Kunda, M., McGreggor, K., & Goel, A. K. (2012). Reasoning on the Raven's Advanced Progressive Matrices Test with Iconic Visual Representations. In Proceedings of the 34th Annual Meeting of the Cognitive Science Society, Sapporo, Japan. pp. 1828-1833.

Reasoning on the Raven's Advanced Progressive Matrices Test with Iconic Visual Representations

Maithilee Kunda*, Keith McGreggor*, and Ashok Goel
Design & Intelligence Laboratory, School of Interactive Computing, Georgia Institute of Technology
85 Fifth Street NW, Atlanta, GA 30332 USA
{mkunda,keith.mcgreggor}@gatech.edu, goel@cc.gatech.edu
**these authors contributed equally to this work*

Abstract

Although the problems on Raven's Progressive Matrices intelligence tests resemble geometric analogies, studies of human behavior suggest the existence of two qualitatively distinct types of strategies: verbal strategies that use propositional representations and visual strategies that use iconic representations. However, all prior computational models implemented to solve these tests have modeled only verbal strategies: they translate problems into purely propositional representations. We examine here the other half of what may be a dual-process mechanism of reasoning in humans: visual strategies that use iconic representations. In particular, we present two different algorithms that use iconic visual representations to address problems found on the Advanced Progressive Matrices test, the best of which yields performances at levels equivalent to the 75th percentile for human test takers aged from 20 to 62 years-old. We discuss implications of our work for understanding the computational nature of Raven's and visual analogy in problem solving.

Keywords: Analogy; intelligence tests; knowledge representations; mental imagery; Raven's Progressive Matrices; visual reasoning.

Introduction

The Raven's Progressive Matrices (RPM) test is a standardized intelligence test. The test consists of geometric analogy problems in which a matrix of geometric figures is presented with one entry missing, and the correct missing entry must be selected from a set of answer choices. Figure 1 shows an example of a matrix problem of this kind.

There are currently three published versions of the RPM: the original Standard Progressive Matrices (SPM), the Advanced Progressive Matrices (APM), developed as a more difficult test than the SPM for individuals in high IQ ranges, and the Colored Progressive Matrices (CPM), intended as a simpler test than the SPM to be used with children, the elderly, or other individuals falling into lower IQ ranges (Raven et al., 2003). The RPM tests are considered to be the single best psychometric measures of general intelligence, outside of multi-domain IQ tests like the Wechsler scales (Snow et al., 1984), and all three versions of the RPM are widely used in clinical, educational, occupational, and scientific settings.

Neuroimaging and behavioral studies suggest that humans recruit qualitatively different strategies on the RPM regarding what types of mental representations are used, specifically in terms of visual versus verbal strategies. Visual strategies use iconic mental representations rooted in the visual perceptual modality, such as mental imagery. Verbal strategies use amodal propositional mental representations, such as linguistic description.

From factor analyses of both the SPM (Lynn et al., 2004; van der Ven & Ellis, 2000) and the APM (Dillon et al., 1981; Mackintosh & Bennett, 2005; Vigneau & Bors, 2005) as well as from fMRI data (Prabhakaran et al., 1997) comes evidence for various categories of RPM problems differentially eliciting from people either visual or verbal strategies. Studies of patients with focal brain lesions have also found linkages between brain regions associated with visual or verbal processing and successful performance on certain RPM problems (Berker & Smith, 1988; Villardita, 1985). Individuals with autism, who may exhibit a general bias towards using visual strategies over verbal ones (Kunda & Goel, 2007, 2011), tend to do particularly well on the RPM (Bölte et al., 2009; Dawson et al., 2007) and have been observed with fMRI to prefer predominantly visual strategies on the RPM (Soulières et al., 2009).

Despite this breadth of evidence for the existence of both visual and verbal RPM strategies, most computational RPM accounts have presumed to translate visual inputs into propositional representations, over which various kinds of reasoning then take place. One reason for this may be the general preponderance of propositional representations in computational accounts of cognition; in many models of visual reasoning across various task domains, visual knowledge too is represented using propositions (Carpenter et al. 1990, Lovett et al. 2010, Davies et al. 2008).

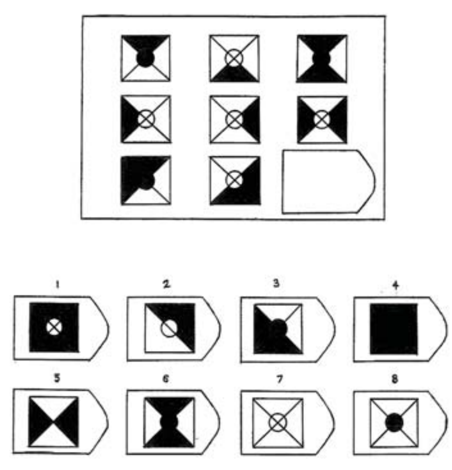

Figure 1: Example RPM Problem.

Another reason may stem from the practice of using verbal reporting protocols to study RPM problem solving. By their very nature, verbal reports are better suited to describing verbal strategies than visual strategies, which may introduce bias into the results of such protocols. Of even greater significance are findings across multiple task domains that the act of verbal reporting actually biases individuals towards using verbal strategies and/or impairs their use of visual strategies, a phenomenon known as "verbal overshadowing" (Schooler & Engstler-Schooler, 1990; Schooler et al., 1993). DeShon, Chan, and Weissbein (1995) found that a verbal reporting protocol on the APM significantly impaired accuracy on about half of the problems, and specifically on those typically solved using visual strategies.

The goal of our work is to develop computational models of a dual cognitive strategy that uses both verbal and visual representations. This first requires the development of computational models of the visual strategy itself. Once such computational models have been developed, they then may potentially be coupled with existing models of the verbal strategy. We have developed two such computational models of reasoning on the RPM using iconic visual representations. In earlier work, we tested these models against the SPM (Kunda, McGreggor, & Goel, 2010). In this paper, we apply these computational models to the APM.

In so far as we know, this work represents several firsts: it is the first report of any computational model addressing the entirety of the APM test, the first in which the problems are attempted using purely iconic visual representations, and the first to tackle the test using scanned images of each test page, without any re-rendering or representational change of inputs from those presented to a human test-taker.

Computational Accounts of the RPM

Hunt (1974) proposed the existence of two different RPM strategies that varied primarily in how problem inputs were represented. The "Analytic" algorithm used propositions to represent problems as lists of features and logical operations to evaluate rules such as constancy and addition/subtraction. The "Gestalt" algorithm, akin to mental imagery, used iconic representations and perceptual operations like continuation and superposition. However, neither algorithm was actually implemented. All of the computational RPM models that have since been developed resemble Hunt's Analytic algorithm in that they rely on a conversion of problem inputs into amodal propositional representations.

Model 1 Carpenter, Just, and Shell (1990) used a production system that took hand-coded symbolic descriptions of certain problems from the Advanced Progressive Matrices (APM) test and then selected from a set of predefined rules to solve each problem. The rules were generated by the authors from a priori inspection of the APM. The rules were experimentally validated using a verbal reporting protocol, but the potential confound of a verbal overshadowing effect was not addressed. Differences between low- and high-scoring participants were modeled by developing two different versions (FairRaven and BetterRaven) of the production system; the more advanced system (BetterRaven) contained an increased vocabulary of rules and a goal monitor. Both systems were tested against 34 of the 48 problems from the APM and solved 23 and 32 problems, respectively.

Model 2 Bringsjord and Schimanski (2003) used a theorem-prover to solve selected RPM problems stated in first-order logic, though no specific results were reported.

Model 3 Lovett, Forbus, and Usher (2010) combined automated sketch understanding with the structure-mapping analogy technique to solve SPM problems. Their system took as input problem entries sketched in Powerpoint as segmented shape objects and then automatically translated these shapes into propositional descriptions, using a sketch understanding system based on work by Biederman (1987). A two-stage structure-mapping process, following the theory of Gentner (1983), was then used to select the answer that most closely fulfilled inferred analogical relations from the matrix. This system was tested against 48 of the 60 problems on the SPM and solved 44 of these 48 problems.

Model 4 The system of Cirillo and Ström (2010), like that of Lovett et al. (2010), took as input hand-drawn vector graphics representations of test problems and automatically generated propositional representations. Then, like the work of Carpenter et al. (1990), the system drew from a set of predefined patterns, derived by the authors from an a priori inspection of the SPM, to find the best-fit pattern for a given problem. This system was tested against 36 of the 60 problems on the SPM and solved 28 of these 36 problems.

Model 5 Rasmussen and Eliasmith (2011) used a spiking neuron model to induce rules for solving RPM problems. This system took as input hand-coded vectors of propositional attribute-value pairs. While the system was said to correctly solve RPM problems, no specific results were reported.

Our Approach

As mentioned above, despite considerable differences in architecture and problem-solving focus, all five of these computational models of the RPM have reasoned over amodal propositional representations of test inputs. We believe that Raven's problems may be solved via computational models that use purely iconic visual representations of test inputs, and we present these models as a complementary view of reasoning on the RPM.

The two models that we have developed are the affine model and the fractal model, both of which use image transformations to solve RPM problems without converting the input images into any sort of propositional form. Previously, we described each of the models along with an analysis of their performance on all 60 problems from the SPM (Kunda, et al., 2010).

Iconic Visual Reasoning

The affine and fractal methods differ in important ways, but share two intuitions: comparing images under a variety of transformations, and judging the similarity based upon features which arise from the images.

Similitude Transformations

Each of our algorithms compares images (or fragments of images) under a variety of transformations. We use similitude transformations, similarity-preserving transformations which are a subset of affine transformations. Similitude transforms are a linear composition of a dilation, an orthonormal transformation, and a translation. Our implementation presently examines images under eight orthonormal transformations, specifically dihedral group D4, the symmetry group of a square. The translation is determined as a consequence of the searching each algorithm performs. The affine method restricts dilation to a value of one, i.e. no scaling, whereas the fractal method uses a short sequence of progressively smaller dilation values. Thus, the fractal method's similitude transformations are contractive.

There is evidence that human visual processing can apply some of these types of transformations to mental images, or at least operations that are computationally isomorphic in some sense. In the theory of mental imagery proposed by Kosslyn, Thompson, and Ganis (2006), transformations of mental images include scanning (i.e. translation), zooming (i.e. scaling), and rotation, among others.

A Model of Similarity

Our models must judge the similarity between images. The nature of this similarity may be determined by any number of means, many of which might associate visual or geometric features to points in a coordinate space, and compute similarity as a distance metric (Tversky 1977). Tversky developed an alternate approach by considering objects as collections of features, and similarity as a feature-matching process. We adopt Tversky's interpretation, and seek to derive a set of features for use in our matching process.

We desire a metric of similarity which is normalized, one where the value 0.0 means entirely dissimilar and the value 1.0 means entirely similar. We use the ratio model of similarity as described in (Tversky 1977), wherein the measure of similarity **S** between two representations **A** and **B** is calculated by the formula:

$$\mathbf{S(A,B) = f(A \cap B) / [f(A \cap B) + \alpha f(A\text{-}B) + \beta f(B\text{-}A)]}$$

where **f(X)** is the number of features in the set **X**. Tversky notes that the ratio model for matching features generalizes several set-theoretical models of similarity proposed in the psychology literature, depending upon which values one chooses for the weights **α** and **β**.

Although the same equation is used for similarity calculations, each of our models has its own interpretation of what constitutes a feature. In the affine method, a feature is defined to be a single pixel, and intersection, union, and subtraction operations are defined as the minimum, maximum, and difference of pixel values. This formulation assumes that pixels are independent features within the pixel sets represented by images A and B. While this notion of pixel independence is a strong simplification, it matches assumptions made by basic template theories of visual similarity that define similarity based purely on evaluations of the extent of overlapping figural units (Palmer, 1978), e.g. individual pixels. The fractal method uses features derived from different combinations of elements from the fractal representation of the image comparison (McGreggor, Kunda, & Goel, 2010).

For each base transform t_i:

- Apply $\mathbf{t_i}$ to image **A** to create image $\mathbf{t_i(A)}$.
- Search all possible translation offsets between images $\mathbf{t_i(A)}$ and **B** to find single offset **(x,y)** yielding highest similarity between them.
- Calculate similarity s between images $\mathbf{t_i(A)_{(x,y)}}$ and **B**
- For set-theoretic addition and subtraction, determine image composition operation ⊕ and operand **X** as follows:
  - If **Σ(A-B) = 0**, then ⊕ and **X** are null.
  - If **Σ(A-B) = Σ(B-A)**, then ⊕ refers to image addition and $\mathbf{X = B - t_i(A)_{(x,y)}}$.
  - If **Σ(A-B) > Σ(B-A)**, then ⊕ refers to image subtraction and $\mathbf{X = t_i(A)_{(x,y)} - B}$.

The composition transformation $\mathbf{T_i}$ is thus defined as precisely the transformation that changes image **A** into image **B**:

$$\mathbf{T_i(A) = t_i(A)_{(x,y)} \oplus X = B}$$

Algorithm 1. Inducing a composite transform

The Affine Model

Given a matrix problem, the affine model makes two basic assumptions: (a) that collinear elements are related by a composition of a similitude and/or set-theoretic transform, and (b) that parallel sets of elements share identical or analogous transforms. The model proceeds in three steps:

1) Induce a best-fit composite transform for a set of collinear elements in the matrix.
2) Apply this transform to the parallel set of elements containing the empty element; the result is a predicted answer image.
3) Compare this predicted image to the given answer choices for maximum similarity.

Algorithm 1 shows how, for a pair of images **A** and **B**, the "best-fit" composite transform is induced. The base unary transforms are the eight orthonormal symmetry transforms mentioned above (image rotations and mirrors), along with image addition (union of sets) and image subtraction (complement of sets). The base binary transforms are the five set operations of union, intersection, subtraction (both directions), and exclusive-or.

There are two places at which the affine model computes visual similarity, first in the induction of a best-fit composite transform, and second in the selection of the answer choice that most closely matches the predicted image. In addition to using Tversky's ratio model of similarity, as defined above, we also implemented a sum-squared-difference measure, which we converted to a measure of similarity (with minimum value of 0.0 and maximum value of 1.0) as:

$$\mathbf{SSDsimilarity = 1 / (1 + SSD)]}$$

These two similarity measures exhibit different behaviors. The Tversky measure privileges matches that share more pixel content. In contrast, the SSD similarity measure

Kunda, M., McGreggor, K., & Goel, A. K. (2012). Reasoning on the Raven's Advanced Progressive Matrices Test with Iconic Visual Representations. In Proceedings of the 34th Annual Meeting of the Cognitive Science Society, Sapporo, Japan. pp. 1828-1833.

effectively ignores any pixel content that is shared; similarity is calculated only as a function of pixels that are different.

Once the transformation is found that maximizes similarity, the transformation is applied to the first entry or entries in the last row or column, as shown in Figure 2. The resulting image represents the algorithm's best guess as to the missing entry. This image is compared to the answer choices, and the best match is chosen as the final answer.

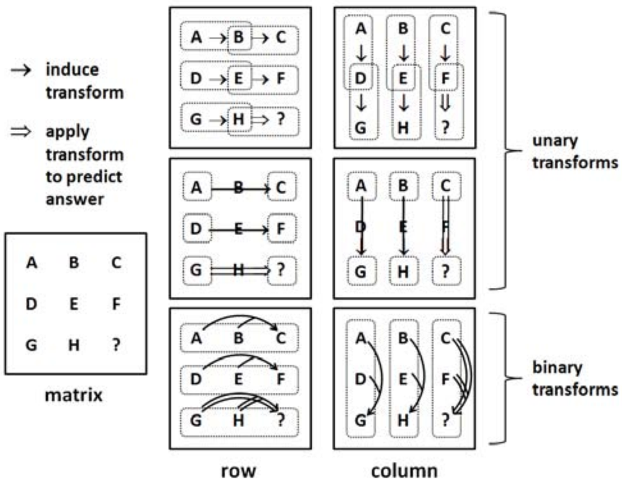

Figure 2. Sets of elements examined by the affine method

The Fractal Method

Like the affine method, the fractal method seeks to find a re-representation of the images within a Raven's problem as a set of similitude transformations. Unlike the affine method, the fractal method seeks these representations at a significantly finer partitioning of the images, and uses features derived from these representations to determine similarity for each possible answer, simultaneously, across the bulk of relationships present in the problem.

For visual analogy problems of the form **A : B :: C : ?**, each of these analogy elements are a single image. Some unknown transformation **T** can be said to transform image **A** into image **B**, and likewise, some unknown transformation **T′** transforms image **C** into the unknown answer image. The central analogy in the problem may then be imagined as requiring that **T** is analogous to **T′**. Using fractal representations, we shall define the most analogous transform **T′** as that which shares the largest number of fractal features with the original transform **T**.

To find analogous transformations for **A : B :: C : ?**, the fractal algorithm first visits memory to retrieve a set of candidate solution images **X** to form candidate solution pairs in the form **<C, X>**. For each candidate pair of images, we generate a fractal representation of the pairing from the fractal encoding of the transformation of candidate image **X** in terms of image **C**. We store each transform in a memory system, indexed by and recallable via each associated fractal feature.

First, systematically partition D into a set of smaller images, such that D = {d_1, d_2, d_3, ... }.

For each image d_i:

- Examine the entire source image S for an equivalent image fragment s_i such that an affine transformation of s_i will likely result in d_i.
- Collect all such transforms into a set of candidates C.
- Select from the set C that transform which most minimally achieves its work, according to some predetermined metric.
- Let T_i be the representation of the chosen transformation associated with d_i.

The set T = {T_1, T_2, T_3, ... } is the fractal representation of the image D.

Algorithm 2. Fractal Representation of D from S

To determine which candidate image results in the most analogous transform to the original problem transform **T**, we first fractally encode that relationship between the two images **A** and **B**. Next, using each fractal feature associated with that encoding, we retrieve from the memory system those transforms previously stored as correlates of that feature (if any). Considering the frequency of transforms recalled, for all correlated features in the target transform, we then calculate a measure of similarity.

Determining Fractal Similarity The metric we employ reflects similarity as a comparison of the number of fractal features shared between candidate pairs taken in contrast to the joint number of fractal features found in each pair member (Tversky 1977). The measure of similarity **S** between the candidate transform **T′** and the target transform **T** is calculated using the ratio model. This calculation determines the similarity between unique pairs of transforms. However, the problems from the Raven's test, even in their simplest form, poses an additional concern in that many such pairs may be formed.

Reconciling Multiple Analogical Relationships In 2x2 Raven's problems, there are two apparent relationships for which analogical similarity must be calculated: the horizontal relationship and the vertical relationship. Closer examination of such problems, however, reveals two additional relationships which must be shown to hold as well: the two diagonal relationships. Furthermore, not only must the "forward" version of each of these relationships be considered but also the "backward" or inverse version. Therefore for a 2x2 Raven's problem, we must determine eight separate measures of similarity for each of the possible candidate solutions.

The 3x3 matrix problems from the APM introduce not only more pairs for possible relationships but also the possibility that elements or subelements within the images exhibit periodicity. Predictably, the number of potential analogical relationships blooms. At present, we consider 48 of these relationships concurrently.

Relationship Space and Maximal Similarity For each candidate solution, we consider the similarity of each potential analogical relationship as a value upon an axis in a

Kunda, M., McGreggor, K., & Goel, A. K. (2012). Reasoning on the Raven's Advanced Progressive Matrices Test with Iconic Visual Representations. In Proceedings of the 34th Annual Meeting of the Cognitive Science Society, Sapporo, Japan. pp. 1828-1833.

Model	Representation	Input	Strategy	Set of APM	
				I (12 problems)	II (36 problems)
FairRaven	propositional	hand-coded propositions	prediction	7*	16*
BetterRaven	propositional	hand-coded propositions	prediction	7*	25*
Affine	iconic	scanned images	prediction	7	14
Fractal	iconic	scanned images	estimate / refine	12	26

*out of 7 and 27 problems attempted, respectively

Table 1: Affine and fractal results on the APM compared with results of the Carpenter et al. (1990) model.

large "relationship space." To specify the overall fit of a candidate solution, we construct a vector in this multidimensional relationship space and determine its Euclidean distance length. The candidate with the longest vector length is chosen as the solution to the problem.

The fractal method is described in more detail in McGreggor, Kunda, and Goel (2010, 2011).

Method

We tested our affine and fractal models on all 48 problems from the Raven's Advanced Progressive Matrices test, 12 on Set I, and 36 on Set II. To obtain visual inputs, we scanned paper copies of each test at 200 dpi and manually corrected for small (+/- 3°) rotational misalignments. Thus, the input to the models was grayscale images in the PNG format, with each image containing a single problem (matrix and answer choices).

The models used a semi-automated procedure to extract individual sub-images from each problem image. Each 3x3 problem contained 8 sub-images (plus one target blank) for the matrix entries and 8 sub-images for the answer choices.

The models were run against two variations of the test inputs: raw inputs and quantized inputs. For the raw inputs, grayscale values were extracted directly from the original PNG images, and no color correction of any kind was performed. The raw inputs contained numerous pixel-level artifacts and some level of noise. For the quantized inputs, each grayscale value was rounded to be either white or black, thus turning the inputs into pure black-and-white images as opposed to grayscale.

In addition, each model considered multiple strategies when solving the problems. The affine method used two different similarity measures (Tversky and SSD). The fractal method used three different groupings of relationships (horizontal, vertical or both).

Results

Across all input variations and strategies, the affine model correctly solved 7 of the 12 problems on Set I, and 14 of the 36 problems on Set II. These levels of performance generally correspond to the 25th percentile for both sets, for 20- to 62-year-olds (US norms) (Raven et al. 2003). Looking at input variations individually, the scores were 7 and 10 on each set for raw input, and 6 and 12 on each set for quantized input. Of the similarity measures used, the best scores were achieved using the Tversky measure on the quantized set, with scores 6 and 12 on sets I and II respectively.

Likewise, the fractal algorithm correctly solved all 12 of the problems on Set I, and 26 of the 36 problems on Set II. This level of performance corresponds to the 95th percentile for set I, and the 75th percentile for set II, for 20- to 62-year-olds (Raven et al. 2003). Looking at input variations individually, the scores were 10 and 21 on each set for raw input, and 7 and 18 on each set for quantized input. Of the groupings used by the fractal method, the best scores were achieved by considering both horizontal and vertical groupings on raw input, at 7 and 17 on sets I and II respectively.

In comparison, Carpenter et al. (1990) report results of running two versions of their algorithm (FairRaven and BetterRaven) against a subset of the APM problems. Their results, and ours, are given in Table 2. On the ones not attempted by Carpenter et al. (1990), our methods score 4 and 5 on set I (of the 5 skipped), and 4 and 7 on set II (of the 9 skipped), for affine and fractal respectively.

Discussion

We have presented two different models that use purely iconic visual representations and transformations to solve many of the problems on the Raven's Advanced Progressive Matrices test. Our results align strongly with evidence from typical human behavior suggesting that multiple cognitive factors underlie problem solving on the APM, and in particular, that some of these factors appear based on visual operations. Additionally, in so far as we know, this work is the first report of any computational model addressing the entirety of the APM test, the first in which the problems are attempted using purely iconic visual representations, and the first to tackle the test using scanned images of each test page, without any re-rendering or representational change of inputs from those presented to a human test-taker.

This robust level of performance calls attention to the visual processing substrate shared by the affine and fractal algorithms: similitude transforms as a mechanism for image manipulation, and the ratio model of similarity as a mechanism for image comparison. Of course, there are many other types of visual processing that may or may not be important for accounts of visual analogy, such as non-similitude shape transformations or image convolutions, which certainly bear further investigation.

While it has been shown (Davies et al. 2008) that visuospatial knowledge alone may be sufficient for addressing many analogy problems, the representations used in that work were still propositional. In contrast, the methods described here use only visual representations. We believe the visual methods we have presented for solving the APM can be generalized to visual analogy in other domains, such as other standardized tests (e.g. the Miller's Geometric Analogies test), as well as to tests of visual oddity. We conjecture that these methods may provide insight into general visual recognition and recall. Cognitively, we hold that these strategies are a reflection of what Davis et al. (1993) referred to as the deep, theoretic manner in which representation and reasoning are intertwined.

Acknowledgments

This research has been supported by NSF (RI) Grant #1116541, entitled "Addressing Visual Analogy Problems on the Raven's Intelligence Test," by the ONR through a NDSEG fellowship, and by the NSF GRFP fellowship program.

References

Biederman, I. 1987. Recognition-by-components: A theory of human image understanding. Psych. Rev., 94, 115-147.

Berker, E., & Smith, A. (1988). Diaschisis, site, time and other factors in Raven performances of adults with focal cerebral lesions. Int. J. Neuroscience, 38(3-4), 267-285.

Bölte, S., Dziobek, I., & Poustka, F. (2009). Brief report: The level and nature of autistic intelligence revisited. J. Autism and Developmental Disorders, 39(4), 678-682.

Bringsjord, S., & Schimanski, B. (2003). What is artificial intelligence? Psychometric AI as an answer. In IJCAI (Vol. 18, pp. 887–893).

Carpenter, P., Just, M., & Shell, P. (1990). What one intelligence test measures: a theoretical account of the processing in the Raven Progressive Matrices Test. Psychological Review, 97(3), 404-431.

Cirillo, S., & Ström, V. (2010). An anthropomorphic solver for Raven's Progressive Matrices (No. 2010:096). Goteborg, Sweden: Chalmers University of Technology.

Davies, J., Goel, A., & Yaner. P. (2008). Proteus: A theory of visual analogies in problem solving. Knowledge-Based Systems, 21(7), 636-654.

Davis, R. Shrobe H., and Szolovits, P. 1993. What is a Knowledge Representation? AI Magazine 14.1:17-33.

Dawson, M., Soulières, I., Gernsbacher, M. A., & Mottron, L. (2007). The level and nature of autistic intelligence. Psychological Science, 18(8), 657-662.

DeShon, R., Chan, D., & Weissbein, D. (1995). Verbal overshadowing effects on Raven's Advanced Progressive Matrices: Evidence for multidimensional performance determinants. Intelligence, 21(2), 135-155.

Dillon, R., Pohlmann, J., & Lohman, D. (1981). A factor analysis of Raven's Advanced Progressive Matrices freed of difficulty factors. Educational and Psychological Measurement, 41, 1295–1302.

Gentner, D. 1983. Structure-mapping: A theoretical framework for analogy. Cognitive Science 7(2), 155-170.

Hunt, E. (1974). Quote the raven? Nevermore! In L. W. Gregg (Ed.), Knowledge and Cognition (pp. 129–158). Hillsdale, NJ: Erlbaum.

Kosslyn, S., Thompson, W., & Ganis, G. (2006). The case for mental imagery. New York, NY: Oxford U. Press.

Kunda, M., and Goel, A.K. (2008). "How Thinking in Pictures can explain many characteristic behaviors of autism." In Proceedings of the 7th IEEE International Conference on Development and Learning, pp. 304-309.

Kunda, M., & Goel, A. K. (2011). Thinking in Pictures as a cognitive account of autism. J. Autism and Developmental Disorders. 41(9): 1157-1177.

Kunda, M., McGreggor, K., & Goel, A. K. (2010). Taking a look (literally!) at the Raven's intelligence test: Two visual solution strategies. In Proc. 32nd Annual Conf. Cognitive Science Society, pp. 1691-1696.

Lovett, A., Forbus, K., & Usher, J. (2010). A structure-mapping model of Raven's Progressive Matrices. In Proc. 32nd Annual Conf. Cognitive Science Society.

Lynn, R., Allik, J., & Irwing, P. (2004). Sex differences on three factors identified in Raven's Standard Progressive Matrices. Intelligence, 32(4), 411-424.

Mackintosh, N., & Bennett, E. (2005). What do Raven's Matrices measure? An analysis in terms of sex differences. Intelligence, 33(6), 663-674.

McGreggor, K., Kunda, M., & Goel, A. K. (2010). A fractal approach towards visual analogy. In Proc. 1st Int. Conf. Comp. Creativity, Lisbon, Portugal, January, 2010.

McGreggor, K., Kunda, M., & Goel, A.K. (2011). Fractal analogies: Preliminary results from the Raven's test of intelligence. In Proc. Second International Conference on Computational Creativity, Mexico City, April 2011.

Palmer, S.E. (1978). Structural aspects of visual similarity. Memory and Cognition, 6(2), 91-97.

Prabhakaran, V., Smith, J., Desmond, J., Glover, G., & Gabrieli, J. (1997). Neural substrates of fluid reasoning: An fMRI study of neocortical activation during performance of the Raven's Progressive Matrices test. Cognitive Psychology, 33(1), 43-63.

Rasmussen, D., & Eliasmith, C. (2011). A neural model of rule generation in inductive reasoning. Topics in Cognitive Science, 3(1), 140-153.

Raven, J., Raven, J. C., & Court, J. H. (2003). Manual for Raven's Progressive Matrices and Vocabulary Scales. San Antonio, TX: Harcourt Assessment.

Schooler, J., & Engstler-Schooler, T. (1990). Verbal overshadowing of visual memories: Some things are better left unsaid. Cognitive Psychology, 22(1), 36-71.

Schooler, J., Ohlsson, S., & Brooks, K. (1993). Thoughts beyond words: When language overshadows insight. J. Experimental Psychology: General, 122(2), 166-183.

Snow, R., Kyllonen, P., & Marshalek, B. (1984). The topography of ability and learning correlations. Advances in the Psychology of Human Intelligence, 2, 47-103.

Soulières, I., Dawson, M., Samson, F., Barbeau, E., Sahyoun, C., Strangman, G., et al. (2009). Enhanced visual processing contributes to matrix reasoning in autism. Human Brain Mapping. 30(12), 4082-107

Tversky, A. (1977). Features of similarity. Psychological Review, 84(4), 327-352.

van der Ven, A. & Ellis, J. (2000). A Rasch analysis of Raven's standard progressive matrices. Personality and Individual Differences, 29(1), 45-64.

Vigneau, F., & Bors, D. (2008). The quest for item types based on information processing: An analysis of Raven's Advanced Progressive Matrices, with a consideration of gender differences. Intelligence, 36(6), 702-710.

Villardita, C. (1985). Raven's Colored Progressive Matrices and intellectual impairment in patients with focal brain damage. Cortex, 21(4), 627-634.