Algebraic Patterns as Ensemble Representations

Anna Leshinskaya (alesh@sas.upenn.edu) and **Sharon L. Thompson-Schill (sschill@psych.upenn.edu)**
Department of Psychology, University of Pennsylvania
Stephen A. Levin Building, 425 S. University Ave
Philadelphia, PA 19104

Enoch Lambert (elambe03@tufts.edu)
Center for Cognitive Studies, Tufts University
115 Miner Hall
Tufts University
Medford, MA 02155 USA

Abstract

Observers rapidly extract summary statistics from sets of visually presented items, like the mean size of a set of circles, or the mean expression of a set of faces. Their excellent ability to report summary statistics stands in contrast to near-chance representation of any of the individuals. Here we asked to what extent this 'ensemble perception' signature extends to a more abstract property: *relations* among elements. Participants watched ten unique animations of visually patterned objects (hereafter, 'shapes') colliding with each other and producing a new shape. Collisions conformed to ABA patterns, such that the result shape always matched one of the collider shapes. Recognition tests showed that participants accurately recalled the collisions they saw, but also falsely accepted foils which conformed to the ABA pattern but which were not in fact specifically seen (were rearrangements of the original shapes across collisions). On the other hand, they were much less likely to accept foils which did not conform to the pattern, but were equally distinct rearrangements (e.g., AAB). This suggests that participants represented the overall, common pattern better than the specifics of what they saw; the superior encoding of the summary relative to the individuals thus applies to summaries of relations. However, in contrast to prior findings with visual features, we did not find that recall of individual patterns was entirely at chance. Our paradigm offers a way to pursue future questions such as the pressures and motivations which might govern the trade-off between summarizing evidence vs. retaining individual experiences.

Keywords: ensemble perception; artificial grammar learning; pattern recognition; episodic memory; semantic memory

Introduction

Rather than encoding experiences in perfect detail, the mind naturally uses regularities and summary statistics to compress them. We can keep more items in working memory if items are predictive of each other (Brady, Konkle, & Alvarez, 2009); it becomes faster to find images in search display if they appear in predictable spatial configurations (Chun & Jiang, 1998); and we spontaneously and obligatorily register the mean orientation of sets of gabor patches (Parkes, Lund, Angelucci, Solomon, & Morgan, 2001). Although contingencies and averages are distinct statistics, all of these cases demonstrate that we spontaneously compress experiences, by encoding a summary of what is common across them.

In principle, representational systems can differ in the extent to which they compute summaries and discard individual observations (Dennett, 1991). Curiously, human participants sometimes represent summaries better than the observations composing them. When we see sets—like a series of differently-sized circles—we recall their mean (here, size) substantially better than we can recall any particular individual (Ariely, 2001; Chong & Treisman, 2003; Haberman & Whitney, 2009). This is true even when the number of items is relatively small (4) and when the items are presented sequentially. This suggests that we compute summaries and update them rapidly, discarding the items that went into this computation along the way. This 'ensemble perception' signature is true for visual properties like size, orientation, or facial expression (see Alvarez, 2011 and Whitney & Yamanashi Leib, 2018 for reviews). Here we asked whether this signature also applies to a property which is not a visual feature, but rather an abstract rule.

Algebraic rules (Marcus, 2001) are patterns based on relations among elements, such as *same* and *different*. For example, triplets of syllables can be readily seen as belonging to patterns like ABA— "ga di ga", "ku la ku", or "do re do"—vs AAB—"ga ga di". Learners (adults or infants) can recognize such patterns even with entirely distinct syllable sets and in both auditory and visual modalities (Ferguson, Franconeri, & Waxman, 2018; Marcus, Vijayan, Rao, & Vishton, 1999; Saffran, Pollak, Seibel, & Shkolnik, 2007). Algebraic rules are hallmarks of relational thinking, requiring relatively advanced computational architecture (Marcus, 2001; Overlan, Jacobs, & Piantadosi, 2017). They are also excellent compressions: recognizing that the last element always matches the first reduces the number of bits needed to represent the triplet by 1/3. Thus, despite the possible computational cost, encoding relations among stimuli is adaptive for circumventing limited memory capacity.

Here we asked how a representation of a shared algebraic pattern relates to the representation of the diverse individuals exhibiting the pattern. Specifically, we asked whether we would see the signature of ensemble perception. If so, participants should not only recognize that the set of items tends to follow an ABA pattern, but they should find it easier to recall that abstract pattern than the particular items they specifically saw (for instance, “ga di ga”, but not “ga ku ga”). Alternatively, due to the computational demands of inferring algebraic patterns, participants may be less reliable at recognizing the common ABA pattern across items, failing to summarize the data, and may be better at discerning which individual items they specifically saw.

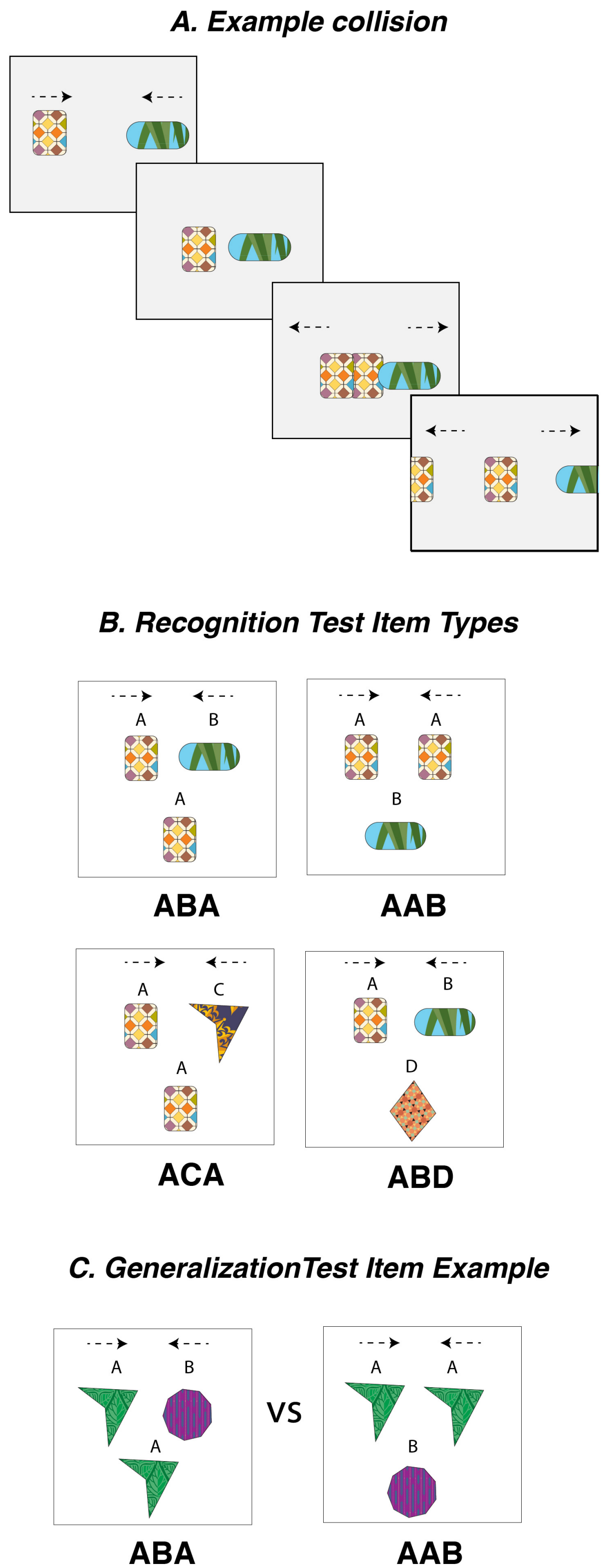

Figure 1. Top: Example of a collision between an A and B shapes and an A shape result. Bottom: Foil types for this collision, represented schematically. C shapes were taken from other collisions presented during the demonstration.

We used a novel paradigm in which participants watched pairs of novel shapes collide with each other and ‘produce’ a third shape (Figure 1). Participants’ task was to watch the collisions and see how many they could remember. As there was no instruction to look for patterns, the choice to summarize or not had to be intrinsically motivated.

Each collision was in fact governed by an ABA pattern: two distinct shapes, A and B, collided, producing another A. (We use ABA to denote the abstract elements, and lowercase *aba* to denote specific shapes). After watching 10 unique ABA collisions once each, participants performed a recognition test where each test item was either a collision they had seen, or one of three kinds of foils (unseen collisions). ACA foils swapped shapes between collisions: if specific collisions *aba* and *dcd* were shown, foils were *aca* and *dbd*. AAB foils were rearrangements of the same shapes in seen collisions, so that two A’s collided to produce a B. We reasoned that if participants recalled the common pattern better than the individual items, they should accept ACA foils at a higher rate than AAB foils. This is because ACA is pattern-consistent while AAB is not, though in terms of individual shapes composing the collisions, ACA is in fact more different from the original. ABD foils were also used as these were equal in the number of element-wise changes from ABA as ACA, but were also pattern-inconsistent.

We were also able to ask whether participants recalled *only* the summary pattern, and lost all item representations, by seeing whether they accept ACA foils at the same rate as ABA correct items. In ensemble perception, tests of individual recognition are often at chance (Ariely, 2001; Haberman & Whitney, 2009). Finally, a forced-choice test with new items directly tested whether learners represented the pattern in generalizable form.

Methods

Participants

30 participants were recruited and tested via Amazon Mechanical Turk. Participants provided electronic consent and procedures were approved by the Institutional Review Board of the University of Pennsylvania. Compensation was \$2. Three participants were excluded for failing an attention measure, and one for missing data. The included sample had 15 females and 15 males, with age M = 37, range 21 – 64). The task took an average of 15.62 minutes.

Stimuli

Stimuli were animated shape collisions (Figure 1). In each animation, two shapes approached each other from the left and right sides of the screen, met in the middle, and a third 'result' shape appeared between them as they moved away. Individual frames were created in Adobe Illustrator and concatenated into GIFs. Each GIF was composed of 23 frames shown at a 180 ms framerate and 4.14 s duration. GIFs were interspersed with 660 ms of blank screen for a 4.8 s total ISI. The majority of shapes used in the displays is shown in the Appendix.

Procedure

The task was presented to participants using a custom JavaScript webpage. It began with a *demonstration phase.* Participants were shown the following instructions: "You will play a game where you will see pairs of shapes collide with each other and see how many you can remember." They then watched 10 unique demonstration collisions in randomized order (lasting ~ 1 minute). Each of these collisions followed an ABA pattern: the two collider shapes, A and B, were distinct, and the result shape was a duplicate of A. A total of 20 different shapes were used, so that no shapes were repeated across collisions.

They then saw the *specific recognition test*. On each trial, a collision was shown, and participants had to decide whether or not they had seen it in the demonstration phase, by clicking 'yes' or 'no' after it ended. They were allowed to replay the collision. Apart from all 10 demonstration collisions, test items also showed three types of foils, created by rearranging the shapes across or within the demonstration collisions. ACA foils swapped the 'B' shapes between two different collisions, so that if specific shapes *aba* and *dcd* had been shown, foils were *aca* and *dbd*. AAB foils rearranged the shapes within an original collision, so that now two A's collided to produce B. ABD foils produced a result shape taken from another collision. The swaps were selected by pairing the 10 collisions into 5 foil-pairs. One of each of the three foil types was shown for each of the 10 original collisions; thus, there were 10 data points for each participant for each test item type.

We also added three attention check items, which showed previously unseen shapes in which two of the same shape collided, producing another duplicate (i.e., an AAA pattern). Participants had to respond 'no' to all three attention items to be included in the further analyses. Overall, there were 43 specific recognition test trials, shown in randomized order. There was no trial-level feedback, but an overall score was shown at the end of the test.

Participants were then given the *generalization test*. The instructions read, "The collisions you first watched followed certain patterns or rules. Now you will see new collisions and be asked to decide which ones follow similar patterns or rules." A two-alternative forced-choice test asked them to choose between pairs of collisions, shown one at a time, side by side. We used previously unseen shapes to create two new sets of ABA, AAB, and ABD items. Critical questions asked participants to choose between a pattern-consistent collision (ABA) and one of the two foils (AAB or ABD). Filler items showed the two foils, in order to balance the number of times each collision was shown overall. Each question type was shown once for each novel shape set, creating a total of 8 trials.

Finally, we asked participants whether or not they took any notes during the task. No participant reported taking notes.

Results

Specific Recognition Test

We computed the percent acceptance rate ('yes' response) for each type of test item; results are shown in Figure 2. The correct test item (ABA) was identical to the collision previously shown; this was (correctly) accepted at a high rate (M = 85%, SE = 0.05%). The ACA foil item maintained the pattern but its middle shape was swapped across previously seen collisions; this was (falsely) accepted at a high rate (M = 73%, SD = 0.05%). The AAB foil item was accepted at a low rate (M = 15%, SE = 0.06%) as was the ABD foil item (M = 13%, SE = 0.04%).

A 4- way ANOVA indicated a significant effect of item type, $F(75,3) = 59.43$, $p < .001$. Planned t-tests were used to probe these differences pairwise. We found that ACA foils were accepted at a higher rate than AAB foils, $t(25) = 6.77$, $p < .001$, CI [40 75] and ABD foils, $t(25) = 6.16$, $p < .001$, CI [42 86], indicating that participants indeed represented the pattern better than the specifics. Nonetheless, we also found higher acceptance rates for the correct (ABA) items than the ACA foils, $t(25) = 3.30$, $p = .002$, CI [4 19], indicating that item information was not completely lost.

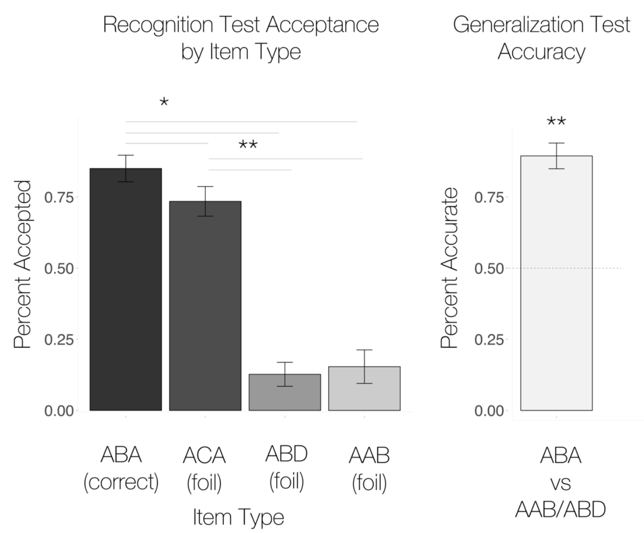

Figure 2. Rate of acceptance on the specific recognition test for each type of test item, and accuracy on the generalization test. Statistical comparisons are shown with * indicating $p < .01$ and ** indicating $p < .001$.

One potential account of these effects is that individual item representations decayed more rapidly or were more susceptible to interference from the question presentations. It should be noted that the majority of test items are also pattern-inconsistent, and so the amount of interference should be equal for both specific and pattern recall; however, it could still be that the item representations are more susceptible. We thus looked at responses from the first 10 test items only (thus, a random subsample from each participant). The critical effect of greater ACA than AAB acceptance was still significant, $t(25) = 7.10$, $p < .001$, CI [45 82], and the difference between ABA and ACA was marginal, $t(25) = 1.89$, $p = .07$, CI [-1 28].

Generalization Test

Participants were reliably above chance on choosing the pattern-consistent, entirely novel ABA collision relative to both foils; AAB: $M = 89\%$, $SE = 0.05\%$; $t(25) = 7.63$, CI [78 99], $p < .001$; ABD: $M = 90\%$, $SE = 0.05\%$; $t(25) = 8.38$, CI [81 100], $p < .001$ (Figure 2). They were thus highly reliable in learning a generalizable representation of the algebraic rule. As there was no difference between the two foil types ($t < 1$), accuracies for both were collapsed into a composite generalization score ($M = 89\%$). We found that this composite accuracy was not significantly different from the rate at which participants accepted the correct ABA items on the specific recognition test ($t < 1$), indicating that the representation of the abstract pattern was no worse than specific recall. We also found that accuracy on the generalization test was substantially higher than participants' ability to accurately reject the ACA foils (i.e., inverse of their acceptance rate; $t(25) = 7.45$, CI [47 82], $p < .001$. This is in line with the findings from the specific recognition test that the representation of the abstract pattern was superior to specific recall.

Discussion

We investigated the relationship between the ability to recall specific items (unique collisions of three shapes) and to identify and recall the common pattern governing them (here, an ABA algebraic rule). We found that analogously to signatures in ensemble perception, participants recalled the common pattern substantially better than the specifics of the individual items. Nonetheless, some memory of the individuals persisted, in contrast to certain findings with visual feature ensembles.

Our results indicate that the core signature seen in ensemble perception—superior fidelity of summary statistics over individual items—generalizes beyond visual features like size, facial expression, or line orientation (Whitney & Yamanashi Leib, 2018) and similarly applies to relational properties over visual events, like algebraic rules. This substantially extends the repertoire where such ensemble signatures might be found.

Our findings also speak to the question of how much a pattern-based summary relies on the representation of the individuals being summarized. Individual items must of course be processed at *some* level, but showing that their details can be quickly forgotten in spite of near-ceiling summary representations suggests that this level is relatively minimal. Because items were shown sequentially, and were short-lived, learners had to encode the pattern and update the summary with each subsequent representation—otherwise, it would be too late. It could therefore be the case that the item representation is discarded almost immediately after it is perceived.

The literature on episodic memory has similarly investigated whether summary recall is dependent on item recall, and has separated out these representations using delay paradigms and studies of amnesia. With multi-day delays, animals' reliance on the locations of specifically experienced platforms in a water maze declines, and is replaced by a representation of their mean location (Richards et al., 2014). Patients with amnesia (impairment to episodic memory) are as able as controls to extract patterns in artificial grammar learning studies, but unlike them, fail on recognition tasks of individual items from which they learned that grammar (Knowlton, Ramus, & Squire, 1992). Here we offer an elegant way to show this dissociation in healthy participants within a few minutes of testing, and to directly quantify the amount of information preserved about the individual items and the overall patterns. This opens an avenue of research investigating the circumstances and pressures that may motivate our cognitive system to rely on one or the other.

What might such pressures be? If learning is an attempt to infer the underlying model that generates observations, specific experiences serve as evidence towards hypotheses about that model—for example, a mean value or an underlying structure (Tenenbaum, Kemp, Griffiths, & Goodman, 2011). Once their beliefs over relevant models are updated, data points could be discarded (Nagy, Török, & Orbán, 2018). In this light, it has been argued that the choice to update a model vs. keep the data may be informed by factors such as the number of relevant models or how likely the relevant model to update might eventually change (Nagy et al., 2018; Richards & Frankland, 2017). Here, the right model was the ABA pattern, which explained all observations reliably. We might predict that if the pattern sometimes changed, recalling the specifics of all collisions might be enhanced, as this suggests to the learner that the model may not tell the full story or might change. We plan to test this in future work.

Another factor may be the computational cost of that update. Representing items in terms of their relations may be inferentially complex (Frank & Tenenbaum, 2011; Kuehne, Gentner, & Forbus, 2000; Overlan et al., 2017) and appears optional: one could perceive and remember a specific collision without ever representing the relations among its elements. If hypotheses about relations are computationally costly to update, the compression benefit of computing a relation may not outweigh the costs. The qualitative divergence we saw between algebraic patterns here vs. visual features in the past is consistent with this possibility: in the case of algebraic patterns, representations of individuals were not entirely lost, while for visual feature summaries, they often are (Ariely, 2001; Haberman & Whitney, 2009). If visual features require fewer inferential steps to encode than relational patterns, this could be consistent with that idea. Our paradigm offers a way to test some of these questions directly in future work.

Acknowledgments

This work was supported by NIH grant R01DC015359 to S.L.T-S. This project was also made possible through the support of a grant from the John Templeton Foundation, to A.L, E.L., and S.L.T-S. The opinions expressed in this publication are those of the authors and do not necessarily reflect the views of the John Templeton Foundation.

References

Alvarez, G. A. (2011). Representing multiple objects as an ensemble enhances visual cognition. *Trends in Cognitive Sciences*, *15*(3), 122–131.

Ariely, D. (2001). Seeing sets: Representation by statistical properties. *Psychological Science*, *12*(2), 157–162.

Brady, T. F., Konkle, T., & Alvarez, G. A. (2009). Compression in visual working memory: Using statistical regularities to form more efficient memory representations. *Journal of Experimental Psychology: General*, *138*(4), 487–502.

Chong, S. C., & Treisman, A. (2003). Representation of statistical properties. *Vision Research*, *43*(4), 393–404. https://doi.org/10.1016/S0042-6989(02)00596-5

Chun, M. M., & Jiang, Y. (1998). Contextual cueing: Implicit learning and memory of visual context guides spatial attention. *Cognitive Psychology*, *36*(1), 28–71.

Dennett, D. C. (1991). Real patterns. *The Journal of Philosophy*, *88*(1), 27–51.

Ferguson, B., Franconeri, S. L., & Waxman, S. R. (2018). Very young infants learn abstract rules in the visual modality. *PLoS ONE*, *13*(1), 1–14.

Frank, M. C., & Tenenbaum, J. B. (2011). Three ideal observer models for rule learning in simple languages. *Cognition*, *120*(3), 360–371.

Haberman, J., & Whitney, D. (2009). Averaging facial expression over time. *Journal of Vision*, *9*(2009), 1–13.

Knowlton, B. J., Ramus, S. J., & Squire, L. R. (1992). Intact artificial grammar learning in amnesia: Dissociation of category-level knowledge and explicit memory for specific instances. *Psychological Science*, *3*(3), 172–179.

Kuehne, S., Gentner, D., & Forbus, K. (2000). Modeling infant learning via symbolic structural alignment. *Proceedings of the twenty-second annual conference of the cognitive science society*, 286–291.

Marcus, G. F. (2001). *The Algebraic Mind*. Cambridge, MA: MIT Press.

Marcus, G. F., Vijayan, S., Rao, B., & Vishton, P. (1999). Rule learning by seven-month-old infants. *Science*, *283*(January), 77–80.

Nagy, D. G., Török, B., & Orbán, G. (2018). Semantic compression of episodic memories. *ArXiv Preprint ArXiv:1806.07990.*

Overlan, M. C., Jacobs, R. A., & Piantadosi, S. T. (2017). Learning abstract visual concepts via probabilistic program induction in a Language of Thought. *Cognition*, *168*, 320–334.

Parkes, L., Lund, J., Angelucci, A., Solomon, J. A., & Morgan, M. (2001). Compulsory averaging of crowded orientation signals in human vision. *Nature Neuroscience*, *4*(7), 739–744.

Richards, B. A., & Frankland, P. W. (2017). The persistence and transience of memory. *Neuron*, *94*(6), 1071–1084.

Richards, B. A., Xia, F., Santoro, A., Husse, J., Woodin, M. A., Josselyn, S. A., & Frankland, P. W. (2014). Patterns across multiple memories are identified over time. *Nature Neuroscience*, *17*(7), 981–6.

Saffran, J. R., Pollak, S. D., Seibel, R. L., & Shkolnik, A. (2007). Dog is a dog is a dog: Infant rule learning is not specific to language. *Cognition*, *105*(3), 669–680.

Tenenbaum, J. B., Kemp, C., Griffiths, T. L., & Goodman, N. D. (2011). How to grow a mind: statistics, structure, and abstraction. *Science*, *331*(6022), 1279–85.

Whitney, D., & Yamanashi Leib, A. (2018). Ensemble perception. *Annu. Rev. Psychol*, *69*, 12.1-12.25.

Appendix