stringent tests of segmentation ability and statistical learning. Observers viewed a single dot as it moved along simple sequences of paths, and were later able to discriminate these sequences from novel and partial ones shown at test. Observers demonstrated this ability when there were potentially helpful trajectory-segmentation cues such as a common origin for all movements (Experiment 1); when the dot's motions were entirely continuous and unconstrained (Experiment 2); when sequences were tested against partial sequences as a more stringent test of statistical learning (Experiment 3); and finally, even when the element trajectories were in fact *pairs* of trajectories, so that abrupt directional changes in the dot's motion could no longer signal

inter-trajectory boundaries (Experiment 4). These results suggest that observers can automatically extract regularities in movement — an ability that may underpin our capacity to learn more complex biological motions, as in sport or dance.

**Keywords** Statistical learning · Visual perception · Motion perception

Picture a dancer in fluid motion. The stream of choreographed action is comprised of simpler movements strung together to convey emotions, intentions or stories. Our environment is filled with such sequences of meaningfully dynamic stimuli, and we must parse their component motions to understand them and respond accordingly. For example, parsing continuous movement is relevant in training for sports (Hossner et al., 2015; Romeas & Faubert, 2015), analyzing gestures (Goldin-Meadow & Beilock, 2010; Roseberry et al., 2011), and ultimately, in facilitating expression and communication (Goldin-Meadow, 2000).

But how do we perceive and segment complex and fluid patterns of movement? While complex movements can be thought of as sequences of simpler actions, it remains an open question how humans segregate dynamic events into meaningful parts. Studying dynamic sequences can be difficult because we lack a vocabulary for the basic units of movement. Here, we ask whether continuously dynamic stimuli can be understood as sequences of element motion trajectories.

We propose that statistical regularities underlie the parsing and learning of meaningful sequences from continuous complex movements. Just as speech is comprised of a continuous stream of words (objects), which in turn are strings of syllables, dynamic motion can be viewed as a continuous stream of meaningful trajectory sequences (objects), which in turn are strings of elemental motion trajectories. Statistical learning is

---

Joan Danielle Khonghun Ongchoco and Stefan Uddenberg contributed equally to this work.

✉ Stefan Uddenberg  
stefan.uddenberg@yale.edu

<sup>1</sup> Division of Social Sciences, Yale-NUS College, Singapore, Singapore

<sup>2</sup> Department of Psychology, Yale University, Box 208205, New Haven, CT 06520-8205, USA

proposed to be the mechanism by which infants (Saffran, Aslin, & Newport, 1996) and adults (Saffran, Johnson, Aslin, & Newport, 1996) can segment words within continuous speech. Speech segmentation is difficult because there are no acoustic cues or temporal gaps that can reliably signal where one word ends and another begins (Cole & Jakimik, 1980). Despite this lack of acoustically invariant cues to word boundaries, words can be parsed based on the transitional probabilities between speech sounds (i.e., the conditional probability that one sound follows another). For example, in a seminal study, 8-month-old infants were presented with a sample speech stream made up of a seemingly random sequence of syllables (e.g., “bi-da-ku-pa-do-ti-go-la-bu-bi-da-ku- ...”). This sequence was actually not entirely random, but contained “words” made up of triplets of syllables repeated throughout the stream (namely “bidaku,” “padoti,” and “golabu”). After only 2 min of exposure to such a stream, infants were able to significantly discriminate the “words” embedded within the sequence from “non-words” (i.e., other triplets of syllables they had never heard before), as well as “part-words” (i.e., triplets that combined syllables from different words). This is because the transitional probabilities of syllables within words (e.g., “bi” to “da”) were higher than the probabilities of those between words (e.g., “ku” to “pa”) (Saffran, Aslin, et al., 1996).

Further work suggests that statistical learning may be a domain-general learning mechanism (Kirkham et al., 2002), capable of operating under different contexts and in other sense modalities. Statistical learning paradigms have demonstrated that we can automatically parse visual sequences of shapes to extract temporal regularities in the world (Fiser & Aslin, 2002). In such visual statistical learning (VSL) paradigms, speech sounds (such as “pa” or “bo”) played in a continuous stream are substituted for simple shapes (such as squares or crosses).

While VSL paradigms convincingly demonstrate that humans can perform statistical learning in the visual domain, such sequences of discrete shape stimuli are not direct analogues to continuous speech. In looking at relationships between shapes across time, each shape is clearly separate from the other — that is, the next shape does not arise from a smooth morphing of the previous shape, but enters as a new object right after another disappears from view. This deviates from continuous speech paradigms that are often devoid of pauses or other segmentation cues between syllables and words.

Therefore, a more direct visual analogue of continuous speech would be a continuously moving object with complex trajectories. While other studies have tried to study continuous biological motion with complex hand gestures (Roseberry et al., 2011), dance sequences (Opacic et al., 2009), or other dynamic action sequences (Baldwin et al., 2008; Meyer & Baldwin, 2011), they used biologically significant stimuli with which people already have a great deal of experience. Here,

we distill the statistical learning and segmentation of dynamic motion input down to its maximally sparse form: a moving dot. We ask if participants can implicitly learn “word” sequences of moving dot trajectories.

## Experiment 1

In this first experiment, we developed a basic motion trajectory “alphabet” as input for a movement statistical learning paradigm. Our procedures then followed Fiser and Aslin’s (2002) seminal visual statistical learning paradigm as closely as possible. This first experiment tested if people could learn triplets of trajectories just as they can learn triplets of shapes (Fiser & Aslin, 2002).

## Method

### *Participants*

Twelve naïve participants (a number chosen to be in line with past statistical learning studies, such as Fiser & Aslin, 2002) were recruited using Amazon’s Mechanical Turk (MTurk) online labor market (for discussion of this pool’s nature and reliability, see Crump et al., 2013). Each participant took part in a single session lasting approximately 26 min on average. On completion of the task, they were given a small monetary reward of US\$1.50, and were excluded from repeat participation in this (or any related) study.

### *Apparatus*

The experiment was conducted using the Qualtrics online survey platform (<http://www.qualtrics.com>) and custom software written using PHP, Javascript, CSS, and HTML. Worker ID screening was implemented via TurkGate (Goldin & Darlow, 2013) to ensure that individuals could not participate more than once or in any related studies. Because our experiment required the use of a relatively large display and the viewing of embedded videos, mobile devices such as phones and tablet computers were disallowed, and attempts to access the experiment from such a device led to its immediate termination along with an error message. Before the experiment began, each participant also completed a simple browser compatibility check to ensure that their system was able to process the required animations and response measures, and only participants who passed that check were allowed to continue.

### *Stimuli*

Animations were presented as embedded videos in .mp4 format near the center of the display (800 px by 600 px) at a rate of 60 frames/s. Individual frames of the animations were

created using MATLAB and the Psychophysics Toolbox (Brainard, 1997), and then compiled into movies using Adobe Photoshop. Each animation was presented against a uniform black background, and consisted of a sequence of motion trajectories performed by white discs presented one at a time. Animations always began with a white disc (20 px diameter) which appeared at the center of the video display for one frame, then moved continuously against a uniform black background along a given motion trajectory at a speed of 4 px/frame for 1 s (or 60 frames). The moving disc always moved away from and back toward the center along the same path in both directions. This way, the disc never disappeared and always began any given trajectory from the center. Thus, the resultant percept was of a single object continuously moving the entire time.

Each disc's motion was constrained to an "alphabet," or repertoire, of 12 different 1-s-long motion trajectories. The basic alphabet is composed of directional trajectories, moving in a linear path for the cardinal directions of north, south, east, and west, and a curvilinear path for the inter-cardinal (or secondary) directions of northeast, northwest, southeast, and southwest, giving us a total of 12 trajectories (see Fig. 1a). To create motion sequences for our animations, the 12 motion trajectories were first randomly grouped into four "base" triplets—that is, sets of three distinct trajectories which always appeared one after the other in the same order, as shown in Fig. 1b. Another four sets of "impossible" triplets—that is, sets of three distinct trajectories that were never shown to participants in sequence during the experiment's "learning" animations, which are described in more detail below—were also generated. These two kinds of triplets were then used to create two categories of animations—learning animations and test animations.

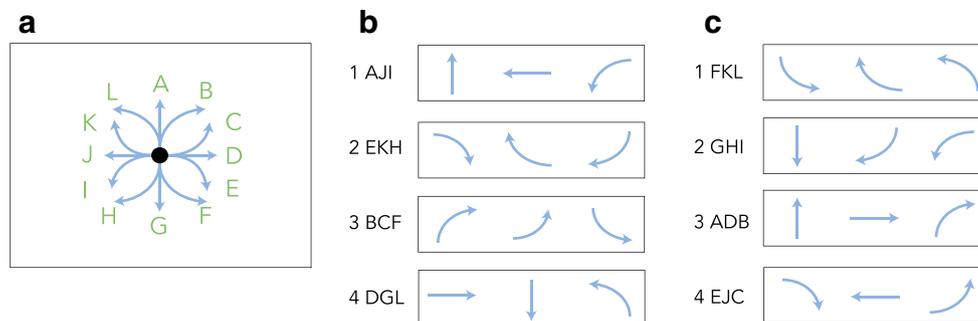
**Learning animations** In learning animations, the base triplets were chained together to create a pseudo-random 96-triplet sequence, with the only constraints being that (1) no immediate repetitions of a triplet were allowed (e.g., ABCABC), and

(2) no immediate repetitions of a *pair* of triplets were allowed (e.g., ABCDEFABCDEF). This semi-random generation of the motion sequences ensured that all motion trajectories and all base triplets appeared an equal number of times in the learning sequences. Consequently, the joint probability of any given base triplet was .083, and the joint probability of any sequence of three element trajectories spanning triplets was .027, mirroring the probability of shape and shape triplet presentations in Fiser and Aslin (2002). A total of three such learning sequence animations — each lasting 9 min 46 s, and which always began with a single frame showing a white fixation cross — were generated and randomly assigned to participants.

**Test animations** In test animations, each of the four base triplets was paired with each of the four impossible triplets in two different orders to yield 32 test pairs which were presented in a forced-choice task phase of the experiment. In each test pair video, the number "1" first appeared at the center of the video frame for 1 s, followed by the first triplet, then the number "2" appeared at the center of the video for 1 s, followed by the second triplet, and finally, a blank background.

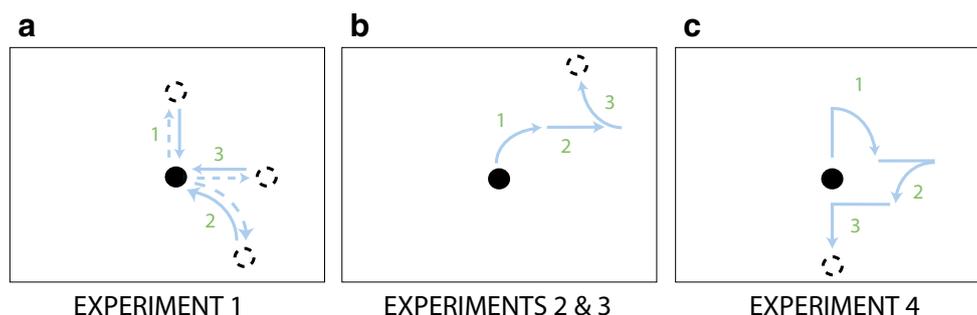
### Procedure

Participants were shown one of the three  $\approx$ 10-min learning sequence animations (randomly assigned across participants) and were instructed to view the movie such that they would be able to answer questions about what they saw. No other instructions were included that might hint at the predetermined patterns of the trajectories. See Fig. 2 for depictions and links to the learning animations. After watching the movie, participants then completed a series of two alternative forced-choice questions comparing 32 pairs of trajectory sequences, each pair containing one base triplet and one impossible triplet in some random order. Participants were instructed to judge which of the two triplets was more familiar based on the learning sequence movie they had watched, choosing either



**Fig. 1** **a** A depiction of the motion trajectory "alphabet" created for our statistical learning paradigm. This alphabet was used across all experiments, and the disc's motion was constrained to follow some sequence of the linear and curvilinear paths shown. **b** A sample set of four base triplets shown during the learning phase of the experiment.

Each base triplet is composed of three different trajectories from the motion alphabet, chosen at random. **c** A sample set of four "impossible" triplets — random motion sequences which were never shown to participants during the learning stage. These are again composed solely of trajectories from the motion alphabet



**Fig. 2** A depiction of the stimulus' motion in Experiments 1–4. **a** In Experiment 1, the disc first traveled away from, and then back toward the center of the video frame for each trajectory in the sequence, producing a percept of continuous motion (<http://camplab.psych.yale.edu/demos/Exp1.mov>); **b** In Experiments 2 and 3, the disc traveled around the screen without being forced to return to a given point

onscreen at any time (<http://camplab.psych.yale.edu/demos/Exp2and3.mp4>); **c** In Experiment 4, the disc traveled around the screen like in Experiments 2 and 3, but this time, with abrupt breaks in every letter of the triplet (<http://camplab.psych.yale.edu/demos/Exp4.mp4>). Mean performance in each experiment is depicted below the relevant panel; all were significantly above chance

the first or the second triplet displayed by means of radio buttons presented in a different random order below the video for each trial. No time limit was imposed for responses.

## Results

Twelve participants successfully discriminated base triplets from impossible triplets ( $M = 68.75\%$ ,  $SD = 16.43\%$ ),  $t(11) = 3.95$ ,  $p = .002$ ,  $d = 1.14$ . Thus, participants were capable of parsing patterns from a continuous stream when tracking a single constantly moving object with a fixed trajectory origin. Because the movement here was continuous without any perceptual breaks, these results demonstrate participants' ability to recognize patterns from a continuous stream that is more analogous to continuous speech.

## Experiments 2a and 2b

Experiment 1 tested movement sequences with trajectories that were constrained to always begin from the center of the screen. However, in the real world, moving objects are typically not constrained to return to a fixed origin. More critically here, the common center point for all movements in the previous experiment may have served as a perceptual cue for the boundaries of element trajectories (although not sufficient to parse triplet trajectories, dependent on statistical learning). In order to test the limits of statistical learning of continuous movement, we removed this constraint and allowed the single moving object to roam around the screen.

## Method

The design and procedure were identical to that used in Experiment 1 except where noted. Here, each participant took part in a single session lasting 19 min on average, and compared to the previous experiments, they earned a slightly

larger compensation of US\$2.14. The same primary and secondary directional motions from the last two experiments were used, except that here, a blue disc (rather than a white disc) now appeared at the center of the screen for one frame, and subsequently roamed around a textured background generated from random visual white noise. Trajectories were unconstrained in their starting position, and as such, the disc did not return to the screen's center between trajectories, instead moving continuously around the frame.

To ensure that the disc would remain visible and never leave the screen despite its pseudo-random continuous motion, the disc travelled at a slower speed of 3 px/frame, while each element trajectory now lasted for 30 frames, or 500 ms. Learning animations lasted a total of 2 min 24 s. Whenever the disc's next position would come close to the border of the display (i.e., within 50 px of the screen's edge), the frame was adjusted to follow the disc, so that the disc always remained visible as it continued to move. The textured background helped give the impression of a camera tracking the object (this impression can be confirmed via visual inspection of the following video, also referenced in Fig. 2c: <http://camplab.psych.yale.edu/demos/Exp2and3.mp4>). This "tracking" procedure entailed adjusting the disc's position at a speed of 3 px/frame back toward the center of the video frame while the textured background was adjusted in the opposite direction that disc moved, again at the same speed (3 px/frame). In the test animations, the first test triplet would be displayed via a disc starting from the center of the screen, followed by a 1-s blank screen inter-stimulus interval, and finally the second test triplet would be displayed with another disc starting from the center of the screen again.

We performed two versions of this experiment — Experiments 2a and 2b — due to a forced "update" in the way Qualtrics handled the custom Javascript used to control the videos. Whereas in the other experiments Qualtrics would load the custom Javascript for video playback on a page-by-page basis, under this new scheme all the Javascript for the

entire experiment was loaded once at the outset. As a result, this “update” caused problems with test pair presentation for a subset of participants, as revealed by examining participants’ debriefing questionnaires.

In response to these problems, we developed a second version of the experiment (i.e., Experiment 2b) in which we simply added a third option for each of the questions in the test phase of the experiment: “The video did not play properly.” This allowed participants to report any problems with video playback for each video shown, and allowed us to further ensure that all test pairs properly played for all 12 participants to be included in the final analyses for each experiment. In Experiment 2a, two out of 14 participants were excluded, while in Experiment 2b, five out of 17 participants were excluded because of one or more self-reported issues with video playback.

## Results

In Experiment 2a, participants successfully discriminated base triplets from impossible triplets ( $M = 57.03\%$ ,  $SD = 10.67\%$ ),  $t(11) = 2.28$ ,  $p = .043$ ,  $d = .66$ . We also obtained similar results in Experiment 2b where participants successfully discriminated base triplets ( $M = 59.11\%$ ,  $SD = 11.50\%$ ),  $t(11) = 2.75$ ,  $p = .019$ ,  $d = .79$ , and so we opted to combine the data from the two independently samples for further analysis. The combined sample (24 participants) again showed successful discrimination of base triplets from impossible triplets ( $M = 58.07\%$ ,  $SD = 10.90\%$ ),  $t(23) = 3.63$ ,  $p = .001$ ,  $d = .74$ . These results demonstrate participants’ ability to detect implicit patterns even when the object is engaged in truly continuous movement.

Mean performance in this freely moving object paradigm was significantly lower than the mean performance in Experiment 1, where the object moved continuously while constrained to return to the center (58.07 % vs. 68.75 %;  $t(34) = 2.33$ ,  $p = .026$ ,  $d = .80$ ).

## Experiment 3

To provide an even more stringent test of the statistical learning of movement demonstrated above, we tested recognition for base triplets versus “part-base” triplets — “accidental” combinations of trajectories that span the boundary between two base triplets. Such “part-base” triplets were made up of the final trajectory of one base triplet, and the first two trajectories of a second base triplet. Unlike “impossible” triplets, these “part-base” triplets were in fact seen occasionally over the course of the video, but they violate the fully intact structure that distinguishes base triplets from one another. This is therefore a more robust test of statistical learning, because successful performance requires sensitivity to the full underlying

structure of the learned triplets (Baldwin et al., 2008; Saffran, Aslin, et al., 1996).

## Method

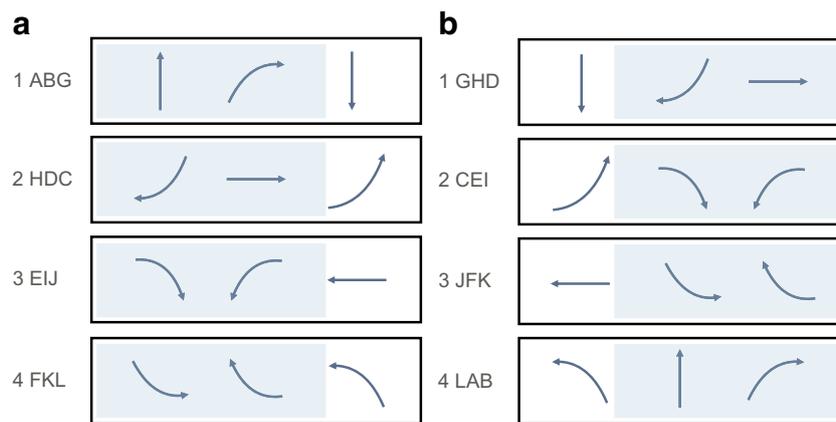
The design and procedure were identical to that used in Experiment 2b, except as noted below. Each participant once again participated in a single session lasting 12.5 min on average, and compared to the previous experiments, they were compensated US\$2.85. However, in the test phase, instead of impossible triplets, base triplets in the learning phase were paired with part-base triplets. Following the procedures used in previous statistical learning experiments (e.g., Baldwin et al., 2008; Saffran, Aslin, et al., 1996), these part-base triplets were formed by combining the final letter of a given base triplet with the first two letters of a second base triplet. In this way, four unique part-base triplets were constructed, in which (1) each base triplet’s final trajectory was represented once, and (2) each base triplet’s final two trajectories were represented once. See Fig. 3 for a representation of the base and part-base triplets from this experiment.

## Results

Participants successfully discriminated base triplets from part-base triplets ( $M = 55.08\%$ ,  $SD = 10.34\%$ ),  $t(23) = 2.41$ ,  $p = .025$ ,  $d = .49$ . This experiment confirms that participants extracted the full statistical structure of the base triplets shown throughout the learning animation. Mean performance in this experiment was not significantly different from that in Experiment 2 (55.08 % vs. 58.07 %;  $t(46) = .98$ ,  $p = .334$ ,  $d = .28$ ) but was once again significantly lower than mean performance in Experiment 1 (55.08 % vs. 68.75 %;  $t(34) = 3.06$ ,  $p = .004$ ,  $d = 1.05$ ).

## Experiment 4

In prior experiments, trajectory letters may have been parsed or individuated via a correlated cue — the low-level changes in velocity that emerged between individual motions (i.e., motion breaks, or angular discontinuities). It is worth noting that such correlated cues are also present in natural speech (e.g., transitions between syllables are correlated with changes in prosody). Although not sufficient to parse triplets, which require statistical learning, the discontinuities between elements may have aided the learning of sequences. Thus, to further test the robustness of the statistical learning of continuous movement, Experiment 4 disrupted letter segmentation by introducing motion breaks within the letters themselves. We redesigned our base triplets so that each “letter” here was in fact a *pair* of trajectories from the movement alphabet. This allowed us to generate letters that have motion discontinuities,



**Fig. 3** **a** A sample set of four base triplets shown during the learning phase of the experiment. Each base triplet is composed of three different trajectories from the motion alphabet, chosen at random. **b** A sample set of four “part-base” triplets which were created by combining the final

letter of a base triplet (shown here with a white background) with the first two letters of another base triplet (shown here with a light-blue/light-grey background)

yielding animations that increased the demands on statistical learning to not only parse triplets, but the letters themselves (see Hunt & Aslin, 2001).

## Method

The design and procedure were identical to that used in Experiment 2b, except as noted below. The single experimental session now lasted 18.4 min on average. Triplets were made up of motion letters that had discontinuities within them, designed by putting two motion trajectories together with the following constraints: (1) no letter should appear as a continuous, smooth motion (e.g., two trajectories that form a semi-circle cannot go together); (2) no trajectory should be repeated within a letter; and (3) no trajectory should be followed by its opposite, reverse trajectory (i.e., the second trajectory must not follow the same path of the first in reverse; although such trajectories have a motion break, the reversal was visually salient). Hence, instead of triplets being composed of one trajectory per letter (e.g., A-B-J), each triplet was now composed of two trajectories per letter (e.g., EA-BL-CJ). To equate the probability of any given element trajectory appearing, each such trajectory only appeared twice across the four bases. Each element trajectory was still 500 ms long, making each letter 1 s in duration.

## Results

Participants successfully discriminated base triplets with abrupt breaks within the letters themselves, ( $M = 54.04\%$ ,  $SD = 9.05\%$ ),  $t(23) = 2.18$ ,  $p = .039$ ,  $d = .45$ . These results show that people can parse out motion regularities from the continuous learning sequence even when the element motions themselves required statistical learning, since discontinuities in motion were present both within and between letters. Mean

performance in this experiment was significantly lower than the mean performance in Experiment 1 (54.04 % vs. 68.75 %;  $t(34) = 3.48$ ,  $p = .001$ ,  $d = .99$ ; also significant at  $p = .012$  when correcting for unequal variances as revealed by Levene’s test,  $F(1,31) = 4.32$ ,  $p = .045$ ), once again suggesting greater difficulty in parsing out regularities in a freely moving object paradigm versus one where the disc is constrained to the center. However, mean performance was not significantly different from mean performance in Experiment 2 (54.04 % vs. 58.07 %;  $t(46) = 1.40$ ,  $p = .170$ ,  $d = .40$ ) or Experiment 3 (54.04 % vs. 55.08 %;  $t(46) = .37$ ,  $p = .712$ ,  $d = .11$ ).

## General discussion

Using an “alphabet” of basic elements of movement combined into continuously dynamic sequence “objects,” our four experiments newly demonstrated the ability to parse movement sequences using statistical learning. Across all four experiments, which provided increasingly stringent tests of statistical learning and motion segmentation, participants identified embedded trajectory sequences that could only be differentiated based on statistical learning of transitional probabilities.

Just as words can be extracted from continuous speech without temporal gaps to signal perceptual boundaries (Saffran, Aslin, et al., 1996; Saffran, Johnson, et al., 1996), there were no gaps between element trajectories in the experiments, which both made use of a single continuously moving disc. Statistical learning of dynamic motion patterns in these experiments indicates that motion trajectories may serve as perceptual objects (Scholl, 2001), and that such dynamic objects can be learned and parsed from otherwise continuous input. Understanding trajectories as perceptual objects can be related to event (Zacks & Tversky, 2001) and gesture (Goldin-Meadow, 2000) processing, given that they also need to be

parsed from continuous input. Our study presents such dynamic objects without the possible confound of semantics or intentionality that events and gestures tend to have.

Considering motion trajectories as perceptual objects raises several questions surrounding object-based attention selection and capacity. First, one could ask whether some types of motion trajectories are learned and parsed more easily than others. For example, would animacy cues enhance statistical learning and parsing of dynamic sequences?

Statistical learning performance was highest for Experiment 1, but did not differ across Experiments 2, 3, and 4. This suggests that statistical learning of movement is robust across increasing levels of continuity (absent helpful discontinuities) – only improved by the overt segmentation cues in Experiment 1 that involved all trajectories returning to a common origin point. Future parametric work can characterize the perceptual variables that affect the efficacy of trajectory parsing.

Does the manner and efficacy with which we learn and parse dynamic patterns influence the way we perform other more complex activities? Statistical learning of moving objects has implications for event comprehension, memory and problem solving. Studies have shown that event segmentation creates breakpoints for memory encoding and facilitates the recognition of event sequences for future action planning (Kurby & Zacks, 2008). Statistical learning of motion sequences can therefore be crucial in appreciating or training for dance or sports. It can influence the emotions one feels and remembers from watching a dancer, or determine how an athlete can perform a better tennis or golf swing by breaking complicated motion sequences into manageable parts.

In conclusion, the present experiments demonstrate the statistical learning of structures in truly continuous motion patterns. Our results highlight a mechanism by which humans may learn more complex biological motions or event sequences, suggesting a new approach to understanding motion as perceptual objects, and to consider the ways parsing movement might impact how we perceive (and perform in) the continuously dynamic world around us.

**Author Note** JDKO and SU contributed equally to this work. For helpful conversation and/or comments on the project, we thank the members of the Chun Cognitive Neuroscience lab at Yale University. This project was funded by a Yale-NUS College Summer Independent Research Grant awarded to JDKO and an NSF Graduate Research Fellowship awarded to SU.

## References

- Baldwin, D., Andersson, A., Saffran, J., & Meyer, M. (2008). Segmenting dynamic human action in statistical structure. *Cognition*, *106*, 1382–1407.
- Brainard, D. H. (1997). The psychophysics toolbox. *Spatial Vision*, *10*, 433–436.
- Cole, R. A., & Jakimik, J. (1980). A model of speech perception. In R. A. Cole (Ed.), *Perception and production of fluent speech* (pp. 133–136). Hillsdale: Erlbaum.
- Crump, M. J. C., McDonnell, J. V., & Gureckis, T. M. (2013). Evaluating Amazon's Mechanical Turk as a tool for experimental behavioral research. *PLoS ONE*, *8*, e57410.
- Fiser, J., & Aslin, R. N. (2002). Statistical learning of higher-order temporal structure from visual shape sequences. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *28*, 458–467.
- Goldin, G., Darlow, A. (2013). TurkGate (version 0.4.0) [Software]. Available from <http://gideongoldin.github.com/TurkGate/>
- Goldin-Meadow, S. (2000). Beyond words: The importance of gesture to researchers and learners. *Child Development*, *71*, 231–239.
- Goldin-Meadow, S., & Beilock, S. (2010). Action's influence on thought: The case of gesture. *Perspectives on Psychological Science*, *5*, 664–674.
- Hossner, E.-J., Schiebl, F., & Göhner, U. (2015). A functional approach to movement analysis and error identification in sports and physical education. *Frontiers in Psychology*, *6*, 1339.
- 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.
- 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.
- Kurby, C. A., & Zacks, J. M. (2008). Segmentation in the perception and memory of events. *Trends in Cognitive Sciences*, *12*, 72–79.
- Meyer, M., & Baldwin, D. (2011). Statistical learning of action: The role of conditional probability. *Learning and Behavior*, *39*, 383–398.
- Opac, T., Stevens, C., & Tillmann, B. (2009). Unspoken knowledge: Implicit learning of structured human dance movement. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *35*, 1570–1577.
- Romeas, T., & Faubert, J. (2015). Soccer athletes are superior to non-athletes at perceiving soccer-specific and non-sport specific human biological motion. *Frontiers in Psychology*, *6*, 1343.
- Roseberry, S., Richie, R., Hirsh-Pasek, K., Golinkoff, R. M., & Shipley, T. F. (2011). Babies catch a break: 7-to 9-month-olds track statistical probabilities in continuous dynamic events. *Psychological Science*, *22*, 1422–1424.
- Saffran, J. R., Aslin, R. N., & Newport, E. L. (1996). Statistical learning by 8-month-old infants. *Science*, *274*, 1926–1928.
- Saffran, J. R., Johnson, E. K., Aslin, R. N., & Newport, E. L. (1996). Word segmentation: The role of distributional cues. *Journal of Memory and Language*, *35*, 606–621.
- Scholl, B. J. (2001). Objects and attention: The state of the art. *Cognition*, *80*, 1–46.
- Zacks, J. M., & Tversky, B. (2001). Event structure in perception and conception. *Psychological Bulletin*, *127*, 3–21.