Ferguson, R. W., & Forbus, K. D. (1998). Telling juxtapositions: Using repetition and alignable difference in diagram understanding. In K. Holyoak, D. Gentner, & B. Kokinov (Eds.), *Advances in Analogy Research* (pp. 109-117). Sofia: New Bulgarian University.

Telling juxtapositions: Using repetition and alignable difference in diagram understanding

Ronald W. Ferguson **Kenneth D. Forbus**
The Institute for the Learning Sciences
Northwestern University
Evanston, IL 60201 USA

Abstract

Diagrams often use repetition to convey points and establish contrasts. This paper shows how MAGI, our model of repetition and symmetry detection, can model the cognitive processes humans use when reading repetition-based diagrams. MAGI, which is based on the Structure Mapping Engine, detects repetition by aligning both visual and conceptual relational structure. This lets visual regularity of form support an understanding of the conceptual regularity such forms often depict. We describe JUXTA, which uses this insight to critique a class of diagrams that juxtapose similar scenes to demonstrate physical laws.

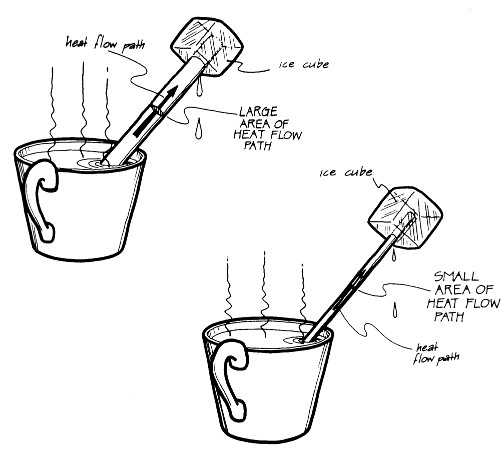
Introduction

In explanatory diagrams, repeated structures often have special significance. To underscore a point or emphasize a difference, diagrams often juxtapose events, scenes, or objects. Examples include a "before and after" display of shirts in a laundry detergent ad and a point-by-point comparison of pumps in a physics text. In such cases, visual repetition heightens contrasts and encourages deeper comparisons. This effect is an instance of what we have termed *analogical encoding* (Ferguson, 1994), because it uses repetition and symmetry detection to support other reasoning processes.

Diagram designers have long known the utility of repetition. Edward Tufte writes that repeating structure "takes advantage of our notable capacity to compare and reason about multiple images that appear simultaneously within our eyespan. We are able to canvas, identify, reconnoiter, select, contrast, review—ways of seeing quickened and sharpened by the direct spatial adjacency of parallel elements." (Tufte, 1997, p. 80). Repetition, detectable at a glance, aids the reader in exploring, and thus understanding, a diagram.

An example illustrates this point. Figure 1 is from a solar energy text (Buckley, 1979). This diagram illustrates a principle of heat transfer by juxtaposing two scenarios. In these scenarios, heat flows from a hot liquid, along an immersed metal bar, to a melting ice cube. Because heat flows faster in the leftmost scene, its ice cube melts more quickly. This difference between the scenarios shows how increasing a conductor's cross area increases heat transfer.

The diagram uses repetition to good effect. The two scenarios not only contain the same physical elements, but are also visually similar. Before understanding the processes or the physical objects, the diagram reader may sense this visual "echo", which divides the diagram into two parts. This division signals the reader that these two parts are to be compared. Then, the visual correspondence of similar shapes supports the conceptual correspondence of the two cups, two bars, and two heat flows that are key to understanding the point in the caption.



Thick Bar Conducts More Heat

Figure 1: A diagram from Buckley (1979).

If the designer had arranged the two scenes to be similar in conceptual but not visual terms—if, for example, the cup and ice cube were shaped or arranged differently—the reader could still understand the diagram. But she might not instantly recognize the implicit comparison, as before. The diagram's visual repetition allows its conceptual comparison to be quickly grasped.

This diagram is also designed so that all differences are relevant. The sole differences in the diagram are the greater thickness of the left metal bar, and the greater volume of water dripping from the left ice cube. These differences are tied to point of the caption: "Thick bar conducts more heat." The thicker bar is the independent variable, and the increased melting visibly indicates the greater heat flow.

Other differences could have been allowed. The cups could differ in volume or height, or the metal bars could differ not just in thickness, but in length. Intuitively, however, such differences would make the diagram less clear. As Tufte notes, "[i]nformation consists of differences that make a difference." (1997, p. 65)

The two repetition-based techniques used by this diagram—using visual regularity support a conceptual comparison, and limiting differences to only those relevant to the diagram's point—are our starting point for a cognitive model of how humans comprehend repetition in diagrams.

Structural alignment processes in diagrammatic reasoning

Why should visual repetition aid diagram comprehension? How does difference contribute to understanding? We believe the answers may lie in structure-mapping processes.

Our explanation involves two models. The first, MAGI, is a model of repetition and symmetry detection which links regularity detection with analogical mapping. The second, Markman and Gentner's *alignable difference* model, shows how difference detection depends on structural alignment. Based on these two models, we describe three diagram design defects that occur in repetition-based diagrams.

MAGI

Similarity and analogical comparison can be modeled as the structural alignment of propositional descriptions. (Falkenhainer, Forbus, & Gentner, 1989; Forbus, Ferguson, & Gentner, 1994; Gentner, 1983; Gentner, 1989; Goldstone, 1994; Holyoak & Thagard, 1989; Keane & Brayshaw, 1988).

MAGI (Ferguson, 1994, In preparation) is the first model linking regularity detection with similarity. MAGI is based on the idea that symmetry and repetition (both visual and conceptual) can be viewed as a similarity mapping between a description and itself. Using an extension of the Structure Mapping Engine (SME; Falkenhainer et al., 1989; Forbus et al., 1994), MAGI uses structural alignment to detect regularity within a single description. Like SME, MAGI's mapping process is computationally tractable because it operates in a local-to-global fashion. Individual alignments are constructed in parallel and then aggregated into global mappings, mappings governed by systematicity constraints favoring relationally deep, interconnected correspondence sets. MAGI also operates incrementally. As new information is added to a description, MAGI's mapping can be extended appropriately.

To detect regularity, MAGI maps over a visual representation built by GeoRep. GeoRep, given a line drawing, builds a propositional description of its salient perceptual relations. Starting with the drawing's graphical primitives (line segments, arcs, circles, ellipses, and spline curves), a set of visual routines (Ullman, 1984) represent a variety of relationships, including types of object connection, parallelisms, horizontal and vertical relations, and descriptions of polygons and their inflexion points. GeoRep contains a rule engine, and its default rule set can be extended to handle particular domains.

Given a stylized line drawing of our example diagram (Figure 2), MAGI can map over the diagram's perceptual relations to determine object correspondences (Figure 3). If we add information about the physical objects and processes in the diagram, MAGI can extend its mapping accordingly.

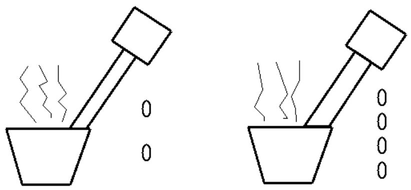


Figure 2: Stylized redrawing of Figure 1

MAGI can helps explain the immediacy and utility of visual regularity. It describes how repetition is detected and the nature of the correspondences produced. More importantly, however, it provides a link between perceptual and conceptual regularity.

Based on MAGI's model, we assume the reader of the diagram begins by detecting the its visual regularity (Figure 3). As conceptual information is also acquired, the reader may attempt to use this information to extend the mapping. However, if the new conceptual information cannot be mapped consistently with the visual information, the reader may either fail to notice the conceptual regularity, or need to ignore the previous visual regularity. Handling this conflict may slow or block diagram comprehension.

Visual repetition and symmetry detection operate very early in perception. Visual symmetry can be detected after display times of less than 100 ms. (Carmody, Nodine, & Locher, 1977; Corballis & Roldan, 1975; Julesz, 1971). Consequently, most models of symmetry detection do not incorporate more complex algorithms such as structural alignment (with the notable exception of the Wageman's Bootstrapping model (1995)). Until recently, it seemed unlikely that visual symmetry detection could involve alignment.

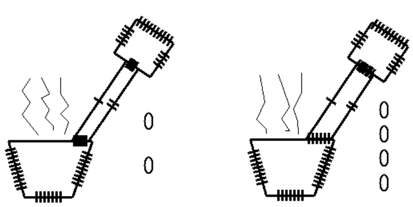


Figure 3: Regularity found by MAGI

However, new results from Aminoff, Ferguson and Gentner (In preparation; 1996) provide evidence that even the earliest forms of symmetry detection may involve alignment. In two experiments, Aminoff *et al.* (in preparation) showed subjects symmetric and asymmetric polygons with display times of 50 ms. In each experiment, subjects were consistently faster or more accurate at judging the asymmetry of polygons containing aligned qualitative differences, including differences in corner concavity and number of vertices. This effect was independent of several other quantitative asymmetry measures, including differences in area and radial length. Thus, these results are new evidence for alignment early in symmetry detection. For this reason, it is entirely possible that structural alignment is used for both very early and much later forms of regularity detection.

Alignable differences in comparison

The MAGI model explains how visual repetition can support an understanding of conceptual repetition. However, we have not yet addressed the utility of differences in a diagram.

Of course, difference detection might be seen as a very different process than repetition detection. In Tversky's influential contrast model of comparison (1977), similarity increases as a function of common features, and decreases as a function of mismatched features. If individual features are assumed independent, the detection of matched features would neither encourage nor block the detection of mismatched features.

Studies by Markman and Gentner, however, found evidence that alignment significantly affected the kinds of differences human participants noticed, with most differences directly linked to preexisting aligned commonalities (and thus called *alignable differences*). This model predicts that increasing the similarity of two concepts also increases the number of alignable differences noticed. This prediction was borne out in their experiments. When human participants were asked to list differences between high and low similarity word pairs (Markman & Gentner, 1993), participants consistently listed more alignable differences for pairs with high similarity (hotels and motels) than for pairs with low similarity (magazine and kitten). A second set of experiments (Markman & Gentner, 1996), generalized the results for word pairs to pairs of pictures, and also showed that alignable differences had a greater effect on participants' judgment of similarity than did nonalignable differences. When determining differences between two things, people seem to focus more on alignable than nonalignable differences.

Because alignable differences are produced more often than nonalignable differences, and because they have a greater influence on participants' similarity judgments, it is safe to assume that alignable differences are critical to the contrasts undertaken in repetition-based diagrams. Because alignable differences are easily generated in the context of structural alignment, visual alignable differences may communicate their points very effectively.

We conjecture that structural alignment has a profound effect on diagram understanding. Visual alignment supports conceptual alignment, and also highlights alignable differences.

Three problems of diagram style

Which factors—by analogy with understanding written prose—make a diagram more comprehensible? As we have seen, repetition in diagrams should be visually apparent, and should draw the reader into a deeper conceptual comparison without causing missteps or misalignments. Alignable differences should be salient and should serve the point of the diagram. These criteria suggest three general types of design defects that may hinder comprehension of repetition-based diagrams.

Visual/conceptual cross-mappings. Cross-mappings (Gentner & Toupin, 1986) occur when surface information and relational information suggest different mappings for the same objects. Visual cross-mappings occur when two objects are visually alignable, but the roles or functions of the aligned objects are not equivalent. For example, if two oblong objects match, but one is a metal bar conducting heat, and another the handle of a container, the initial visual correspondence between the parts might confuse readers. The readers might seek some common functional role between the two objects, and find none, slowing them down.

Alignable differences that are either not salient or not compelling. Some alignable differences are more noticeable than others. In our example diagram, for instance, many people find the difference in the number of water droplets easier to spot than the difference in thickness for the two metal bars.

We do not yet have a theory of what makes alignable differences salient or compelling. Understanding salience alone requires a more complex model of visual attention than we have available. However, we can define techniques to make alignable differences either more salient or more compelling, a process we call *difference amplification*.

We can make alignable differences more salient by either adding additional alignable structure that draws attention to that difference, or by other techniques, such as color. Besides making differences more salient, we can also make them more compelling by making the importance of the difference more evident. We do this by making it easier for the diagram reader to link the visual alignable differences to the conceptual differences underlying the diagram's point. Labeling is the easiest way to accomplish this.

Aligned differences unrelated to, or interfering with, the point of diagram. When alignable differences exist, they should be related to the diagram's point. Some alignable differences may be irrelevant; if our diagram had one cup colored red, and the other blue, this difference would be obvious but unlikely to confuse the reader. Alignable differences may detract from a diagram when they appear to be related to the point of the diagram, but are not. If one cup was being heated with a burner in our diagram, we might be confused about how this particular difference relates to the role of the thicker metal bar, since both the flame and relative bar thickness would affect the rate of heat flow.[1] Such alignable

[1] Tufte (1997) gives an example of how this principle is violated in the "before and after" drawings done by the 19th century architect Humphrey Repton. Repton's "after" views often embellish. For example, a landscaping proposal adds changes to the "after" view that are appealing but are unrelated to the proposed modification, such as stylishly

differences make it more difficult to draw a conclusion from the diagram, and thus hinder the reader's ability to comprehend the point of the juxtaposed situations.

To summarize, the MAGI model and Markman and Gentner's alignable difference model suggest three ways in which a diagram can be confusing. First, it may contain a visual-conceptual cross-mapping. Second, alignable differences may not be salient. Finally, alignable differences may be irrelevant or may interfere with the point of the diagram. These three criteria can be easily characterized in terms of the MAGI model and some simple assumptions about visual representation.

Because these stylistic problems can be cleanly described in terms of the MAGI model, it is possible to build a diagram critic that uses these principles to parse and critique diagrams. We can use mismatches between correspondences at the visual, physical and process levels to determine how well the visual regularity in the figure guides the comparison. If we have a representation of the diagram's point (which often can be derived from the caption), we can also determine if the alignable differences in the figure convey the point, are orthogonal to the point, or get in the way of understanding the point.

We have built such a system, called JUXTA[2]. Given diagrams that juxtapose physical situations, JUXTA can produce a critique of the figure, and note differences that may confuse or distract the reader. JUXTA also amplifies a diagram's relevant alignable differences by labeling them, using its physical knowledge to create and place useful explanatory labels.

We now summarize how JUXTA works.

The JUXTA Architecture

Figure 4 describes JUXTA's architecture. JUXTA's input is a stylized line drawn diagram and a representation of the diagram's caption. It provides three kinds of feedback. First, it amplifies relevant alignable differences by labeling them with process descriptions. Second, it critiques differences that interfere

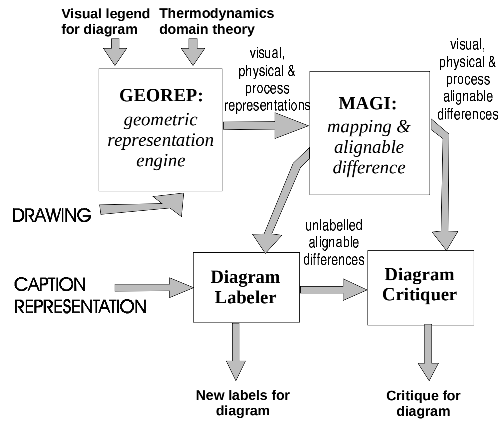


Figure 4: JUXTA's architecture

with the point of the diagram (as given in the caption). Finally, it notes differences that are orthogonal to the point of the diagram, and thus may be removed at the designer's discretion.

Processing the figure

First, JUXTA (using the GeoRep visual representation engine) represents the diagram at three different levels—visual level (e.g. a square), a physical level (an ice cube), and a physical process level (heat flowing into an ice cube) using a set of rules and low-level visual routines (Figure 5).

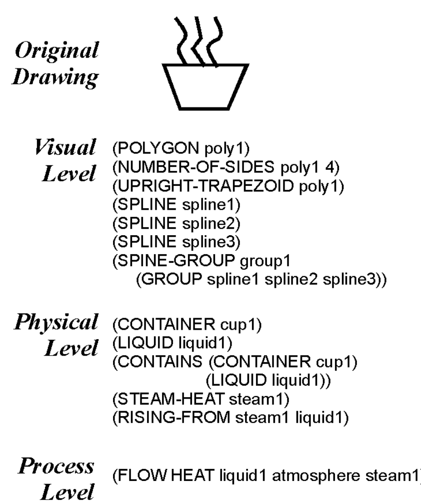


Figure 5: Levels of representation

dressed people on the sidewalks and fine sailing ships in the adjacent harbor.

[2] JUXTA stands for Juxtaposition Understanding and eXplanation Through Analogy.

Class of object	Visual legend	Salient dimensions
Container of liquid	Upright, top-heavy trapezoid	Height and width
Steam or heat	Group of proximate spline curves	Number of curves
Metal bar	Oblong oblique trapezoid	Length and thickness
Ice cube	Square	Width
Water drops	Group of proximate, vertical ellipses	Number of ellipses

Table 1: Visual legend for recognized objects

JUXTA uses a simplified model of object recognition, which depends on a set of rules to determine when a set of visual entities represent a particular type of structured object. The heuristics used for object recognition are summarized in Table 1. This technique requires the use of stylized diagrams, but otherwise retains much of the flexibility of general diagrams. For example, objects can be drawn using a drawing program, object dimensions can vary as needed, and diagram parts can be composed into more comprehensive scenes

Of course, JUXTA also needs a representation of the caption, which is assumed to contain the point of the diagram. To avoid doing natural language interpretation, we give JUXTA the representation of the caption directly. The representations use Qualitative Process Theory (Forbus, 1984). It is useful to identify two parts of captions for juxtaposition diagrams, the *antecedent* and *consequent*. In this caption, the antecedent is the difference in thickness of the bars and the consequent is the difference in the rates of heat flow.

Finding regularity and differences

JUXTA runs MAGI on the figure to detect correspondences (Figure 3). JUXTA then uses a simple mechanism for detecting alignable differences based on finding differences in dimensions predetermined by the object category. For example, when two trapezoids correspond, JUXTA compares their height and length. Invisible differences, such as differences in the rate of a physical process, are inferred from visual differences via rules in a domain-dependent knowledge base. For example, if the two trapezoids represent two cups, and one trapezoid is larger, then JUXTA infers that the cup represented by that trapezoid has greater volume. This way, visible differences enable JUXTA to infer deeper conceptual differences.

Amplifying differences via labeling

At this point, JUXTA now has analyzed the figure at the visual, physical, and process levels. It also has, for each of those levels, computed the representation of that level, the regularity mapping for that level, and the set of aligned differences. It can now begin its critical analysis of the figure. First, JUXTA attempts to link the aligned differences to the antecedents and consequences of the point given in the caption. It then amplifies the aligned differences by labeling them. The labels link each key difference in the caption to some visual difference.

To link the objects in the diagram with the referents in the caption representation, JUXTA matches the caption representation against the physical and process representations of the diagram, and uses this match to fill the caption representation's unfilled slots. This is how, for example, JUXTA figures out which object or objects the caption's "thicker bar" refers to. In this case, JUXTA can find the thicker bar on the right using the common object category (metal bar) to select both metal bars, and using the alignable difference (thicker) to distinguish between them.

Once JUXTA understands which parts of the figure are being referenced in the caption, it labels the differences. This involves constructing paired labels for each alignable difference given in the caption, and then determining where to place each label. To label an alignable difference, JUXTA must find a visible referent to point to. When an alignable difference is along a visible dimension (such as the thickness of a bar), the object itself is the referent of the label, and JUXTA points to the shape which represents the physical object. Alternatively, when a caption relationship is not visible (such as heat flow along the metal bar), JUXTA looks for a consequence of the relationship which is visible difference. In the example figure, the difference in heat flow causes a difference in the

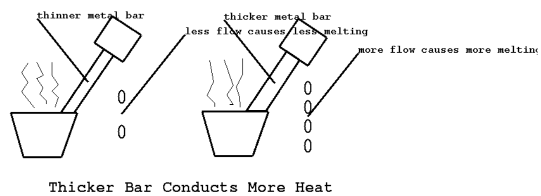


Figure 6: Results of labeling stage of JUXTA on example diagram

rate at which the ice cube melts, causing a visible difference in the number of drops (ellipses), so JUXTA labels this. The result of the labeling stage on the example diagram is given in Figure 6.

Critiquing the diagram

After labeling the figure, JUXTA critiques how well the alignable differences contribute to the point of the caption. To do this, JUXTA looks at all alignable differences left over from the labeling stage. These are differences that are not related to alignable differences referenced in the caption. If a remaining alignable difference is not the result of the caption antecedent, but can have an effect on its consequent, JUXTA notes it as potentially confusing. For example, Figure 7 is a variant of our example diagram that contains this problem. Here, the amount of heat rising from the second cup is larger than the first container. JUXTA notes this difference as confusing because the amount of heat from the container implies that the second container may contain a hotter liquid, which would also increase the heat flow rate.

Of course, remaining alignable differences may not relate to the caption at all. In this case, JUXTA will not mark it as confusing, but will note the orthogonal status of the alignable difference. For example, in Figure 7, JUXTA will note that one spline curve in the leftmost group is longer. Removing this differences might make interpretation somewhat simpler, but it will not cause problems if left unchanged.

Conclusion

Analogical encoding techniques, based on current models of analogy and similarity, can provide key insights into diagrammatic

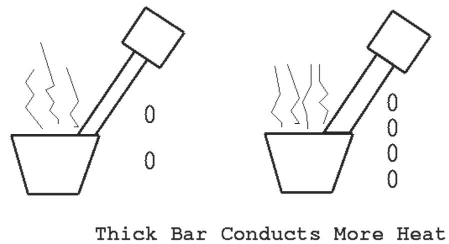


Figure 7: A faulty variant of Figure 1

reasoning. We have shown how MAGI, which uses structure mapping to detect repetition and symmetry, may explain how visual and conceptual regularity support one another, and how alignable differences emphasize relevant points. This model is strong enough to build a system, JUXTA, that can parse, analyze, and critique a diagram by analyzing how correspondences and differences interact between the visual, physical and process levels.

While JUXTA demonstrates the basic principles behind a whole class of diagrammatic reasoners, the current implementation is limited. JUXTA has only been used on a handful of figures. The recognition of objects and processes remains brittle. We are exploring similarity-based feature re-interpretation as one mechanism for improving the system's flexibility.

JUXTA also deals solely with diagrams that use binary repetition to demonstrate physical laws. In practice, diagrams use many types of regularity, including matrices, multiple repeated items, sequences, and symmetry. To expand the kinds of regularity JUXTA handles, MAGI itself may need to be extended to handle some forms of n-ary symmetry and repetition. This problem relates more to psychology than programming—although it is relatively simple to configure a version of MAGI that recognizes small multiples of a scene, it does not yet do so in an efficient way, nor does it reflect our understanding of how humans recognize other forms of regularity. We expect, however, that JUXTA soon handle symmetry as well as repeating diagrams.

The just-mentioned variety of diagrammatic regularity speaks to the fascinating richness of this particular sub-area of cognition. If the simple mechanisms of JUXTA can be extended to a larger range of diagrams, they may not only provide a foundation for computer systems that can understand diagrams in a more human-like fashion, but may also have interesting consequences for our understanding of diagrammatic reasoning, regularity, and analogy.

Acknowledgements

Thanks to Dedre Gentner and Alex Aminoff for many helpful comments on this research. This research was supported by grants from the Office of Naval Research and DARPA.

References

Aminoff, A., Ferguson, R. W., & Gentner, D. (In preparation). Early detection of qualitative symmetry.

Buckley, S. (1979). *Sun Up to Sun Down*. New York: McGraw-Hill.

Carmody, D. P., Nodine, C. F., & Locher, P. J. (1977). Global detection of symmetry. *Perceptual and Motor Skills*, *45*, 1267-1273.

Corballis, M. C., & Roldan, C. E. (1975). Detection of symmetry as a function of angular orientation. *Journal of Experimental Psychology: Human Perception and Performance, 1*, 221-230.

Falkenhainer, B., Forbus, K. D., & Gentner, D. (1989). The Structure-Mapping Engine: Algorithm and examples. *Artificial Intelligence, 41*, 1-63.

Ferguson, R. W. (1994). MAGI: Analogy-based encoding using symmetry and regularity. In *Proceedings of the 16th Annual Conference of the Cognitive Science Society* (pp. 283-288). Atlanta: Erlbaum.

Ferguson, R. W. (In preparation). MAGI: A model of symmetry and repetition detection. .

Ferguson, R. W., Aminoff, A., & Gentner, D. (1996). Modeling qualitative differences in symmetry judgments. In *Proceedings of the 18th Annual Conference of the Cognitive Science Society*. Hillsdale: Erlbaum.

Forbus, K. D. (1984). Qualitative Process Theory. *Artificial Intelligence, 24*, 85-168.

Forbus, K. D., Ferguson, R. W., & Gentner, D. (1994). Incremental structure mapping. In *Proceedings of the 16th Annual Conference of the Cognitive Science Society* (pp. 313-318). Atlanta, GA: Lawrence Erlbaum Associates.

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

Gentner, D. (1989). The mechanisms of analogical learning. In S. Vosniadou & A. Ortony (Eds.), *Similarity and Analogical Reasoning* (pp. 199-241). London: Cambridge University Press.

Gentner, D., & Toupin, C. (1986). Systematicity and surface similarity in the development of analogy. *Cognitive Science, 10*, 244-300.

Goldstone, R. L. (1994). Similarity, interactive activation, and mapping. *Journal of Experimental Psychology: Memory and Cognition*, *20*, 3-28.

Holyoak, K. J., & Thagard, P. (1989). Analogical mapping by constraint satisfaction. *Cognitive Science, 13*, 295-355.

Julesz, B. (1971). *Foundations of Cyclopean Perception*. University of Chicago Press.

Keane, M., & Brayshaw, M. (1988). The Incremental Analogy Machine. In *Proceedings of the Third European Working Session on Learning* (pp. 53-62). London: Pitman.

Markman, A. B., & Gentner, D. (1993). Splitting the differences: A structural alignment view of similarity. *Journal of Memory and Language, 32*(4), 517-535.

Markman, A. B., & Gentner, D. (1996). Commonalities and differences in similarity comparison. *Memory and Cognition, 24,* 235-249.

Tufte, E. R. (1997). *Visual Explanations*. Cheshire, CT: Graphics Press.

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

Ullman, S. (1984). Visual routines. In S. Pinker (Ed.), *Visual Cognition* (pp. 97-159). Cambridge: MIT Press.

Wagemans, J. (1995). Detection of visual symmetries. *Spatial Vision, 9*(1), 9-32.