A Standard Model of Mind:
AAAI Technical Report FS-17-05

Understanding the Role of Visual Mental Imagery in Intelligence: The Retinotopic Reasoning (R2) Cognitive Architecture

Maithilee Kunda
Department of Electrical Engineering and Computer Science, Vanderbilt University
PMB 351679, 2301 Vanderbilt Place, Nashville, TN 37235-1679, USA

Abstract

This paper presents a new Retinotopic Reasoning (R2) cognitive architecture that is inspired by studies of visual mental imagery in people. R2 is a hybrid symbolic-connectionist architecture, with certain components of the system represented in propositional, symbolic form, but with a primary working memory store that contains visual "mental" images that can be created and manipulated by the system. R2 is not intended to serve as a full-fledged, stand-alone cognitive architecture, but rather is a specialized system focusing on how visual mental imagery can be represented, learned, and used in support of intelligent behavior. Examples illustrate how R2 can be used to model human visuospatial cognition on several different standardized cognitive tests, including the Raven's Progressive Matrices test, the Block Design test, the Embedded Figures test, and the Paper Folding test.

Introduction

Consider the following question: How many windows are there in your home? Many people can answer this question after only a few moments thought. (Try it!)

Often, people use **visual mental imagery** to produce an answer to this question. They visualize the interior of their home, mentally traveling from room to room and counting windows as they go. Some people visualize their home from the outside. Other, non-imagery-based strategies are possible as well; for example, someone who has just ordered replacement windows may just know the answer straightaway, from memory. However (at least based on the author's admittedly non-random and informal polling of seminar audiences), many people seem to adopt a strategy that is at least partially, if not completely, imagery-based.

Visual mental imagery can be defined as the use of visual mental images to represent knowledge, together with visual imagery operations, such as mental rotation and translation, to reason about this knowledge. While the existence of visual mental imagery in human cognition was vigorously debated for much of the late 20th century (aptly named "The Imagery Debate"), many convergent findings in neuroscience now support the idea that visual mental imagery is a genuine and useful form of mental representation in humans (Pearson and Kosslyn 2015). Note that mental imagery can occur in other sensory modalities as well, e.g., tactile imagery (Yoo et al. 2003). However, for simplicity, discussions in this paper focus on mental imagery in the visual modality.

In people, visual mental images are represented in brain regions responsible for visual perception, where neurons are organized in a *retinotopic* fashion, i.e. the arrangement and connectivity of neurons in these brain regions preserves the 2D spatial layout of light-sensitive cells in the retina. However, mental images involve neural activations in these regions that are **not** directly tied to concurrent perceptual inputs (Slotnick, Thompson, and Kosslyn 2005). In addition, the neural activity associated with visual mental imagery has been found to play a functional role: if this neural activity is artificially suppressed, then performance on certain tasks will decrease (Kosslyn et al. 1999).

Beyond just laboratory-based evidence that humans *can* use visual mental imagery, extensive scientific and anecdotal evidence additionally suggests that humans *do* use visual mental imagery, often in highly complex and creative tasks. Both Richard Feynman (Gleick 1992) and Albert Einstein (Feist 2008) observed that they thought about physics concepts primarily using mental images and only secondarily using words and equations. Temple Grandin, a professor of animal science who is also on the autism spectrum, fixes problems in livestock facilities by performing visual mental simulations of complicated equipment (Grandin 2008). The importance of visual mental imagery has repeatedly been emphasized across numerous domains of human expertise, including surgery (Luursema, Verwey, and Burie 2012), mathematics (Giaquinto 2007), engineering (Ferguson 1994), language comprehension (Pressley 1976), programming (Petre and Blackwell 1999), scientific discovery (Nersessian 2010), and more.

While psychology and neuroscience have told us much about the behavioral and neural markers of visual mental imagery, we know much less about its *computational* properties. Many basic questions—like how mental rotation might be represented as an operator, or how imagistic information is encoded and retrieved in long-term memory—remain important open issues for scientific inquiry.

This paper presents a new Retinotopic Reasoning (R2) cognitive architecture that, inspired by studies of visual mental imagery in people, aims to work towards finding answers to these questions. R2 is a hybrid symbolic-connectionist architecture, with certain components of the system represented in propositional, symbolic form, but with

Kunda, M. (2017). Understanding the role of visual mental imagery in intelligence: The Retinotopic Reasoning (R2) cognitive architecture. Proceedings of the AAAI Fall Symposium Series (AAAI-FSS): A Standard Model of the Mind, Washington, DC. AAAI Technical Report FS-17-05. p. 401-407.

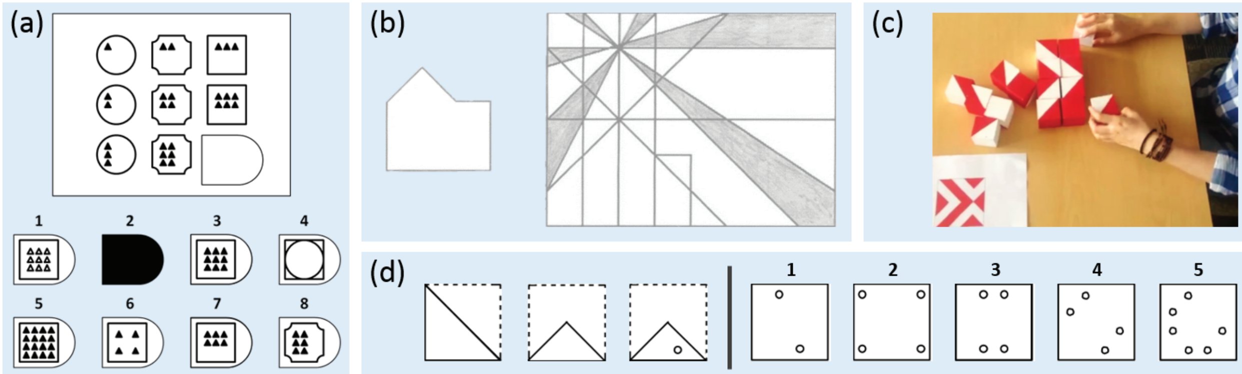

Figure 1: Example problems from the four standardized visuospatial cognitive tests: (a) Raven's Progressive Matrices test, (b) Embedded Figures test, (c) Block Design test, and (d) Paper Folding test.

a primary working memory store that contains visual "mental" images that can be created and manipulated by the system. R2 is not intended to replace non-imagery-based, propositional AI architectures, but rather to complement them, just as human intelligence relies on many different cognitive modalities.

Developing robust computational accounts of visual mental imagery will not only serve to illuminate certain aspects of human intelligence, like those described above, but also will enable AI systems to better achieve human-like and/or human-level performance on many tasks. Interactive, imagery-based AI systems have potential to augment human capabilities in these areas and to improve education and interventions for people with cognitive conditions like autism.

Existing Imagery-based AI Systems

Unlike visual perception, i.e., interpreting what is currently being seen, visual mental imagery involves creating and manipulating a new set of images in the mind. For an AI system to employ visual mental imagery, it must create and use data structures that are imagistic in nature, and use imagery-based operators to manipulate these data structures.

Most AI systems built to date do not use any visual mental imagery at all. Even AI systems that specialize in visual tasks—e.g., interpreting a visual diagram—often use images only as perceptual inputs, and store their internal knowledge using propositional (i.e., abstract, symbolic) formats. A simple example of an imagery-based representation showing the relationship between two shapes is $\square - \triangle$, while a propositional representation of the same information might look like `[LEFT-OF(TRIANGLE, SQUARE)]` (Nersessian 2010).

Examples of AI systems that do use visual mental imagery span many different problem domains, including:

- Diagrammatic/spatial reasoning (Kosslyn and Shwartz 1977; Glasgow and Papadias 1992; Tabachneck-Schijf, Leonardo, and Simon 1997).
- Geometric analogies (Kunda, McGreggor, and Goel 2013; McGreggor, Kunda, and Goel 2014).
- Naive physics (Funt 1980; Gardin and Meltzer 1989; Narayanan and Chandrasekaran 1991).
- Commonsense question answering (Bigelow et al. 2015; Lin and Parikh 2015).

In addition, there have been efforts to include visual mental imagery as a component within a larger cognitive architecture, including for the following architectures: SOAR (Lathrop, Wintermute, and Laird 2011), NEVILLE (Bertel et al. 2006), PRISM (Ragni and Knauff 2013), and Casimir (Schultheis, Bertel, and Barkowsky 2014).

The Retinotopic Reasoning (R2) cognitive architecture that is presented here is not incompatible with many of these previous efforts. R2 is designed with an initial focus on modeling low-level visuospatial cognitive operations of the type measured in people using standardized cognitive tests, with near-term research goals of improving the understanding of typical and atypical human cognitive development. However, ultimately, R2 is intended to be a modular, extensible, and continually evolving platform for broader research into visual mental imagery.

Problem Domain: Visuospatial Cognitive Tests

To date, variants of the R2 architecture have been constructed to address four different visuospatial cognitive tests: the Raven's Progressive Matrices (RPM) test, the Embedded Figures Test (EFT), the Block Design Test (BDT), and the Paper Folding Test (PFT). Example problems from these tests are shown in Figure 1.

The **Raven's Progressive Matrices** (RPM) test problems resemble geometric analogies, as shown in Figure 1a. The RPM is widely used and has been identified as the best available single-format test of general intelligence for people (Snow, Kyllonen, and Marshalek 1984). Previous human research suggested that people typically use visual mental imagery to solve only the easiest RPM problems, using verbal strategies on harder problems (Lynn, Allik, and Irwing 2004). However, a version of R2 was used to prove that visual-imagery-based approaches could solve many difficult RPM problems (Kunda, McGreggor, and Goel 2013), which is significant as one potential explanation for why some individuals on the autism spectrum show increased levels of activity in visuospatial brain regions while solving the RPM relative to neurotypical individuals (Soulières et al. 2009).

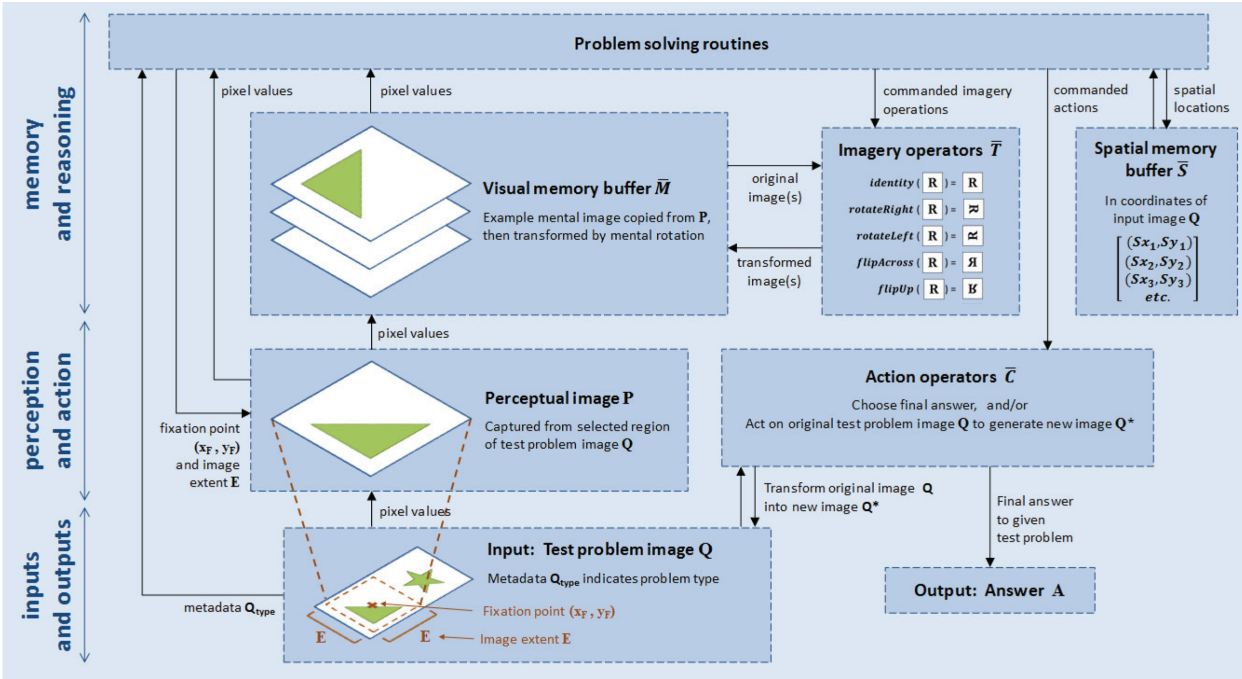

Figure 2: Overview diagram of Retinotopic Reasoning (R2) cognitive architecture.

Follow-on work has looked at how image-based representations of problem information can support problem decomposition (Kunda 2015), and also at how patterns of errors on the test may be indicative of certain cognitive strategies (Kunda et al. 2016).

The **Embedded Figures Test** (EFT) involves finding a given target inside a search environment, as shown in Figure 1b. The EFT measures a cognitive construct called *field independence*, i.e., the extent to which a person can separate a goal-related perceptual signal from a complex background, which has been linked to diverse aspects of human intelligence such as creativity and STEM achievement (Witkin et al. 1975). In previous work, a version of R2 was used to provide a new, process-level explanation of individual differences in field independence as measured by the EFT (Kunda and Ting 2016).

The **Block Design Test** (BDT) involves assembling a set of colored blocks to match a given design, as shown in Figure 1c. The BDT measures visuospatial ability and forms a part of many widely-used cognitive test batteries such as the Wechsler IQ scales. A version of R2 is currently under development to investigate how patterns of visual attention on the BDT can predict individual differences in cognitive strategies (Kunda, El Banani, and Rehg 2016).

The **Paper Folding Test** (PFT) involves mentally predicting how a series of folds and hole punches applied to a piece of paper would appear after the paper is unfolded, as shown in Figure 1d. The PFT is frequently used in research studies as a measure of visuospatial ability (Silvia 2008). A version of R2 is currently under development to investigate how visual mental imagery contributes to PFT performance, and whether patterns of errors can serve as markers for different cognitive strategies (Ainooson and Kunda 2017).

While R2 capabilities to date have been designed primarily to address the problem domain of visuospatial cognitive tests, R2 is intended to be a generalizable, extensible architecture that, in future work, will be extended to other problem domains as well.

Overview of R2

Because the Retinotopic Reasoning (R2) cognitive architecture is designed primarily to model visuospatial, and not verbal or other modalities of, cognition, **visual mental imagery is R2's primary format for representing knowledge**. What this means is that after R2 receives an input problem, typically in the form of one or more images, all problem information remains in imagistic form throughout the duration of R2's problem-solving processes.

The majority of AI systems that model similar types of visuospatial problem solving use propositional (i.e., abstract and symbolic) representations of problem information, going all the way back to Evans' seminal ANALOGY system in 1964 that solved geometric analogy problems (Evans 1964) and including more recent examples of systems that use propositional representations to solve problems from the Raven's Progressive Matrices test (Carpenter, Just, and Shell 1990; Lovett, Forbus, and Usher 2010).

R2 contains separate modules that correspond to functionally different subsystems in human cognition, such as visual memory, spatial memory, etc., as shown in Figure 2. A single R2 problem-solving episode begins when R2 receives a cognitive test problem as input, and ends when R2 generates an answer (or decides it cannot produce an answer) for that problem. Each R2 module in Figure 2 is now described in more detail, using the Raven's Progressive Matrices (RPM) problem in Figure 1a as a running example.

Input: Test problem image Q. The input to R2 is a single test problem in the form of a test problem image Q. Each Q will be accompanied by problem metadata Q_{type}, a categorical variable that identifies the test problem type (see examples in Figure 1) and is used to select the appropriate problem-solving routine. Information from Q is accessed by R2's perceptual image P. For certain cognitive tests (e.g., Block Design), R2 can change Q using action operators.

Running example: For the Raven's Progressive Matrices (RPM) example problem in Figure 1a, the input image Q would be a scanned image showing the entire problem, and Q_{type} would be set to `RPM-3x3` (RPM problem of size 3 rows by 3 columns).

Perceptual image P. R2 uses an internal data structure called the perceptual image P to access information from the test problem image Q received as input. P is analogous to the retina in human vision, in that R2 has no direct control over the contents of P. However, like human gaze, P can be pointed at different locations *within* a given test problem image Q. Then, information in P can be accessed by R2's problem-solving routines, and can also be sent to the visual memory buffer for temporary storage.

P is a fixed-size, 2D matrix of pixels and can be directed to different locations in Q by specifying a fixation point (x_F, y_F) and an image extent E (see Figure 2). In particular, let (x_F, y_F) be a fixation point defined using pixel indices in input image Q , and let E be an image extent value (defined in terms of number of pixels). Then, for perceptual image P of size $m \times m$, each pixel p_{ij} in column i and row j of P is populated with values from Q according to:

$$p_{ij} = f\left(\left[x_F - \frac{E}{2} + \frac{i}{m}E\right], \left[y_F - \frac{E}{2} + \frac{j}{m}E\right]\right) \quad (1)$$

Using the fixation point (x_F, y_F) and image extent E to specify what is essentially R2's gaze direction is analogous to a person's top-down visual attention. In people, visual attention is directed by a combination of top-down, i.e., task-driven, and bottom-up, i.e., salience-driven, factors (Buschman and Miller 2007). R2 does not currently use an explicit model of visual salience (Itti, Koch, and Niebur 1998; Bruce and Tsotsos 2006), though one will be added in future work. Salience effects are currently modeled as part of R2's test-specific problem solving routines.

Running example: During the course of solving the RPM example problem in Figure 1a, R2 would shift its gaze to "look at" various regions within the 3x3 problem matrix as well as the various answer choices. Each gaze shift would be commanded by the problem solving routine.

Visual memory buffer $\overline{M}$. People can remember visual information in many different ways. R2 models just one of these: storing visual information in the form of visual mental images. R2 has a visual memory buffer $\overline{M}$ that can store a finite number of mental images M_i. Each mental image M_i is a data structure of the same format as R2's perceptual image P but which can contain visual information that is *different* from what is currently being received through P. A mental image M_i can be created from information in P or from information in another mental image M_j, or from combinations of the two using the imagery operators in $\overline{T}$. Information in $\overline{M}$ can be populated, manipulated, or accessed at any time by R2's problem solving routines.

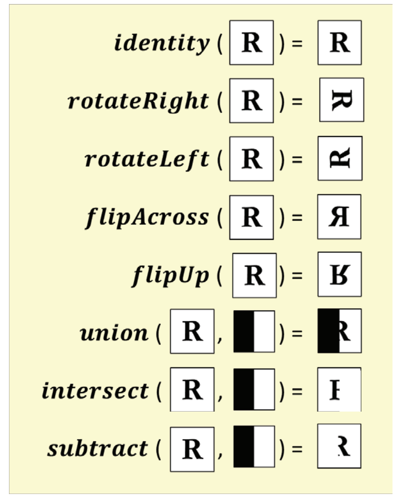

Figure 3: Examples of imagery operators in $\overline{T}$.

Running example: While solving the RPM example problem in Figure 1a, R2 might store a mental image M_1 of one matrix entry, then shift its gaze to another entry to store a second mental image M_2, and then compare the two using an imagery operator to obtain a third mental image M_3. These kinds of imagery operations can be used to solve many RPM problems (Kunda, McGreggor, and Goel 2013).

Imagery operators $\overline{T}$. R2 constructs and transforms mental images in $\overline{M}$ using a set of imagery operators $\overline{T}$. $\overline{T}$ contains several individual operators T_i, where each T_i is a function that takes one or more images as input and returns a new image as output. R2 uses its imagery operators to create new mental images using combinations of existing mental images in $\overline{M}$. Examples of operators in $\overline{T}$ are illustrated in Figure 3.

Running example: During the course of solving the RPM example problem in Figure 1a, R2 might use a visual subtraction operator to compare the middle-left matrix entry to the top-left matrix entry.

Spatial memory buffer $\overline{S}$. Just as people have different working memory buffers for visual and spatial information (Baddeley 2003), R2's spatial memory buffer $\overline{S}$ stores a list of references to individual location coordinates (Sx_i, Sy_i) in the test problem image Q.

Running example: While solving the RPM example problem in Figure 1a, the spatial memory buffer might be used to store the locations of answer choices previously considered but discarded by the current problem solving routine.

Problem solving routines. For each visuospatial cognitive test in its repertoire, R2 contains one or more test-specific problem solving routines. These routines access other R2 data structures and operators to solve test problems.

Running example: Problem solving routines for the RPM follow two different high-level strategies, based on research in humans (Bethell-Fox, Lohman, and Snow 1984). In one approach, called *constructive matching*, people first inspect the problem matrix, mentally construct an image as their candidate answer, and then find the matching answer choice. In the second approach, called *response elimination*, people first look at the answer choices and mentally plug each one into the matrix, and then choose the answer that fits best.

Additional variants include, for example, looking at rows vs. columns vs. diagonals, taking additional steps to confirm an answer after the initial guess, or even retaining information about which strategy seemed to be most successful from one problem to the next, (i.e., inter-problem learning).

Action operators $\overline{C}$. R2's action operators $\overline{C}$ can be used by R2's problem solving routines as needed. One important action in $\overline{C}$ is R2's selection of a final answer. Some tests, like the Block Design Test, involve additional actions such as manipulating blocks or other physical objects. R2 takes manual actions using test-specific operators that alter the test problem image Q.

For example, while solving a Block Design test problem, R2 has access to action operators C_i that can "move" a block from one location in the "tabletop" problem image Q to a new location, or rotate a block in place. Each action operator C_i generates a new image $Q*$ that represents the new state of the "tabletop" after the action has been performed.

To take another example, while solving an Embedded Figures test problem, R2 has access to an action operator C_i that "traces" an outline of its proposed answer onto the original test problem image Q, generating a new image $Q*$ that also serves as its answer A.

Output: Answer A. The output of R2 is its answer A to the given problem Q, or a null value if R2 is unable to produce an answer. The format of A is test-specific.

Discussion and Future Work

Results from computational experiments using R2 variants on several different visuospatial cognitive tests have been published elsewhere, for the Raven's Progressive Matrices test (Kunda, McGreggor, and Goel 2013), the Embedded Figures test (Kunda and Ting 2016), the Block Design test (Kunda, El Banani, and Rehg 2016), and the Paper Folding test (Ainooson and Kunda 2017).

Other work in progress includes investigating: 1) the use of R2 to better interpret human behaviors on cognitive tests (Kunda et al. 2016); 2) how R2 can be adapted to address the real-world task of human visual data exploration, i.e., how humans may use visual mental imagery to support the construction of conceptual narratives from multiple visual views of large datasets (Eliott, Stassun, and Kunda 2017); and 3) computational learning mechanisms to support the automated acquisition of R2 knowledge, in particular knowledge of visual imagery operators as shown in Figure 3, by watching real-world object transformations (Mel 1986).

The overall R2 research effort has a central goal of elucidating the computational building blocks of visual mental imagery and understanding how these building blocks can be composed (either in humans or by AI systems) to produce intelligent behavior. R2 is still in early stages, and there are many important aspects of visual mental imagery that have not yet been incorporated into the architecture, such as how mental images are collapsed into and then retrieved from long-term memory, which is still also an open question in human mental imagery.

Ultimately, it is the hope of the author that R2 research can be integrated with other, e.g., propositional, approaches to intelligence in computational architectures, in order to more fully capture the myriad cognitive processes that contribute to human intelligence, and to improve the state of the art in AI systems that achieve human-level and/or human-like performance on complex and difficult tasks.

R2 may also help us to better understand neurodiversity in people. For example, many individuals on the autism spectrum, like Temple Grandin (Grandin 2008) and others (Hurlburt, Happe, and Frith 1994), have expressed a preference for or bias towards thinking using visual mental imagery, as compared to neurotypical individuals, and recent research in neuropsychology has also identified individuals with aphantasia who experience little-to-no mental imagery (Zeman, Dewar, and Della Sala 2015). Better understanding qualitative variations of human intelligence, in its many different neurodiverse manifestations, not only has the potential to profoundly affect how we measure and conceptualize human cognitive abilities but also will help inform the design of improved, evidence-based practices for helping people with atypical cognitive conditions achieve positive outcomes in education, employment, and more.

Acknowledgments

This research builds on collaborative work done with A. Goel and K. McGreggor, and was originally inspired by neuropsychology research performed by M. Dawson, L. Mottron, and I. Soulières on cognition in autism.

References

Ainooson, J., and Kunda, M. 2017. A computational model for reasoning about the paper folding task using visual mental images. In *To appear in Proceedings of the 39th Annual Conference of the Cognitive Science Society, London, UK*.

Baddeley, A. 2003. Working memory: looking back and looking forward. *Nature reviews neuroscience* 4(10):829–839.

Bertel, S.; Barkowsky, T.; König, P.; Schultheis, H.; and Freksa, C. 2006. Sketching mental images and reasoning with sketches: Neville–a computational model of mental & external spatial problem solving. In *Proceedings of the 7th International Conference on Cognitive Modeling, Trieste (ICCM 2006)*, 349–350.

Bethell-Fox, C. E.; Lohman, D. F.; and Snow, R. E. 1984. Adaptive reasoning: Componential and eye movement analysis of geometric analogy performance. *Intelligence* 8(3):205–238.

Bigelow, E.; Scarafoni, D.; Schubert, L.; and Wilson, A. 2015. On the need for imagistic modeling in story understanding. *Biologically Inspired Cognitive Architectures* 11:22–28.

Bruce, N., and Tsotsos, J. 2006. Saliency based on information maximization. *Advances in neural information processing systems* 18:155.

Buschman, T. J., and Miller, E. K. 2007. Top-down versus bottom-up control of attention in the prefrontal and posterior parietal cortices. *science* 315(5820):1860–1862.

Carpenter, P. A.; Just, M. A.; and 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.

Eliott, F. M.; Stassun, K.; and Kunda, M. 2017. Visual data exploration: How expert astronomers use flipbook-style visual approaches to understand new data. In *Proceedings of the 39th Annual Meeting of the Cognitive Science Society*.

Evans, T. G. 1964. A program for the solution of a class of geometric-analogy intelligence-test questions. Technical report, DTIC Document.

Feist, G. J. 2008. *The psychology of science and the origins of the scientific mind*. Yale University Press.

Ferguson, E. S. 1994. *Engineering and the Mind's Eye*. MIT press.

Funt, B. V. 1980. Problem-solving with diagrammatic representations. *Artificial Intelligence* 13(3):201–230.

Gardin, F., and Meltzer, B. 1989. Analogical representations of naive physics. *Artificial Intelligence* 38(2):139–159.

Giaquinto, M. 2007. *Visual thinking in mathematics*. Oxford University Press.

Glasgow, J., and Papadias, D. 1992. Computational imagery. *Cognitive science* 16(3):355–394.

Gleick, J. 1992. *Genius: The life and science of Richard Feynman*. Vintage.

Grandin, T. 2008. *Thinking in pictures, expanded edition: My life with autism*. Vintage.

Hurlburt, R. T.; Happe, F.; and Frith, U. 1994. Sampling the form of inner experience in three adults with Asperger syndrome. *Psychological medicine* 24(02):385–395.

Itti, L.; Koch, C.; and Niebur, E. 1998. A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on pattern analysis and machine intelligence* 20(11):1254–1259.

Kosslyn, S. M., and Shwartz, S. P. 1977. A simulation of visual imagery. *Cognitive Science* 1(3):265–295.

Kosslyn, S. M.; Pascual-Leone, A.; Felician, O.; Camposano, S.; Keenan, J.; Ganis, G.; Sukel, K.; Alpert, N.; et al. 1999. The role of area 17 in visual imagery: Convergent evidence from PET and rTMS. *Science* 284(5411):167–170.

Kunda, M., and Ting, J. 2016. Looking around the minds eye: Attention-based access to visual search templates in working memory. *Advances in cognitive systems* 4:113–129.

Kunda, M.; Soulières, I.; Rozga, A.; and Goel, A. K. 2016. Error patterns on the raven's standard progressive matrices test. *Intelligence* 59:181–198.

Kunda, M.; El Banani, M.; and Rehg, J. M. 2016. A computational exploration of problem-solving strategies and gaze behaviors on the block design task. In *38th Annual Conference of the Cognitive Science Society, Philadelphia, USA*.

Kunda, M.; McGreggor, K.; and Goel, A. K. 2013. A computational model for solving problems from the ravens progressive matrices intelligence test using iconic visual representations. *Cognitive Systems Research* 22:47–66.

Kunda, M. 2015. Computational mental imagery, and visual mechanisms for maintaining a goal-subgoal hierarchy. In *Proceedings of the Third Annual Conference on Advances in Cognitive Systems ACS*, 4.

Lathrop, S. D.; Wintermute, S.; and Laird, J. E. 2011. Exploring the functional advantages of spatial and visual cognition from an architectural perspective. *Topics in cognitive science* 3(4):796–818.

Lin, X., and Parikh, D. 2015. Don't just listen, use your imagination: Leveraging visual common sense for non-visual tasks. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2984–2993.

Lovett, A.; Forbus, K.; and Usher, J. 2010. A structure-mapping model of ravens progressive matrices. In *Proceedings of CogSci*, volume 10, 2761–2766.

Luursema, J.-M.; Verwey, W. B.; and Burie, R. 2012. Visuospatial ability factors and performance variables in laparoscopic simulator training. *Learning and individual differences* 22(5):632–638.

Lynn, R.; Allik, J.; and Irwing, P. 2004. Sex differences on three factors identified in raven's standard progressive matrices. *Intelligence* 32(4):411–424.

McGreggor, K.; Kunda, M.; and Goel, A. 2014. Fractals and ravens. *Artificial Intelligence* 215:1–23.

Mel, B. W. 1986. A connectionist learning model for 3-d mental rotation, zoom, and pan. In *Proceedings of the Eighth Annual Conference of the Cognitive Science Society*, 562–71.

Narayanan, N. H., and Chandrasekaran, B. 1991. Reasoning visually about spatial interactions. In *IJCAI*, 360–365.

Nersessian, N. J. 2010. *Creating scientific concepts*. MIT press.

Pearson, J., and Kosslyn, S. M. 2015. The heterogeneity of mental representation: ending the imagery debate. *Proceedings of the National Academy of Sciences* 112(33):10089–10092.

Petre, M., and Blackwell, A. F. 1999. Mental imagery in program design and visual programming. *International Journal of Human-Computer Studies* 51(1):7–30.

Pressley, G. M. 1976. Mental imagery helps eight-year-olds remember what they read. *Journal of Educational Psychology* 68(3):355.

Ragni, M., and Knauff, M. 2013. A theory and a computational model of spatial reasoning with preferred mental models. *Psychological review* 120(3):561.

Schultheis, H.; Bertel, S.; and Barkowsky, T. 2014. Modeling mental spatial reasoning about cardinal directions. *Cognitive science* 38(8):1521–1561.

Silvia, P. J. 2008. Another look at creativity and intelligence: Exploring higher-order models and probable confounds. *Personality and Individual differences* 44(4):1012–1021.

Slotnick, S. D.; Thompson, W. L.; and Kosslyn, S. M. 2005. Visual mental imagery induces retinotopically organized activation of early visual areas. *Cerebral cortex* 15(10):1570–1583.

Snow, R. E.; Kyllonen, P. C.; and 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. B.; Sahyoun, C. P.; Strangman, G. E.; Zeffiro, T. A.; and Mottron, L. 2009. Enhanced visual processing contributes to matrix reasoning in autism. *Human Brain Mapping* 30(12):4082–4107.

Tabachneck-Schijf, H. J.; Leonardo, A. M.; and Simon, H. A. 1997. CaMeRa: A computational model of multiple representations. *Cognitive Science* 21(3):305–350.

Witkin, H. A.; Moore, C. A.; Goodenough, D. R.; and Cox, P. W. 1975. Field-dependent and field-independent cognitive styles and their educational implications. *ETS Research Report Series* 1975(2):1–64.

Yoo, S.-S.; Freeman, D. K.; McCarthy III, J. J.; and Jolesz, F. A. 2003. Neural substrates of tactile imagery: a functional mri study. *Neuroreport* 14(4):581–585.

Zeman, A.; Dewar, M.; and Della Sala, S. 2015. Lives without imagery-Congenital aphantasia. *Cortex* 73:378–380.