Uncovering visual priors in spatial memory using serial reproduction

Thomas A. Langlois* (thomas.langlois@berkeley.edu)
Department of Psychology, University of California, Berkeley
Berkeley, CA 94720 USA

Nori Jacoby* (kj2413@Columbia.edu)
Center for Science and Society, Columbia University
New York, NY 10027 USA

Jordan Suchow (suchow@berkeley.edu)
Thomas L. Griffiths (tom_griffiths@berkeley.edu)
Department of Psychology, University of California, Berkeley
Berkeley, CA 94720 USA

*These authors contributed equally to this work

Abstract

Visual memory can be understood as an inferential process that combines noisy information about the world with knowledge drawn from experience. Biases can arise during encoding of information from the outside world into internal representations, or during retrieval. In this work, we use the method of serial reproduction, in which information is passed along a chain of participants who try to recreate what they observed. We apply this method to the study of visual perception in the context of spatial memory biases for the remembered position of dots inside different geometric shapes. We present the results of non-parametric kernel density estimation of the end result of serial reproduction to model visual biases. We confirm previous findings, and show that memory biases revealed with our method are often more intricate and complex than what had previously been reported, suggesting that serial reproduction can be effective for studying perceptual priors.

Keywords: Vision; spatial memory; inductive biases; serial reproduction; iterated learning;

Introduction

Retrieving detailed visual information from memory requires efficient representations of often complex and noisy visual scenes. In Bayesian accounts of reconstruction from visual memory, the memory system integrates sensory information with knowledge acquired from previous experience ("priors"). Effective use of this information may reduce variability in visual memory and improve overall reconstruction accuracy (Weiss et al., 2002). Using priors is usually advantageous because they capture regularities in the structure of the world that are innate or observed over our lifetimes. However, this can lead to substantial biases during reconstruction. This is because prior information may deviate significantly from our observations, especially when a visual scene is unexpected given previous experience.

In many cases, visual priors are categorical (or prototypical), represented in memory as schematic or simplified objects (Huttenlocher et al., 1991). In one experimental paradigm that reveals categorical effects, participants are asked to remember the location of a dot presented within a circle or other bounding shape. After a brief presentation and a delay, participants reproduce the dot's location by placing it in the recalled position (see Figure 1). Huttenlocher et al. (1991) found that participants tend to misplace dots toward a central (prototypical) location in each of the quadrants of a circle. Following these results, Wedell et al. (2007) tried to predict prototypical positions in spatial memory for dots presented inside a variety of geometric shapes (circle, square, triangle, vertical oval, horizontal oval, and pentagon). In the study, participants were shown thirty-two dots aligned along two concentric circles within each shape. A parametric model with four components (prototypes) was fitted to the remembered positions of the dots, confirming that visual memory of these shapes shows substantial categorical effects.

The approach to characterizing categorical biases used in Huttenlocher et al. (1991) has a number of limitations—specifically, a relatively small number of to-be-remembered locations (32) and a weak measurement of the biases, resulting in limited resolution for capturing the locations of the categories. In addition, Wedell et al. (2007) used a parametric model with a fixed number of categories. The choice of the model, and the number of categories that were used were not fully justified, requiring certain *a priori* assumptions. Here, we propose to use a paradigm based on serial reproduction to characterize visual memory biases without needing to rely on parametric modeling and with substantially better resolution and accuracy.

The method of serial reproduction

Serial reproduction has a long history in experimental psychology, where it has been used to study how various biases distort information when it is transmitted from person to person (Bartlett, 1932). Figure 2 shows a schematic illustration of the experimental paradigm: a participant views a stimulus, such as a dot presented within a bounding shape, and is then asked to reproduce the stimulus as accurately as possible from memory. Critically, the reproduction created by the first participant is used as the stimulus for the second participant, who is then asked to do the same. At each iteration, the reconstruction produced by the previous participant becomes

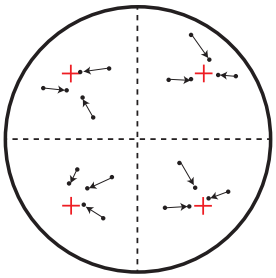

Figure 1: Illustration of prototype effects in memory for points in a circle. The red crosses represent prototypes, and the small points are typically misremembered as being closer to those prototypes.

the stimulus for the next participant to reproduce. Famous early results include the transformation of an owl-like Egyptian hieroglyph into a small cat after ten iterations of a serial-reproduction drawing task (Bartlett, 1932). This result was interpreted in terms of inductive biases in memory: as veridical information from the input becomes degraded following successive iterations, the reconstruction of the ambiguous image is pulled towards a prototypical object with similar visual properties.

Serial-reproduction experiments have long been used to simulate phenomena in cultural transmission, evolutionary biology, anthropology, and cognitive science (Kirby et al., 2008; Claidière et al., 2014), but it wasn't until recently that a rational analysis of serial reproduction considered how information should change as it is transmitted along a chain of rational agents (Xu & Griffiths, 2010). Under a rational analysis, reconstruction from memory is defined as the problem of inferring the most accurate state of the world from noisy data, such as an imperfect memory trace and perceptual noise during encoding of the image. This problem is modeled using the framework of Bayesian statistics. Previous experience is captured by a prior distribution over possible states (a hypothesis space of world states). A posterior is computed, based on the likelihood, which indicates the probability of observing that information, given some hypothesis about the true state of the world. Xu & Griffiths (2010) examined the predictions of this Bayesian account of reconstruction from memory for serial reproduction. They found that serial reproduction by Bayesian agents defines a Markov chain with the following transition probabilities:

$$p(x_{n+1} \mid x_n) = \int p(x_{n+1} \mid \mu) p(\mu \mid x_n) d\mu,$$

where x is a noisy stimulus (such as an imperfect memory trace) and μ is the true state of the world that generated that stimulus (in this case, the veridical image that impinged on the visual system). This Markov chain captures the probability of a new stimulus x_{n+1} being created as a reconstruction of a previously seen stimulus x_n in each iteration in the serial reproduction chain, and has a stationary distribution, called the *prior predictive distribution*, which defines the probability of observing a stimulus x when μ is sampled from the prior:

$$p(x) = \int p(x \mid \mu) p(\mu) d\mu.$$

This process approximates a Gibbs sampler for the joint distribution on x and μ defined by multiplying $p(x \mid \mu)$ and $p(\mu)$. This finding is significant because it provides a mathematical formalism for describing the consequences of serial reproduction: assuming that participants share common inductive biases, the transmission chain will converge to a sample from their shared prior.

In this paper, we explore spatial memory priors in a task where participants were asked to remember the position of a small black dot inside a variety of geometric shapes. Operating under the assumption that people share the same inductive biases, or spatial memory priors, we show that serial reproduction appears to converge on these priors remarkably quickly, revealing patterns that are consistent with some established findings, although in many cases revealing new and intricate patterns that were previously unknown. Finally, we demonstrate the advantages of using a non-parametric kernel density estimation procedure to characterize the prior.

Methods

Participants

Participants were recruited online using Amazon Mechanical Turk. All gave informed consent. The experimental protocol was approved by The Committee for the Protection of Human Subjects (CPHS) at the University of California, Berkeley. Each experiment required approximately 70-100 participants. A total of 570 participants took part in Experiment 1 and an additional 590 took part in Experiment 2.

Stimuli

All images were approximately 400×400 pixels in size. Each shape was a 6-pixel-wide black outline over a white background. The sizes and colors of the backgrounds for the images were intended to ensure that the images would be clearly visible in any standard browser window (unlikely to become occluded), and such that the boundaries of the images would be invisible.

Procedure

We carried out a series of serial reproduction experiments. Participants were presented with timed displays (a shape outline with a dot initialized somewhere within the boundaries of the shape), and were instructed to reproduce the exact location of the dot inside of the shape. Once complete, their response was sent to another worker (again, as a timed display), who was instructed to reconstruct this display from memory, and so on. A total of ten iterations were completed for each chain. See Figure 2 for a schematic diagram of the serial reproduction procedure.

Practice trials. Participants completed ten practice trials in order to become familiar with the user interface. During these practice trials, they were presented with a circle (a black outline of a circle over a white screen), with a dot initialized somewhere within its boundaries. This display was presented for 4000 ms, followed by a blank screen lasting 1000 ms. Next, the circle was presented without the dot and remained on the screen until the participant positioned the dot in the location that they remembered. As soon as the participant clicked, the dot appeared under the mouse cursor. Participants could reposition the dot as many times as they needed. Once done, they pressed a button to proceed to the next trial.

Experimental trials. Following the ten practice trials with the circle, there were ninety-five experimental trials with exactly one of the shapes. In Experiment 1, the shape could be a circle, triangle, square, vertical oval, horizontal oval, or a pentagon. In Experiment 2, the shape could be a regular polygon with more than five vertices. For each of the 95 experimental trials, the presentation time was reduced to 1000 ms. As with the practice trials, the position of the shape on the screen was randomized somewhere inside a larger canvas in order to control for participants resorting to tracking the position of the dot by trivially marking its absolute position on their computer screens. In addition, participants were given trial-by-trial feedback regarding their accuracy. If their responses were within eight percent of the width and height subtended by the shape on the screen, they were told that their response was accurate (a message in green font: "This was accurate"), and received a small monetary bonus. If not, they received no bonus beyond the basic payment for the HIT, in addition to any bonuses accrued from the previous trials, and were presented with a red message ("this was not accurate"). These trials were discarded from the experiment. Participants could not provide multiple responses within a chain.

Experiment 1: We used the same six shapes as Wedell et al. (2007): A circle, equilateral triangle, square, vertical and horizontal ovals, and a regular pentagon. For each shape, we initialized the position of five hundred dots within its boundaries (for the circle, we initialized four hundred dots).

Experiment 2: Because our method revealed a variable number of peaks (prototypes) in the prior for the angular shapes in Experiment 1, and that these appeared to be due in large part to the number of vertices in the polygons (all were regular polygons—an equilateral triangle, square, and pentagon), we wanted to determine the point at which the prototypes begin to merge into the four prototypes in the prior for the circle. We did this by conducting the same experiment with polygons containing increasingly more vertices (approximating a circle more closely as vertices were added).

Results

Our results are presented in two parts. First, we present all our results for Experiment 1, in which we used the same shapes as Wedell et al. (2007). We demonstrate that using a serial reproduction paradigm, as well as non-parametric kernel density estimation, replicates some (but not all) of their key findings. In particular, we find four prototypes arranged in the centers of the four quadrants of the circle, as well as the horizontal and vertical oval shapes, as they did. However, we also show new and intricate patterns in the priors for the angular shapes (triangle, square, and pentagon). We evaluate the predictions of the model by Wedell et al. (2007) on the data we obtained following one iteration, for all the shapes, and compare them to predictions that we obtain from our estimates of the prior following all ten iterations. In addition, we show quantitative evaluations of the change in copying accuracy for the equilateral triangle. Second, we show the results for Experiment 2, where we illustrate the effect of adding vertices to regular polygons on the prior, revealing hitherto unknown grouping effects of the prototypes in spatial memory that occur as regular polygons begin to approximate a circle.

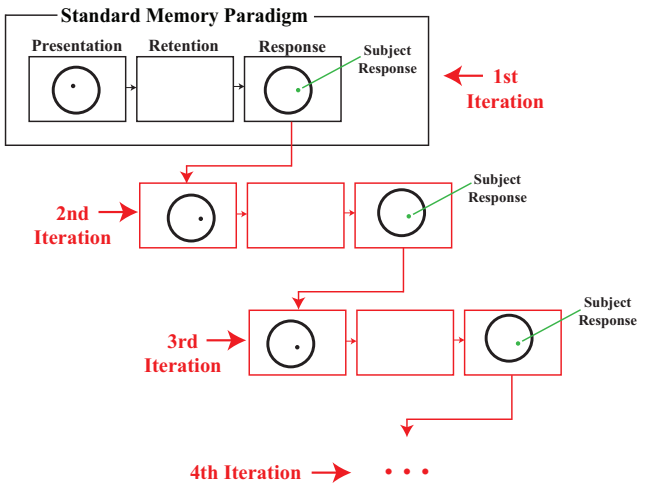

Figure 2: Serial reproduction chain for one trial in the memory task, illustrating the iterative process for a single dot being remembered. The trial in black represents a standard memory paradigm. In red are additional iterations of the task using the result from the previous iteration as the new stimulus, which form the nodes in the serial reproduction process.

Measuring spatial memory priors

Serial reproduction results. Figure 3 shows visualizations of the estimates that we obtained following ten iterations of the serial reproduction experiment using four hundred initial seeds for the circle. Each panel shows the results for each of the ten iterations, including the initial seeds. Notice that the prototypes begin to emerge in as early as the fourth iteration. For the panel showing the results of the tenth iteration, we show an estimate of the prior using our non-parametric kernel density estimate in lieu of plotting the points. Figure 4 shows the results that we obtained for the equilateral triangle. Notice the emergence of bimodal peaks near each of the vertices. This finding suggests that for this shape, there are a total of six prototypes in the prior, grouped in pairs at each corner.

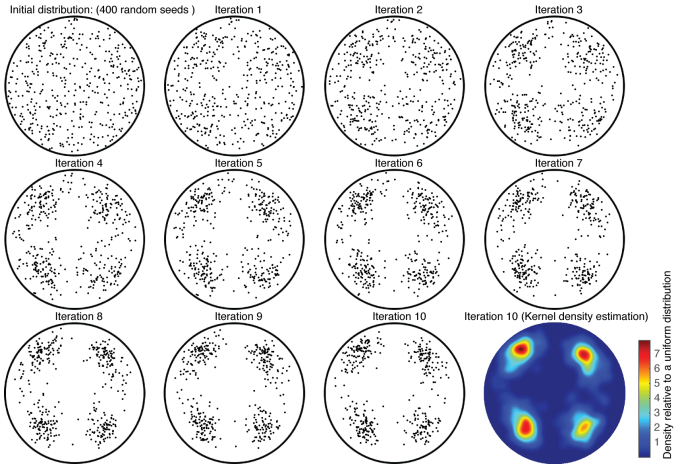

Figure 3: Serial reproduction of 400 dots presented in a circle, for ten generations (iterations) of the process. The top left scatterplot shows the positions of the original seeds (sampled from a uniform distribution) inside the circle shape. The remaining subplots show scatterplots of the results of the serial reproduction chain for iterations 1-10. The subplot of the tenth iteration, in the bottom right, also shows the kernel density estimate. Note that from one iteration to the next, points that were originally scattered uniformally within the circle boundary begin to converge on each of the nearest prototypes at the center of each of the four quadrants in the circle. By the tenth iteration of the process, four clusters are clearly discernable.

Simple shapes. In Figure 5 (panels B and D), we show the kernel density estimates that we obtained for all the shapes. In the case of the circle, vertical oval, and horizontal oval, our results are consistent with past findings (shown in panels A and C). However, we discover bimodal peaks in the vertices of the angular shapes (prototype pairs clustered at each of the corners). This result is particularly striking for the triangle and the square shapes. The same result is present for the pentagon shape, although unlike the peaks in the prior for the triangle and square, those in the pentagon are not quite rotationally invariant, although all three geometric shapes are, suggesting that the shapes and orientations of the modes in the priors are not a simple function of the presence of edges, or the angles at these edges.

Convergence analysis. For the triangle results, we completed a convergence analysis (See Figure 4, panels B and C), using the Jensen-Shannon divergence (JSD). To estimate the variability of these JSDs, we generated 100 bootstrapped data sets sampled from the original data (with replacement). For each one, we computed the JSDs of consecutive iterations (see panel B). The JSD between the initial distribution and iteration 1 was significantly larger than that between the two final iterations (p = 0.02) and there were no significant differences between the distance between iterations 9 and 10 compared with iterations 8 and 9 (p = 0.43).

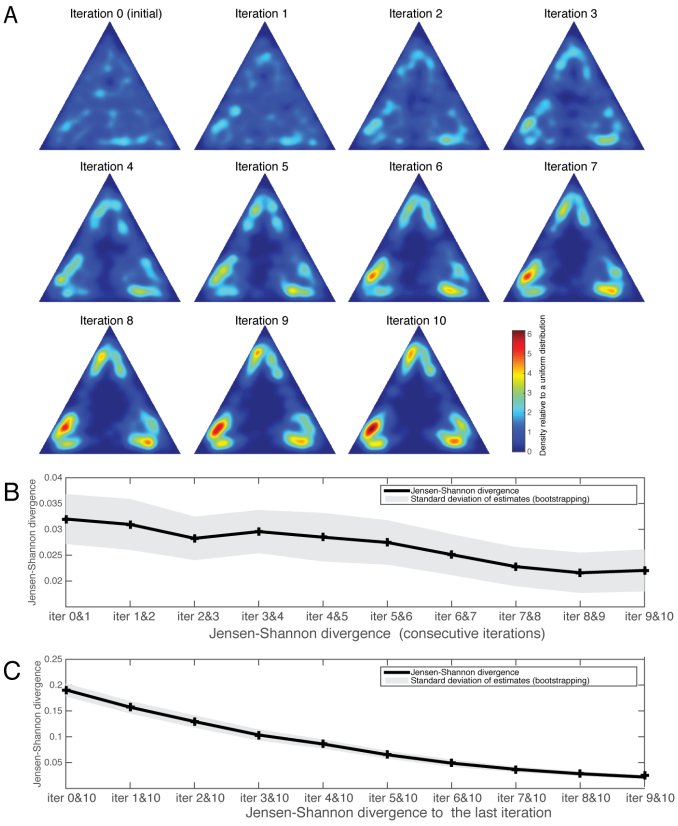

Figure 4: Results we obtained for each iteration in the chain for the triangle shape. A. Kernel Density Estimate (KDE) for the initial distribution and the 10 iterations. B. Convergence analysis using the Jensen-Shannon divergence (JSD) between consecutive iterations. C. JSD between all iterations and the last iteration. Note that both measures decrease with the number of iterations, and suggest that convergence occurs at or near the tenth iteration in the serial reproduction chain.

As another measure of convergence, we also computed the JSD between all iterations and the last iteration (Jacoby & McDermott, 2017) (see panel C). The distance between the last two iterations was significantly smaller than the distances between iteration 10 and each of the remaining iterations (0 through 8). The distance between iteration 10 and 9 was marginally larger than the distance between iteration 10 and 8 (p = 0.041). These analyses suggest that convergence occurs at or near the tenth iteration. To test if the responses of participants became more "prototypical" over the course of the experiment (as they progressed through their trials), we used the estimate of the prior from the final iteration to measure the average log-likelihood of their responses. We used data from the 83% of the participants who performed more than 80% of the trials within the accepted criteria (responses within 8% of the height and width of the shape on the screen). We found that the log-likelihood significantly improved when comparing the first and second half of their responses (t(49) = -2.47,

$p = 0.008$), and when comparing the first 10 trials to the last 10 trials of each of the subjects (t(49) = -2.04, $p = 0.046$).

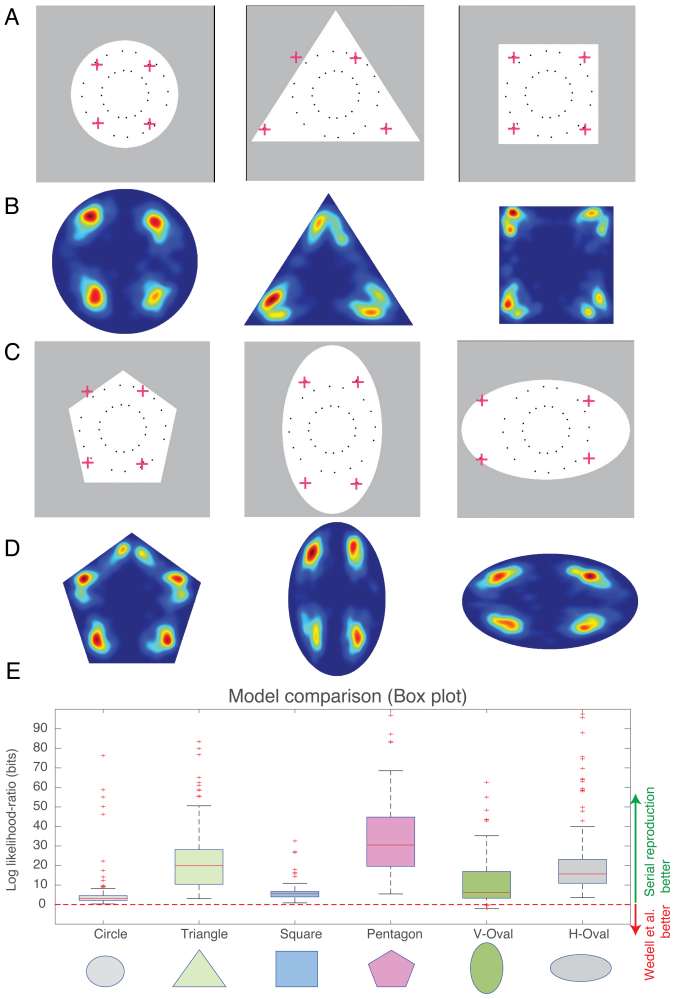

Figure 5: Kernel density estimates for the priors were estimated for all six shapes using the tenth iteration of the serial reproduction chain. A. & C. Original result by Wedell et al. (2007). B & D. Kernel density estimates with serial reproduction. E. Boxplots showing model comparisons. We computed the log likelihood difference for the two models as explained in the main text. In all cases the serial reproduction model was significantly better ($p < 0.01$ for all shapes except vertical oval ($p = 0.03$) resulting in positive log-likelihood ratios.

Model comparisons. Using a combination of non-parametric kernel density estimation and serial reproduction lets us uncover intricacies in the prior for angular shapes (including bimodal peaks at the vertices) that paint a nuanced picture of human spatial memory priors. In addition, our approach enables us to obtain more than just point estimates of the locations of prototypes in spatial memory. Nevertheless, we provide a comparison between point estimates obtained from our method to those obtained from the model by Wedell et al. (2007), for each shape, using the same number of parameters. The model describes the remembered position for a dot i (a response vector $\vec{R}_i$) as a weighted average of the actual location at which the dot was presented, which they refer to as the "fine-grain memory representation", and the weighted sum of the prototype locations, using the following equations:

$$\vec{R}_i = w\vec{S}_i + (1-w)\sum_{j=1}^{4} v_{ij}\vec{P}_j \tag{1}$$

$$v_{ij} = \frac{e^{-c\|\vec{S}_i - \vec{P}_j\|}}{\sum_{k=1}^{4} e^{-c\|\vec{S}_i - \vec{P}_k\|}} \tag{2}$$

where $\vec{S}_i$ and $\vec{R}_i$ are vectors in $\mathbb{R}^2$ containing the x and y coordinates for each point i in the stimulus phase (iteration 0), and in the first response phase (iteration 1), respectively. The $\vec{P}_j$ terms correspond to the four prototype vector coordinates being estimated by the model, in addition to weights w that correspond to the relative strength of the veridical memory (as opposed to the strength of a prototype in the prior).The v_{ij} capture the relevance weight of each of the four j prototypes for each point i. In other words, the strength of the influence of prototype j for each point i. The parameter c corresponds to a "sensitivity" parameter that models the sharpness of the prototype boundaries.

We generated 100 split-half samples of the points for iteration 0 (inital seeds), iteration 1, and iteration 10. Next, for each sample, we obtained estimates of the prototype locations for four prototypes (the same number used by Wedell et al.) by running their model using the training half of iteration 0 and the same points in iteration 1. In order to ensure a fair comparison, we sampled four points under local maxima from the Kernel Density Estimate (KDE) fit to the same points in iteration 10. This gave us four prototype estimates from the Wedell et al. (2007) model, and four points corresponding to local maxima in the KDE we fit to the points in the training half of iteration 10 (which can only be obtained from our paradigm), for each training split half. We evaluated the accuracy of these two sets of four prototype estimates by computing the sum of the negative-log-likelihood values from a KDE that we fit to the remaining points in the testing half of iteration 10. Next, we computed the log likelihood difference for the two models, for each of the shapes. In all cases, the serial reproduction model performed significantly better ($p < 0.01$ for all shapes except the vertical oval ($p = 0.03$) resulting in positive log-likelihood differences. Boxplots showing all the results are displayed in Figure 5E.

Grouping of prototypes. The apparent increase in peaks in the prior for more complex regular shapes afforded the opportunity to consider changes to the prior in the limit, as the shapes begin to approximate a circle. We computed the entropy of the obtained KDEs to quantify their complexity. Complexity increased with the number of vertices (going from a triangle to a heptadecagon, or seventeen-sided regular polygon). However, the prior for a icosihenagon (twenty-one sided regular polygon) begins to reveal the transformation of

the corner peaks into one of the quadrant peaks. Entropy further decreases for the icosipentagon ($p < 0.001$), revealing a prior that appears nearly identical to the prior for a circle, and with similar entropy ($p = 0.68$) (see Figure 6).

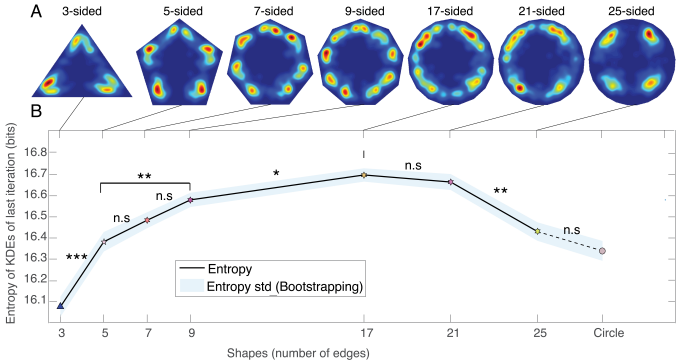

Figure 6: Grouping, and complexity of prior estimates. A. KDEs for regular polygons of increasing complexity. B. Entropy of the last iteration computed for all shapes. Entropy increases steadily with shape complexity (3 to 17 vertices). After the number of vertices exceeds 21, entropy stabilizes, and peaks start grouping toward the nearest quadrant center (as with the circle). We used the Bonferroni correction for multiple comparisons.

Discussion

In this paper, we made a preliminary foray into exploring spatial memory priors using serial reproduction: a process in which information being transmitted through successive participants leaves behind only a signature of the transformation process itself: the perceptual and reconstructive biases of those participants. This iterative process provides an effective tool for greatly amplifying biases in perception and memory.

We used a serial reproduction paradigm in the context of a spatial memory task. KDEs of the dots' final positions revealed detailed structure in priors over location. We found that the priors for circles and ovals show peaks at the center of each of their four quadrants, but also discovered that angular shapes show bi-modal peaks at the vertices in the prior. The modes appear on either side of each vertex, and do not seem to be a simple function of the angle at each vertex, since they are not rotationally invariant in all cases. We provided quantitative comparisons between the performance of a parametric model, and point estimates derived from the KDEs we obtained following ten iterations of the chain. These comparisons demonstrated that our estimates were significantly better than those obtained from the parametric model (we used the same number of parameters—four prototype estimates, even though our method yields kernel density estimates that clearly reveal more than four in some cases). In future work, we intend to determine if priors differ across individuals, by repeating the experiments so that each participant completes a subset of chains in their entirety (within-subject design). While some studies show differences between within and between-subject designs (Claidière et al., 2014), most studies showed high agreement between these versions (Xu & Griffiths, 2010; Jacoby & McDermott, 2017).

Our results suggest that our approach may provide an opportunity to uncover complex priors for a wide range of perceptual phenomena that would otherwise elude traditional experimental approaches, and parametric models. We plan to use it to measure memory biases when there is more than one point to be remembered (Lew & Vul, 2015), and to probe for structured priors in memory for local orientation (Wei & Stocker, 2016). Finally, we intend to uncover perceptual biases in spatial memory using natural complex images, and maps, to explore the effect of higher-order visual features and semantic content on spatial memory biases, and to probe for the emergence of geographic landmarks.

Acknowledgments

This work was funded in part by NSF grant 1456709 to T.L.G., and DARPA Cooperative Agreement D17AC00004 to T.L.G and J.W.S. We would like to thank Thomas Morgan for his help with online experiments.

References

Bartlett, F. C. (1932). Remembering: An experimental and social study. *Cambridge: Cambridge University*.

Claidière, N., Smith, K., Kirby, S., & Fagot, J. (2014). Cultural evolution of systematically structured behaviour in a non-human primate. *Proceedings of the Royal Society of London B: Biological Sciences*, *281*(1797), 20141541.

Huttenlocher, J., Hedges, L. V., & Duncan, S. (1991). Categories and particulars: Prototype effects in estimating spatial location. *Psychological review*, *98*(3), 352.

Jacoby, N., & McDermott, J. H. (2017). Integer ratio priors on musical rhythm revealed cross-culturally by iterated reproduction. *Current Biology*.

Kirby, S., Cornish, H., & Smith, K. (2008). Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences*, *105*(31), 10681–10686.

Lew, T., & Vul, E. (2015). Structured priors in visual working memory revealed through iterated learning. In *Cogsci.*

Wedell, D. H., Fitting, S., & Allen, G. L. (2007). Shape effects on memory for location. *Psychonomic Bulletin & Review*, *14*(4), 681–686.

Wei, X.-X., & Stocker, A. A. (2016). Mutual information, fisher information, and efficient coding. *Neural computation*.

Weiss, Y., Simoncelli, E. P., & Adelson, E. H. (2002). Motion illusions as optimal percepts. *Nature Neuroscience*, *5*(6), 598–604.

Xu, J., & Griffiths, T. L. (2010). A rational analysis of the effects of memory biases on serial reproduction. *Cognitive Psychology*, *60*(2), 107–126.