

# The Role of Symmetry in Geometric Intelligence

Snejana SHEGEVA, Ashok GOEL

Design & Intelligence Laboratory, School of Interactive Computing,  
Georgia Institute of Technology, Atlanta, GA, USA  
[sshegheva3@gatech.edu](mailto:sshegheva3@gatech.edu), [ashok.goel@cc.gatech.edu](mailto:ashok.goel@cc.gatech.edu)

**Abstract.** The exploration of geometrical patterns stimulates the imagination and encourages abstract reasoning, which is a distinctive feature of human-level intelligence. In cognitive science, Gestalt principles such as symmetry have often explained significant aspects of human perception. We present a computational technique for building artificial intelligence (AI) agents that use symmetry as the organizing principle for addressing Dehaene’s test of geometric intelligence. Our work offers symmetry as a core principle for building AI agents capable of geometric intelligence and understanding Gestalt principles in human perception.

**Keywords:** knowledge representation, visual reasoning, geometric perception, symmetry

## 1 Introduction

George Polya argued that symmetry plays an important role in the inductive phase of complex problem solving by reducing and ordering the observable facts (Pólya, 1954). Captivated by visual diagrams of Polya’s work in crystallography, M.C. Escher created a *systematic* organization of geometrical transformations and enshrined symmetry as the principal rule underlying his art (Escher and Schattschneider, 2004). Without systematic knowledge of the mathematics governing patterns of symmetry, he created his own “layman’s theory” of symmetry, duality, infinity, and paradoxes. Escher comes to *the open gate of mathematics* by exploring how concepts like *repetition*, *rotation* and *reflection* shape our interpretation of boundaries between shapes (Haak, 1976).

The artwork Escher produced over his lifetime profoundly challenges our visual perception of the world. Equally impressive, most humans can understand and appreciate the beauty of Escher’s drawings, even in the absence of previous experience with them. Consider, for example, just the two graphics shown in Figure 1. It is easy to see some of the symmetry concepts – such as translation, rotation, and reflection – which invites questions about the nature of cognitive processes when we perceive this kind of art.

Indeed, Gestalt psychology has long proposed symmetry as organizing principle of geometric intelligence (Bornstein et al., 1981), (Li, 2009). Gestalt theories suggest

that human cognition uses repetition, such as translational symmetry (see example in Figure 1, left), or perceptual shift between foreground and background (see example in Figure 1, right) in creating and making sense of art (Tyler, 1995). These theories raise the basic question motivating our work: Might it be possible to build artificial intelligence (AI) agents that use the principles of Gestalt psychology to make complex inferences about geometric patterns and transformations?



Fig. 1: M.C. Escher work inspired by mathematics and nature (Lizards (left) and Bird/Fish (right))

To provide algorithmic answers to this question about designing AI agents, we start with Dehaene's test of geometric intelligence (Dehaene et al., 2006) that shares several themes with Escher's more intricate structures. Dehaene *et. al.* describes symmetry as a *geometrical language* that adults and children can comprehend regardless of their culture and background (Amalric et al., 2017). Dehaene developed the test containing 45 problems to examine whether humans brought up in a technologically advanced civilization with the benefit of formal education, including geometry, performed better than subjects from a technologically primitive society with little formal education. He found that subjects from the Mundurku tribe in the Amazon forests performed about as well on the test as the subjects from a western society. Although Dehaene's experiments were not conclusive, they seemed to indicate that core geometric intelligence might be innate to all humans.

Dehaene's test eschews geometric objects such as triangles and instead relies on more abstract concepts such as *closure*. All 45 problems on the test explore various aspects of core geometry, such as Euclidean geometry, topology, symmetrical figures, metric properties, and geometric transformations (Dehaene et al., 2006). Each problem is an array of six images where one violates the displayed concept, and the test taker attempts to identify it as the one that *breaks* the structure. Figure 2 shows an example that highlights the above-mentioned concept of closure. Although at first glance, symmetry is not explicit in Figure 1, we will show below those specific representations of the drawings in Figure 1 derived from Euclidean transformations capture the latent symmetry and order in the pictures.

Although Dehaene's problems are different from Escher's more intricate drawings, they nevertheless entail similar, if more straightforward, kinds of abstract reasoning. Abstract reasoning on Dehaene's test requires inferences to higher-level concepts such as relations, symmetries, and intricate patterns from low-level pixel representations.

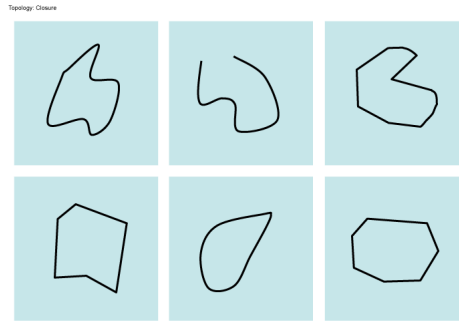


Fig. 2: An example of Dehaene’s geometry problems that explores topological concepts of closure. One image (here, the top center) violates the concept and therefore should be considered as odd-one-out.

According to Dehaene, using geometrical tests with perceptually accessible features such as shapes, positions, and between-object relations, a human capacity to reason abstractly can be measured independent of their culture, language, or experience. This point brings us back to the research question motivating our research. If Gestalt principles underlie human perception, might the same principles form the basis for building AI agents capable of addressing problems on the Dehaene test? Further, if AI agents based on Gestalt principles indeed can be built for Dehaene’s test, what may that tell us about human cognition?

## 2 Related Work

A common feature among many of these studies is a focus on similarity and especially analogy. Carpenter *et al.* provide a detailed cognitive model of problem-solving on the Raven’s Progressive Matrices Test of general human intelligence (Carpenter *et al.*, 1990). Their model is based on the production system architecture in which the agent has access to a variety of rules that capture the range of geometric patterns that occur in Raven’s test.

Lovett, Lockwood & Forbus (2008) view Raven’s test as geometric analogy problems and use the structure-mapping theory of analogy to address them. They describe a cognitively inspired approach that detects geometric shapes from an input drawing on Raven’s test, constructs spatial representations of relations among the objects, and then applies the structure-mapping technique for addressing the problem.

Kunda, McGregor & Goel also view the Raven’s test as a set of visual analogy problems. However, in contrast to Carpenter *et al.* and Lovett *et al.*, they use *affine* transformations, such as translation, rotation, and reflection, directly on pixel-level representations to address the Raven’s test, including the Standard, Color, and Advanced Raven’s test (Kunda *et al.*, 2013). Given an input image, their ASTI model interprets the drawing in terms of linear combinations of affine transformations and completes

the problem in terms of transformation combinations. An exciting aspect of the ASTI computational model is that it does not have prior knowledge of geometric objects and does not need to detect objects. Nevertheless, its performance is comparable to that of earlier methods.

McGreggor, Kunda & Goel describe a method called FAR that makes analogies based on fractal representations. Given an input image, FAR first builds a fractal representation of the input and then uses similarity and analogy to address the problem. The FAR method has been successfully used for Raven's intelligence test, and the Odd-One-Out test (McGreggor and Goel, 2011) and the Dehaene's test (McGreggor and Goel, 2013). An interesting aspect of the self-similar fractal representations is that the FAR technique can automatically change the resolution level to match the given problem. Like ASTI, FAR too does not detect objects in the input images.

Shegheva et al. (2018) have developed a Structural Affinity method to address Raven's intelligence (Shegheva and Goel, 2018). The technique uses Markov Random Fields parameterized by affinity factors to learn the underlying rules described in Carpenter's work and subsequently recognize the pattern to make a prediction. The strong emphasis is on discovering topologies that do not rely on object detection but instead represent features for the type of relationship between images. Not to be confused with image similarity methods, the Structural Affinity captures the *generation* rule that represents the abstract reasoning ability.

Recently, there has been some work on using CNNs to address similar problems (Santoro et al., 2017, 2018; Zhang, Gao, Jia, Zhu and Zhu, 2019). So far, this work has focused on variations of Raven's problems and not yet addressed Dehaene's style of images. For example, Zhang et al. use neural models to generate a dataset of problems similar to Raven's test problems and feed them into a module that reasons based on perceptual contrasts (Zhang, Jia, Gao, Zhu, Lu and Zhu, 2019).

In our earlier work on geometric intelligence, we have used the Gestalt principles of perception to address several classes of problems, including problems on the Standard, Colored, Advanced Raven's tests, and Odd-One-Out tests. The present work builds on our earlier work; in particular, it is similar to a previous computational technique we have described for addressing problems on Raven's Standard test (Shegheva and Goel, 2018).

As noted above, many previous computational models and techniques have viewed geometric intelligence tests in terms of extant theories of similarity and analogy. In our current work, we postulate that *symmetry* plays a fundamental role in how an input image is perceived and forms the basis for analogy. Leyton proposes symmetry as a fundamental element of visual perception and relates it to reconstructing causal histories (Leyton, 1992). In his book "A Generative Theory of Shape," Leyton proposes a constructive representation and can capture symmetry-breaking, or symmetry-building, as a way to reason about the hierarchy of the image components Leyton (2003). Dehaene *et. al.* deliberately minimizes the core concept cues in a problem by randomizing the perceived features. For example, by varying the objects' orientation or modifying the size, the concept becomes obscured. We hypothesize that a suitable representation can *undo* the complexity and highlight the right context for the desired concept.

### 3 Structural Affinity for Core Geometry

In the current work, we build upon the idea of the *agreement* by constructing different types of affinity factors, each representing a property of the geometrical concepts covered in Dehaene’s test (Shegheva and Goel, 2018). The AI agent computes a series of estimations from each image and determines the one *odd-one-out* image in two steps: 1) identify the most relevant properties that attribute to significant deviations 2) rank images by the most considerable contribution of variance.

The rest of this section covers the specifics of our computational technique’s algorithm and individual components, starting with image pre-processing, feature design, and the algorithm for detecting a violation in latent concept.

#### 3.1 Preliminary Processing

Dehaene’s test’s distinctive feature is that each problem with six images demonstrates a single core concept with slight variations. A problem is solved if a test taker identifies the image that contains variation not explained by the concept. Therefore, the proposed method aims to identify the most likely concept by 1) unifying the representation, i.e., reducing superfluous features, and 2) ranking the remaining features by the signal’s strength towards a single pattern.

Appendix A lists all 45 Dehaene’s images grouped by the concept and ordered by difficulty. *Euclidean Geometry*, *Geometrical Figures*, and *Topology* (first three rows in Figure 8) are among the simplest concepts as measured by the performance of participants of the Amazonian indigenous group Mundurukú. This suggests that geometrical primitives are easily identified regardless of their orientation, color, alignment, or size. Problems that test for *symmetry*, *chirality*, and *metric properties*, such as distance, require an increased level of concentration since they involve spacial operations such as rotation and reflection. Unlike previous concepts, the induced orientation change had a more significant impact on making a correct inference. This is especially true for chiral figures (see Figure 8 - fourth row) where the participants’ performance dropped from 90% to 20% by rotating images along random axes. *Geometric Transformations* have the lowest performance score (35%), indicating that reasoning about motion in static images requires considerable analytical judgment.

In general, we observe that Dehaene’s problems exhibit geometrical primitives that fall into specific types of symmetry classes, even in cases where the symmetry concept is not a primary characteristic. For example, problems in the second row of the Figure 8 intended to highlight Euclidean geometry’s properties, such as distance, can also be interpreted with symmetry - reflective and rotational. Thus, with the assumption of preserving the symmetry, our normalization method rotates and translates the original images, re-mapping the pixels to the common axes. This *normalization* intensifies the most relevant features while removing variations intended to obscure the concept.

#### 3.2 Image Segmentation Phase

A Dehaene’s problem is represented as a 3x2 matrix of visual entities that capture geometrical shapes and transformations. To fit the problem into the structural affinity

framework, we first identify the grid and segment the images into its six components. The image is ready to be ingested by our computational technique without additional image pre-processing steps by applying the segmentation.

### 3.3 Representation Phase

The segmentation phase for one problem produces six panels that are subsequently transformed into an  $n \times m$  matrix  $A$  where each cell  $i, j$  takes on a binary value: 1 if the color intensity of the pixel is above a certain threshold  $\epsilon$ , and 0, otherwise. Before reading the pixels into an array, the image is optionally cropped to remove the pixels associated with plain text (typically at the top left of the first image).

As Dehaene's problems target geometry concepts, it makes sense to represent the images in the Cartesian coordinate system instead of a  $n \times m$  matrix where  $n$  and  $m$  are the height and the width of the given image. The binary values (0,1) are mapped to real numbers  $R^+$ . This representation returns a set of points  $\Omega$  in the coordinate system generated by Expression 1.

$$\Omega := \{(i, j) \mid A_{i,j} = 1, \quad 0 \leq i \leq n, \leq j \leq m\} \quad (1)$$

### 3.4 Transformation Phase

We use Principal Component Analysis (PCA) method to *unrotate* the figures and obtain the coordinate axes that contain the maximum variance (Tipping and Bishop, 1999). Figure 3(b) shows the result of a successful re-orientation that removes the randomness in the original axes and brings the *symmetry* feature into focus.

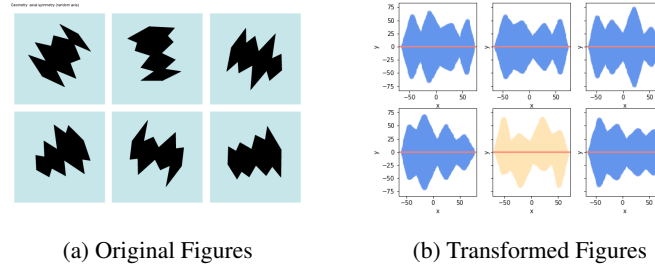


Fig. 3: **Symmetry Concept.** In (a) the concept is shown in its original raw form that include an orientation noise not relevant to the concept. In (b) the transformation applied highlights the single concept of symmetry

The classical principles of Gestalt perception - proximity, similarity, common fate, good continuation, closure, symmetry, parallelism (Wagemans et al., 2012) - inspire the design of features in our method. The method uses a set of simple heuristics used for

building knowledge representation. The goal is to design functions that can serve as cues for the underlying concept.

The **Color** concept can be encapsulated with a *density* function, that counts number of pixels of varying intensity. During the representation phase, the pixels are transformed into coordinate points, reducing the density function to a count of points.

The **Orientation** concept must be captured before the PCA transformation that rotates the figures to the simplest structure is applied. By computing *Pearson* correlation coefficient, we obtain the amount of linear relationships between points (Benesty et al., 2009).

**Topology** concepts, such as *inside/outside*, *closure*, *connectedness*, and *holes*, require several features - *contour count* and *child-parent* relationship between contours. For example, Figure 2 that highlights *closure* contains one figure that is inconsistent with the other figures with regards to the number of contours. The *odd* figure shows a disjointed curve whereas the consistent with the concept figures shows curves that join continuous points.

**Symmetrical Figures** after applying PCA transformation, images are automatically aligned along the axes of symmetry; thus, computing the discrepancies between the alignment points in the upper and lower quadrants will give a clue about how symmetrical is the figure.

First, we *collapse* the points to a single vector by computing the average value per x-coordinate. In symmetrical figures, the vector should contain values close to zero.

$$\hat{Y}_k = \frac{1}{n_k} \sum_{i=0}^{i=n_k} y_i[x_k] \quad (2)$$

where  $\hat{Y}_k$  is the average value of the points where  $x=k$ .

To obtain a scalar measure of the symmetry feature, we subsequently compute the variance of the  $\hat{Y}_k$  vector -  $Var(\hat{Y}_k)$  that captures the overall adherence to the concept. Likewise, we compute an average vector  $\hat{X}_k$  and y-axis along with its scalar representation via variance -  $Var(\hat{X}_k)$ .

The similar features are relevant for the concept of **Geometrical Figures**. A transformation maps specific geometrical figures into themselves along their axes of symmetry. For example, a square has four lines of symmetry, whereas a rectangle has only two. An equilateral triangle has three lines of symmetry, whereas an isosceles triangle has only a bilateral symmetry. Designing several functions that capture the heuristics of symmetry (reflection, translation, rotation), in conjunction with other features, can identify the image most inconsistent with the core concept in the problem.

**Euclidean Geometry** (e.g., line, points, parallelism, and right angle) requires a subset of the features defined above. Figure 4 suggests that a feature that computes a variance along axes can quickly identify the *odd-one-out* image that violates the consistent measure across the remaining figures.

Problems that assess the ability to detect **Metric Properties**, such as distance between objects, center and middle segments, are likewise dependent on the *symmetry* features. In problems where the center of an item is intentionally offset, a feature that creates a mapping between upper and lower quadrant pixels exhibits a *gap*, i.e., several

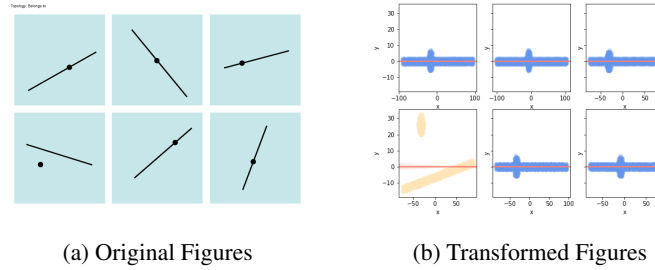


Fig. 4: **Alignment of Points Concept.** In (a) the concept is shown in its original raw form that include an orientation noise not relevant to the concept. In (b) the transformation applied highlights the single concept of alignment

points from that upper quadrant that do not have a pair from the lower quadrant (see Figure 5).

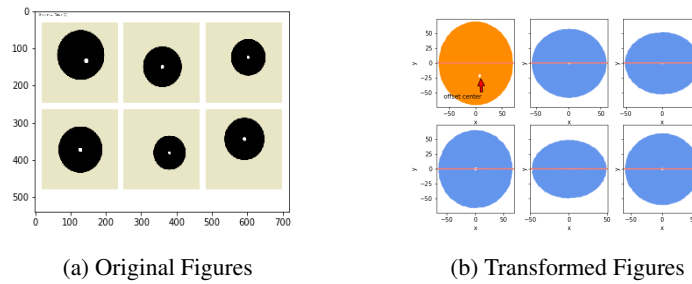


Fig. 5: **Center of Circle.** In (a), the concept is shown in its original raw form that includes circles of different sizes and positions. In (b), the transformation applied highlights the single concept of the location of the center point

**Chiral Figures** require an ability to perform *mental rotation* to align the figures in the same axes for comparison. We note that the ablation of a feature may remove information critical to recognize the concept in some instances. For example, in Figure 6, the concept of chirality is being erased by the transformation where pixels reorient around principal components. This ambiguity results from chirality's underlying property - an object cannot be mapped into its mirror image only by rotation and translation. Therefore, there is no symmetry operation (in 2D) that would preserve the object's invariance.

**Geometrical Transformations** involve *translation*, *homothecy*, *symmetrical reflection*, and *rotation*. Arguably, this is the most difficult mathematical concept, especially when given in static images (Dehaene et al., 2006). All previously described features



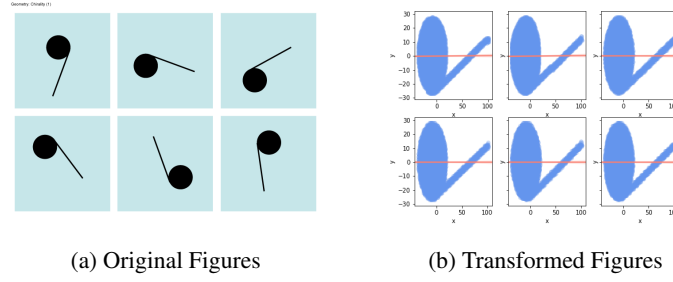


Fig. 6: **Chirality on Oblique Axes.** In (a), the concept is shown in its original raw form that includes chiral figures in random axes. In (b), the PCA transformation removed the *rotation* information which made all the figures identical.

are applicable here as well, although the more confident answers are achieved using a combination of several features.

### 3.5 Problem Solving Phase

Figure 7 illustrates the processing in our computational technique. The algorithm starts with the segmentation and visual encoding, as described in the sections above. After the raw features are extracted and transformed, a filtering method is applied to select the most prominent attribute  $S_i$  that might hold the cues for identifying the discrepant image.

$$S_i = \{f_k(Im_i^*) \mid \forall f \in Features, \quad z(f_k) \geq \delta\} \quad (3)$$

where  $f_k(Im_i^*)$  is a  $k$ -th feature extracted from all six segmented and encoded images for the original problem  $i$ ;

and  $z(f_k) = \frac{x_k - \bar{x}_k}{\sigma_k}$  - the standard score computed for the feature  $z(f_k)$ ; in our experiments  $\delta = 2$  in order to select features where an element is at least two standard deviations from the mean.

The intuition behind applying a z-score to the raw features is that Dehaene's test's core is to detect violations of some desired geometry concept. Therefore, we expect that the most pertinent attribute will contain an anomaly that can be captured with metrics specifically designed to detect outliers. Thus, if the set of figures contain a *odd* attribute, the feature will draw attention to it with an anomalous z-score. If none of the extracted features exceed the threshold, the extraction process is deemed insufficient, and the problem is skipped.

To identify the discrepant image,  $X^*$ , each of the filtered features *votes* for one of the six candidates. The solution is selected by choosing the candidate with the highest number of votes.

$$X^* = \arg \max_j Vote(X_j) \quad (4)$$

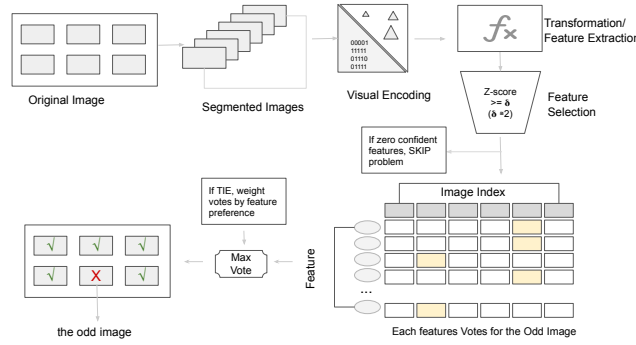


Fig. 7: A diagram of the computational workflow for solving Dehaene's geometrical problems. The process starts with the original image segmented into six panels, subsequently encoded and transformed for visual reasoning. Feature selection is performed based on the z-score exceeding a threshold  $\delta$  (in our case  $\delta = 2$ ). The discrepant image is the one that violates the observed consistency (voting occurs if more than one feature is selected as a candidate for the underlying concept).

Ideal are scenarios in which filtering yields a single feature, and no voting is required. The ambiguity only arises where more than one feature holds an attribute of the underlying geometry concept. For example, the *symmetry* concept can be captured with different high-level heuristic functions, such as one-to-one point mapping or distance to the centerline. For *ties*, voting is repeated with a feature preference weight (more specific features are preferred.)

Additionally, the algorithm applies reduced precision of the computations across all phases - from representation to transformation and scoring. This approach converts scores from exact numerical format to more rounded approximations, which helped with noise when working with raw images (Zadeh, 1984).

## 4 Results

### 4.1 Accuracy

Table 1 summarizes the performance of the designed algorithm per metric as presented in the Dehaene et al. experiment. The total accuracy measure is 89% with 40 out of 45 problems solved correctly. The last column in the table records the performance of Mundurukú participants. Problems involving basic geometry concepts, such as topology, geometrical figures, lines, and angles, have specific features and, therefore, more straightforward (100% correct). Problems that involve transformations are more complex as they involve mental translation, reflections, scaling, and rotations. The lower performance on those problems is consistent with the findings reported by Dehaene et al. Problems that exercise the concept of *chirality* (number of examples = 4) average

to a 50% accuracy with a significant difference between figures shown in vertical vs. random (oblique) axes ( 85% and 23% correspondingly).

Concept	true	total	ratio	human
Color/Orientation	2.0	2.0	1.0	95.0
Chiral Figures Aligned	2.0	2.0	1.0	88.0
Euclidean Geometry	8.0	8.0	1.0	84.0
Geometrical Figures	9.0	9.0	1.0	79.0
Topology	4.0	4.0	1.0	76.0
Symmetrical Figures	3.0	3.0	1.0	67.0
Metric Properties	6.0	7.0	0.86	62.0
Geometrical Transformations	6.0	8.0	0.75	34.0
Chiral Figures Random	0.0	2.0	0.0	23.0

Table 1: Algorithm performance per concept. The index of the table is the name of the concept as presented by Dehaene et. al.; first column is the count of correctly identified images by the structural affinity algorithm; second column is the total count of problem in the given concept; third column is the ratio of correctly solved problems; And the last column is the performance of the Mundurukú participants

## 4.2 Comparison to Other Computational Techniques

This paper demonstrates that by leveraging Gestalt principles of perception in the design of algorithms for problem-solving, our computational technique can approximate the *human-level* behavior on Dehaene’s core geometry tasks. Unlike techniques that rely on objects and borders detection for qualitative representations (such as (Lovett et al., 2008), (Lovett and Forbus, 2011)), we describe a computational technique that 1) *transforms* input images into more perceptually coherent variations, 2) scores the resulting representations against pre-defined properties (such as symmetry, rotation, and other geometrical concepts), and 3) identifies an instance where scores are in disagreement with the rest of images. Lovett et al. (Lovett et al., 2008), and Lovett and Forbus (Lovett and Forbus, 2011) integrate four different systems for addressing visual oddity tasks and solve 39 out of 45 problems correctly. Our approach’s advantage is that it solves a similar number of the problems correctly (40 out of 45). It does so while satisfying the *parsimony* characteristic, i.e., striving for the most straightforward theory that can explain the intent in the discussed geometry problem. Parsimony, in this context, is concerned with problem-solving behavior that characterizes many gifted individuals (Koichu, 2008).

In the computational technique that employs *fractal* representations (McGreggor and Goel, 2013), visual oddity tasks are addressed with a notion of visual similarity that operates on varying levels of image representation - from coarse to refined. In addition to providing solutions, the authors compute confidence measures (30 unambiguously correct and 15 correct but ambiguous) and compare them with human performance. An analysis of ambiguity allows tuning the levels of coarseness to determine the best

strategy for representation. McGreggor et al. claim that their algorithm is parsimonious because it does not use additional mechanisms for problem-solving phases (McGreggor and Goel, 2013). However, simplicity is not always a sufficient measure of parsimony; nor does simplicity always offer a robust explanation of intelligent behavior. Our approach is based on scoring images by their explainability with concepts such as symmetry, topology, and Euclidean geometry properties. An image is considered odd if its characteristics do not match the rest of the images' elements. An analysis of agreements and violations helps highlight the most pertinent attributes of the problem, thus giving an explainable answer.

An algorithm of visual perception that is intuitively based on Gestalt principles can infer higher-level abstractions from raw pixels and their relationships. This observation, in turn, increases the diverse set of capabilities to perceive regularities in images through the lens of symmetry, closure, similarity, and other geometrical concepts.

## 5 Discussion

One of the main questions raised by Dehaene et al. is whether core geometric intelligence is inherent in the human mind (Dehaene et al., 2006). This question is complex; the answer depends on how the geometrical concepts are represented and what type of mental operations are invoked when assessing a concept. Therefore, the research presented here is motivated by two goals: 1) to understand what transformations are ubiquitous for exploring geometrical notions, i.e., what kind of abstract reasoning is most relevant for dissecting a geometry problem, and 2) is there an underlying organizing principle that governs our perception of shapes, and their ability to be transformed into similar, but more perceptually coherent objects.

To answer (1), we developed a computational technique that transforms encoded images that hold geometry concepts using a mathematical mechanism to discover structure and relations among the variables. We found the applying PCA may be interpreted as performing mental transformations, such as rotations of the coordinate axes, to align the studied shapes. In Dehaene's problems, finding the principal components is a viable explanation of how the mind removes irrelevant differences between shapes to identify the central geometry concept. *Forgetting the details* is an idea pervasive in humans and is generally attributed to innate biology Biederman (1987). The resulting structural uniformity brings into focus those aspects that are most relevant to the task at hand, and allows identifying discrepancies that in Dehaene's problem represent the solutions.

To address (2), we observe that most of Dehaene's problems exhibit *symmetry* as a latent (unless directly specified) principle. Gestalt theory suggests that visual perception is frequently driven by the tendency to maximize the shapes' appeal and connection with other surrounding shapes (Spelke, 1990). Therefore, our computational technique benefits from viewing the geometry problems through the lens of symmetry. For example, Figure 5 is probing the concept of *metric properties*, specifically a center of a circle. One way of reasoning about that in concordance with a *metric* concept is to consider the distance of the inside point from all of the circumference points. Alternatively, it may be reasoned with the concept of symmetry - the misalignment of the inner point from the center in one of the figures *breaks* the symmetry group of the circle object, i.e., it is not

possible to map the figure on itself using rotations or reflections. Similar reasoning is applied to Dehaene's other problems, suggesting that symmetry may be the underlying organizing principle for basic geometry concepts. This idea is captured in the generative process, described by Leyton (2003), as a way to recover the necessary operations that can reconstruct any geometrical shape. By applying appropriate symmetry groups, a collection of related images (such as Dehaene's set) can be described in the context of one another, thus revealing the underlying structure of the whole concept.

The complexity of analyzing the perception of images arises from the presence of many possible patterns that are compatible with the observed features. Previous research has proposed parsimony (Epstein (1984)) and simplicity (Chater and Vitányi (2003)) as fundamental principles of cognition: they argue that mind prefers the simplest or most parsimonious explanations - an idea rooted in the Ockham's razor principle. The central theme of our research is an emphasis that perceiving geometrical images through the lens of symmetry lead to more unified and simpler explanations. This approach provides a justification for choosing some patterns over others and appears to be consistent with the Gestalt principles of perception.

## 6 Future Directions

The brain's self-organizing processes described by Gestalt principles enable spontaneous switching between the functions of the figure and ground (Wagemans et al., 2012). Thus, it is reasonably easy for humans to identify the motifs in Escher's drawings without requiring substantial previous exposure and training in geometry concepts. In the next stage of our research, we will probe an AI agent's ability to infer the figure/ground images' organizational functions.

The main focus of this work was to encode the general properties of grouping principles during perceptions of geometrical primitives in Dehaene's images. In future work, we intend to expand the arsenal of geometrical concepts that can be organized with various types of symmetrical operations and groups. In addition to identifying strictly repeating motifs (gliding symmetry), we will explore strategies for extracting motifs for M.C. Escher's tessellation works by analyzing each of the seventeen symmetry groups Schattschneider (2010). We also hope to deepen the interpretative capabilities of AI solutions by building the bridge between Gestalt principles of perception and algorithmic object manipulations for visual reasoning.

## Acknowledgement

The authors thank the members of the Georgia Tech Design & Intelligence Laboratory, especially Keith McGreggor, Tesca Fitzgerald and Priyam Parashar, for many discussions and helpful suggestions.

## References

- Amalric, M., Wang, L., Pica, P., Figueira, S., Sigman, M., Dehaene, S. (2017). The language of geometry: Fast comprehension of geometrical primitives and rules in human adults and preschoolers, *PLoS computational biology* **13**(1), e1005273.

- Benesty, J., Chen, J., Huang, Y., Cohen, I. (2009). Pearson correlation coefficient, *Noise reduction in speech processing*, Springer, pp. 1–4.
- Biederman, I. (1987). Recognition-by-components: a theory of human image understanding., *Psychological review* **94**(2), 115.
- Bornstein, M. H., Ferdinandsen, K., Gross, C. G. (1981). Perception of symmetry in infancy., *Developmental psychology* **17**(1), 82.
- Carpenter, P. A., Just, M. A., Shell, P. (1990). What one intelligence test measures: a theoretical account of the processing in the raven progressive matrices test., *Psychological review* **97**(3), 404.
- Chater, N., Vitányi, P. (2003). Simplicity: A unifying principle in cognitive science?, *Trends in cognitive sciences* **7**(1), 19–22.
- Dehaene, S., Izard, V., Pica, P., Spelke, E. (2006). Core knowledge of geometry in an amazonian indigene group, *Science* **311**(5759), 381–384.
- Epstein, R. (1984). The principle of parsimony and some applications in psychology, *The Journal of Mind and Behavior* pp. 119–130.
- Escher, M. C., Schattschneider, D. (2004). *Visions of Symmetry*, Thames & Hudson.
- Haak, S. (1976). Transformation geometry and the artwork of M.C. Escher, *Mathematics Teacher* **69**(8), 647–652.
- Koichu, B. (2008). On considerations of parsimony in mathematical problem solving, *Proceedings of the 32nd Conference of the International Group for the Psychology of Mathematics Education*, Vol. 3, Cinvestav-UMSNH Mexico, pp. 273–280.
- Kunda, M., McGreggor, K., Goel, A. K. (2013). A computational model for solving problems from the raven's progressive matrices intelligence test using iconic visual representations, *Cognitive Systems Research* **22**, 47–66.
- Leyton, M. (1992). *Symmetry, causality, mind*, MIT press.
- Leyton, M. (2003). *A generative theory of shape*, Vol. 2145, Springer.
- Li, Q. (2009). Interest points of general imbalance, *IEEE Transactions on Image Processing* **18**(11), 2536–2546.
- Lovett, A., Forbus, K. (2011). Cultural commonalities and differences in spatial problem-solving: A computational analysis, *Cognition* **121**(2), 281–287.
- Lovett, A., Lockwood, K., Forbus, K. (2008). Modeling cross-cultural performance on the visual oddity task, *International Conference on Spatial Cognition*, Springer, pp. 378–393.
- McGreggor, K., Goel, A. (2011). Finding the odd one out: a fractal analogical approach, *Proceedings of the 8th ACM conference on Creativity and cognition*, ACM, pp. 289–298.
- McGreggor, K., Goel, A. (2013). Fractal representations and core geometry, *Proceedings of the Second Annual Conference on Advances in Cognitive Systems ACS*, Vol. 3, p. 19.
- Pólya, G. (1954). *Mathematics and plausible reasoning: Induction and analogy in mathematics*, Vol. 1, Princeton University Press.
- Santoro, A., Hill, F., Barrett, D., Morcos, A., Lillicrap, T. (2018). Measuring abstract reasoning in neural networks, *International Conference on Machine Learning*, pp. 4477–4486.
- Santoro, A., Raposo, D., Barrett, D. G., Malinowski, M., Pascanu, R., Battaglia, P., Lillicrap, T. (2017). A simple neural network module for relational reasoning, *Advances in neural information processing systems*, pp. 4967–4976.
- Schattschneider, D. (2010). The mathematical side of m.c. escher, *Notices of the AMS* **57**(6), 706–718.
- Shegheva, S., Goel, A. (2018). The structural affinity method for solving the raven's progressive matrices test for intelligence, *Proceedings of the AAAI Conference on Artificial Intelligence*, Vol. 32.
- Spelke, E. S. (1990). Principles of object perception, *Cognitive science* **14**(1), 29–56.
- Tipping, M. E., Bishop, C. M. (1999). Probabilistic principal component analysis, *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **61**(3), 611–622.

- Tyler, C. W. (1995). Empirical aspects of symmetry perception, *Spatial Vision* **9**(1), 1–8.
- Wagemans, J., Elder, J. H., Kubovy, M., Palmer, S. E., Peterson, M. A., Singh, M., von der Heydt, R. (2012). A century of gestalt psychology in visual perception: I. perceptual grouping and figure–ground organization., *Psychological bulletin* **138**(6), 1172.
- Zadeh, L. A. (1984). Making computers think like people [fuzzy set theory], *IEEE spectrum* **21**(8), 26–32.
- Zhang, C., Gao, F., Jia, B., Zhu, Y., Zhu, S.-C. (2019). Raven: A dataset for relational and analogical visual reasoning, *arXiv preprint arXiv:1903.02741* .
- Zhang, C., Jia, B., Gao, F., Zhu, Y., Lu, H., Zhu, S.-C. (2019). Learning perceptual inference by contrasting, *Advances in Neural Information Processing Systems*, pp. 1075–1087.

## A Dehaene's Core Geometry Tests

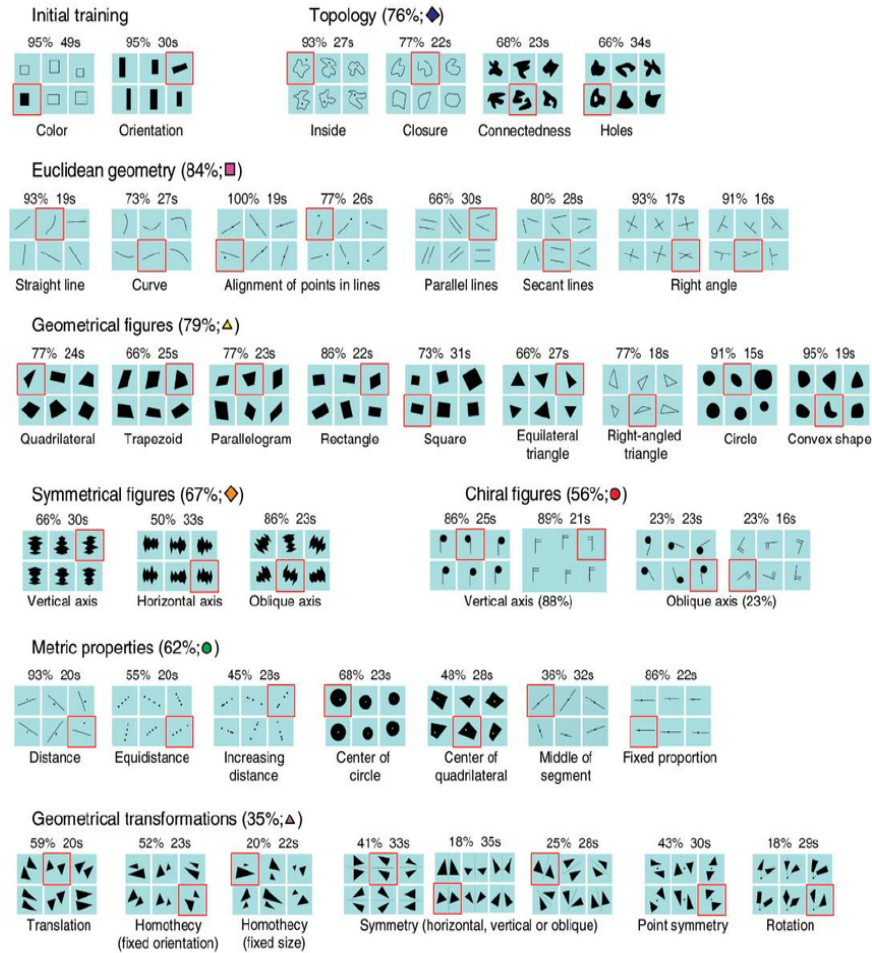


Fig. 8: Dehaene's problem set grouped by concept and presented in order of complexity from simplest (color, orientation) to the most intricate geometric transformations (homothety, symmetry and rotation). The image is retrieved from the Dehaene's research paper published in Research Gate Dehaene et al. (2006). The percentages next the concept measure the performance of Mundurukú participants in each category and individual problem.

Received August 2, 2021 , accepted August 7, 2021



© 2021. This work is published under  
<https://creativecommons.org/licenses/by-sa/4.0/>(the “License”).  
Notwithstanding the ProQuest Terms and Conditions, you may use this  
content in accordance with the terms of the License.