SRRI #125

METHODS FOR EVALUATING THE VALIDITY OF HYPOTHESIZED ANALOGIES

John Clement

Department of Physics and Astronomy
University of Massachusetts/ Amherst

June 12, 1986

In the Proceedings of the Annual Meeting of the Cognitive Science Society, August 15, 1986, pp. 223-234. Hillsdale, NJ: Lawrence Erlbaum Associates.

Preparation of this paper was supported by a grant from the National Science Foundation #MDR8470579. Any opinions, findings, and conclusions or recommendations expressed in this publication are those of the authors and do not necessarily reflect the views of the National Science Foundation.

METHODS FOR EVALUATING THE VALIDITY OF HYPOTHESIZED ANALOGIES*

John Clement

Department of Physics and Astronomy
University of Massachusetts/ Amherst

ABSTRACT

Evidence is presented indicating that spontaneously generated analogies can play a significant role in expert problem solving. Since not all analogies are valid, it is important for the subject to have a way to evaluate their validity. Three methods for evaluating analogical validity are identified using observations from thinking aloud problem solving protocols as well as examples from Newton and Galileo. In particular, this paper focuses on an evaluation strategy called bridging that has been observed in solutions to both science and mathematics problems. In constructing a bridge, the subject finds an intermediate case that is seen as "in between" the analogous case and the problem situation because it shares important features of both. Many of the bridges observed appeared to be novel inventions created by the subject. These empirical studies have led to the construction of a more detailed theory for how analogies can be used effectively in instruction. Some of the strategies observed in experts appear to have high potential for helping science students overcome persistent misconceptions in the classroom.

A number of authors have emphasized the important role of analogies in problem solving and learning, including Gentner and Gentner (1983), Rummelhart and Norman (1980), and Gick and Holyoak (1980). Expert subjects have been observed to generate and use analogies spontaneously during problem solving (Clement, 1981). However, since not all analogies are valid, it is important for the subject to have a way to evaluate their validity. This paper identifies methods used to evaluate and establish confidence in the validity of a hypothesized analogy. "Validity" is used here in a weak sense outside the context of deductive certainty. Since conclusions reached by analogy are viewed as always having a certainty level of less than 100%, establishing validity here means "raising confidence in the appropriateness of the analogy to a high level."

The focus in the study of experts is on qualitative observations from case studies and coarse grained modeling of underlying processes. The primary purpose is to identify important reasoning strategies that occurred across different subjects and problems and to propose an initial description of their form and function. This study contrasts with other studies where the subject is presented with all or part of an analogy and is given the opportunity to use it or complete it since the analogies studied here were not suggested by the experimenter, but were generated spontaneously by the subjects. The paper also attempts to form a direct link between three ordinarily distinct domains: expert problem solving, arguments in the history of science, and strategies for instruction. Methods used to evaluate analogies in each of these domains will be discussed. The methods have important implications for education, because when students are presented with a "clarifying" analogous case, they may very well not understand why it is analogous to the target, unless they can convince themselves intuitively that the analogy is valid.

USE OF ANALOGIES IN EXPERT PROBLEM SOLVING

I first give two brief examples of the role of analogies in problem solving from thinking aloud studies involving experienced professionals in technical fields. The "Wheel Problem" illustrated in Fig. 1, is a question about whether one can exert a more effective uphill force parallel to the slope at the top of a wheel or at the level of the axle (as in pushing on the wheel of a covered wagon, for example). A spontaneous analogy occurs when the subject spontaneously shifts his attention to a different situation B that he believes may be structurally similar to the original problem situation A, with the intent of possibly applying findings from B to A. Subject S1 compared the wheel to the analogous case of pushing on a heavy lever hinged to the hill (Fig. 2b). He reasoned that pushing at the point higher up on the lever would require less force. He then made an inference by analogy that the wheel would be easier to push at the top (the correct answer). Apparently he used the lever to think about what was happening in the wheel.

WHEEL PUSH OR "SYSIPHUS" PROBLEM

YOU ARE GIVEN THE TASK OF ROLLING A HEAVY WHEEL UP A HILL. DOES IT TAKE MORE, LESS, OR THE SAME AMOUNT OF FORCE TO ROLL THE WHEEL WHEN YOU PUSH AT X, RATHER THAN AT Y?

ASSUME THAT YOU APPLY A FORCE PARALLEL TO THE SLOPE AT ONE OF THE TWO POINTS SHOWN, AND THAT THERE ARE NO PROBLEMS WITH POSITIONING OR GRIPPING THE WHEEL. ASSUME THAT THE WHEEL CAN BE ROLLED WITHOUT SLIPPING BY PUSHING IT AT EITHER POINT.

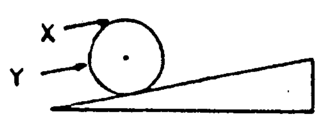

Figure 1

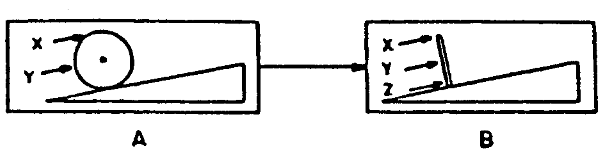

Figure 2

A second example concerns the "Spring Problem" shown in Fig. 3. Several subjects conjectured that this problem was analogous to the simpler case (Fig. 4b) of comparing long and short horizontal rods bent by equal weights hung at their ends. In most cases, a strong intuition that the longer rod bends more was used to predict the correct result that the wider spring stretches more. Thus the bending rod was used as an analogous case for thinking about the spring. That the wide spring will stretch farther seems to correspond to most people's initial intuition about the problem. However, carefully answering the question about <u>why</u> the wide spring stretches more (and explaining exactly where the restoring force of the spring comes from) is a much more difficult task.

Ten professors and advanced graduate students in technical fields were recorded while solving the spring problem in order to study the analogy generation process. They were told that the purpose of the interview was to study problem solving methods and were given instructions to solve the problem "in any way you can". After they reached an answer, subjects were asked to give an estimate of their confidence in their answer. They were then asked if there was any way they could increase their confidence, and this often led to further work on the problem. Probing by the interviewer was kept to a minimum, usually consisting of a reminder to keep talking. Occasionally the interviewer would ask for clarification of an ambiguous statement.

Some of the solutions were quite complex and took up to 90 minutes to complete. All subjects favored the (correct) answer that the wide spring would stretch farther. But the subjects varied considerably in the types of explanations they gave for their prediction. A considerable number of spontaneous analogies were observed, as described below.

<u>SPRING PROBLEM</u>

A WEIGHT IS HUNG ON A SPRING. THE ORIGINAL SPRING IS REPLACED WITH A SPRING

- MADE OF THE SAME KIND OF WIRE,
- WITH THE SAME NUMBER OF COILS,
- BUT WITH COILS THAT ARE TWICE AS WIDE IN DIAMETER.

WILL THE SPRING STRETCH FROM ITS NATURAL LENGTH, MORE, LESS, OR THE SAME AMOUNT UNDER THE SAME WEIGHT? (ASSUME THE MASS OF THE SPRING IS NEGLIGIBLE COMPARED TO THE MASS OF THE WEIGHT). WHY DO YOU THINK SO?

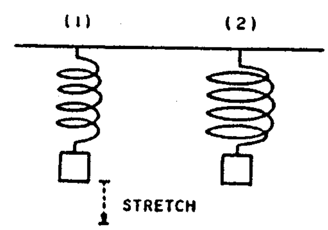

Figure 3

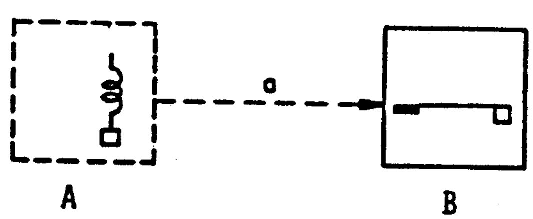

Figure 4

Number of Subjects	10
Total Number of Spontaneous Analogies Generated	38
Total Number of Significant Analogies Generated	31
Number of Subjects Generating at Least One Analogy	8
Number of Subjects Generating a Significant Analogy	7

An analogy was classified as significant if it appeared to be part of a attempt to generate or evaluate a solution, and as non-significant if it was simply mentioned as an aside or commentary. Thus there is evidence that scientifically trained individuals are capable of generating analogies during problem solving.

Subjects indicated varying degrees of certainty about the appropriateness of each proposed analogy. Sometimes they would decide that the new case was not analogous to the original problem in a useful way. In other instances further work would lead them to establish confidence in the validity of an analogy. These observations suggest that the following processes are involved in making a confident inference from a spontaneous analogy (Clement, 1981). (In this description we make a distinction between the analogous case, shown as "B" in Fig. 4, and the analogy relation, shown as the small letter "a".)

a) Generating the analogy. A representation of a situation B that is potentially analogous to A is accessed in memory or constructed. A tentative analogy relation "a" is set up between A and B (e.g. between the wheel and the lever).

b) Evaluating validity: Establishing confidence in the analogy relation. The validity of the analogy relation between A and B is examined critically and is established at a high level of confidence (confidence that the wheel works like the lever).

c) Understanding the analogous case. The subject examines and, if necessary, develops his or her understanding of the analogous case B, so that the behavior of the analogous case is well-understood, or at least predictable (the lever is well-understood).

d) Transferring findings. The subject transfers conclusions or methods from B back to A.

Steps b, c and d above were observed to occur in different orders in different solutions.

METHODS FOR EVALUATING THE ANALOGY RELATION

Methods for generating analogies (step (a) above), are discussed in Clement (1981). The remainder of this paper focuses on step (b) above. Three methods of evaluating validity will be discussed: matching key features; generating a bridging case; and using a conserving transformation. A fourth possible method, analyzing the two cases in terms of a higher level principle, is not treated here. This paper concentrates most on the second method:

generating a bridging case. Examples of evaluation methods will be drawn from a larger sample of 20 expert subjects who solved a variety of problems.

Evaluating the lever analogy for the wheel. A fairly obvious strategy for evaluating the validity of an analogy relation is to assess whether there is a structural match between cases A and B in terms of key features that are important for the behavior of the systems. This involves isolating the key features (especially higher order relationships as defined in Gentner and Gentner, 1983) in each of the cases A and B and comparing them explicitly. In the "Wheel Problem" discussed earlier subject S1 was confident that it would be easiest to move the heavy lever in Fig. 2b by pushing at point X, but he questioned whether there was a valid analogy relation between the wheel and the lever. Can one really view the wheel as a lever, given that the "fulcrum" at the bottom of the wheel is always moving and never fixed? In matching key features, he found a potential mismatch between the stationary fulcrum at the bottom of the lever and the moving fulcrum at the point of contact of the wheel that led him to doubt the analogy. Thus, matching key features is one method for evaluating the validity of analogies.

However, he also used a second, more creative method for evaluating validity. He considered a bridging case in the form of the spoked wheel without a rim shown in Fig. 5C. The spoked wheel allows one to view the original wheel as a collection of many levers. It is a bridge in the sense of being an intermediate case which shares features with both the wheel and the lever. The bridging case reduced the subject's concern about the moving fulcrum issue and raised the subject's confidence in the appropriateness of the lever model.

Presumably, this method works because it is easier to comprehend a "close" analogy than a "distant" one. The bridge divides the analogy into two small steps which are easier to comprehend than one large step. It is easier to see that the real wheel should behave like a rimless spoked wheel, and that the rimless spoked wheel should behave like a lever, than to make this inference in one step. The spoked wheel, then, is an example of a bridging case constructed by the subject in order to establish confidence in the

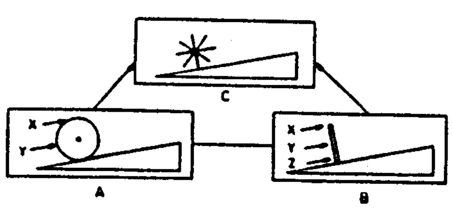

Figure 5

DOUGHNUT PROBLEM

COMPUTE THE VOLUME OF THE TORUS (DOUGHNUT) BELOW WITHOUT TAKING AN INTEGRAL. GIVE AN APPROXIMATE ANSWER IF YOU CANNOT DETERMINE AN EXACT ONE.

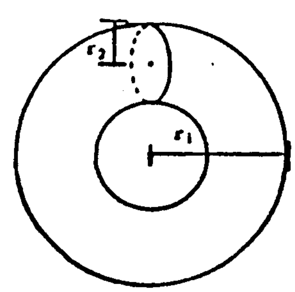

Figure 6

validity of the analogy relation between the lever and the wheel.

Bridging from doughnuts to cylinders. Another example of a bridge occurred in a solution to the mathematics problem (shown in Fig. 6) of finding the volume of a doughnut. Subject S3 first conjectured that the volume might be the same as the answer to the analogous problem of finding the volume of a cylinder (the "staightened out" doughnut). He thought the appropriate length for the cylinder would be equal to the central or "average" circumference of the torus ($2\pi(r_1-r_2)$) but was only "70% sure" of this. However, he then evaluated the plausibility of this choice by considering the bridging case of a square-shaped doughnut shown in Fig. 7. He then showed that the four sides of the square doughnut could be joined end to end to form a single long cylinder with slanted ends. He reasoned that the volume of this long cylinder would be exactly equal to its circular cross section times its central length and that therefore the appropriate length to use in the square doughnut was the average of its inner and outer perimeters. This raised his confidence in his original solution to "85%". He then reached the same conclusion for the case of a hexagonal doughnut, and this raised his confidence to "100%" for the problem. This is an example of a multiple bridge. Thus the bridging cases of a square and hexagonal doughnut helped the subject change his original conjecture about the cylinder into a firm conviction.

Bridging between rods and springs. In the spring problem subject S1 had generated the analogy of a horizontal bending rod. However, he was concerned about the apparent lack of a match between the increasing slope in a bending rod and the constant slope in a stretched spring (that a bug would experience in walking down the spring). This led him to question whether the analogy relation between the rod and the spring was valid.

An extremely successful attempt at a bridge between the case of a single coil of the spring and the bending rod analogy occurred when this subject generated the idea of a square-shaped coil in Figure 8. Visualizing the stretching of a square coil allowed him to recognize that some of the restoring forces in the spring come from twisting in the wire instead of bending — corresponding to the way in which engineering specialists view springs. This discovery of a new causal variable in the system appears to be an example of a significant scientific insight. In this case the square spring not only helped him evaluate the bending rod model, it eventually

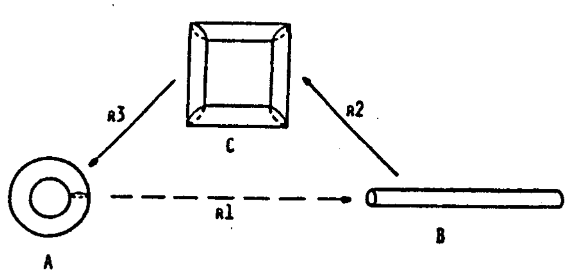

Figure 7

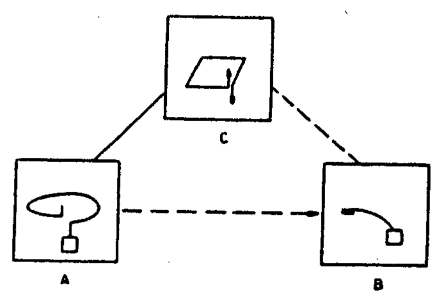

Figure 8

acquired the role of a preferred mental model which changed his conception of how springs work. This significantly increased the subject's confidence concerning the question of whether he had a good understanding of the spring.

Discussion of findings on bridging. Bridging is indicative of the recursive possibilities inherent in reasoning processes like the use of analogies. Since a bridging case is itself an analogous case, it can be described as an analogy used to evaluate a previous analogy (or, more precisely, as a second analogous case used to evaluate the analogy relation between a first analogous case and the original problem). We can summarize our view of bridging as one method for evaluating the analogy relation between a problem situation, A, and an analogous case, B, as follows:

1) The subject constructs a representation for an intermediate bridging situation C which shares important features with both A and B.

2) The subject asks whether the analogy relation between A and C is valid.

3) The subject also asks whether the analogy relation between C and B is valid with respect to the same relationships as in step 2.

4) If the subject can answer yes to both questions with high confidence, this constitutes evidence for the validity of the original analogy.[1]

Here A being analogous to C and C being analogous to B means that A is analogous to B. We can refer to this type of inference as analogical transitivity. However, it should be noted that analogical transitivity is considered a form of plausible reasoning which does not carry the force of a logical implication.

It is clear that many of the bridges discussed here are novel constructions in the sense that they are situations which the subject is unlikely to have studied or worked with before. This indicates that they are invented representations in the form of creative Gedanken experiments that have been constructed, not simply retrieved from memory. In theories of scientific discovery, hypothesis generation is ordinarily seen as a more creative process than hypothesis evaluation. However, in the case of bridging we are faced with a creative, non-empirical evaluation method which generates novel constructions. Thus we seem to have evidence for a type of "creative hypothesis evaluation" process.

ANALOGIES AND BRIDGES IN THE HISTORY OF SCIENCE

Legend has it that Galileo investigated the question of whether light bodies accelerate as rapidly as heavy bodies in an empirical manner by dropping objects from the tower of Pisa, but this legend has come under serious doubt. However it is known that Galileo and his predecessor, Benedetti, did use thought experiments like the following one to argue their side in this issue. Figure 9a shows two equal objects of one unit each being dropped while Figure 9b shows a heavier object being dropped that is equal to the two smaller objects combined. According to Aristotle the one unit objects will fall much more slowly than the larger object. Galileo claimed that they will reach the ground at nearly the same time. In saying this he was effectively proposing an analogy between cases A and B in Figure 9 to the effect that each body falls according to the same rule irrespective of its weight.

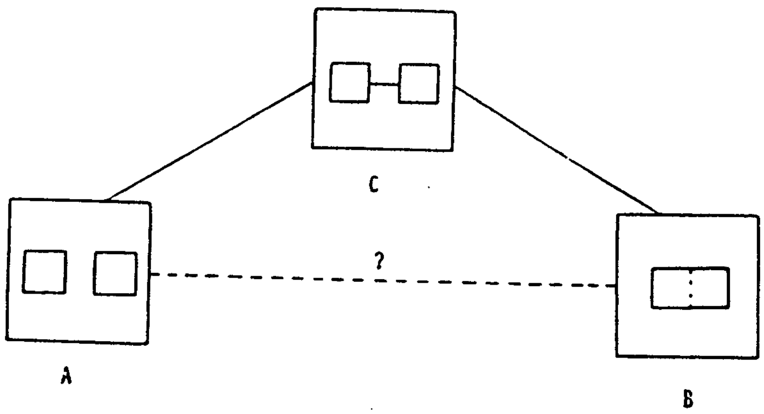

Figure 9

A marvelous bridging case used to support this analogy is the case shown in Figure 9c. The argument was first published by Benedetti (1969) and a similar argument was given by Galileo (1954). Imagine the two unit objects in A to be connected by a thin line or thread. Does the mere addition of this tiny thread, which makes the two objects become one, cause their rate of fall to increase by a large amount? Because this is implausible, the bridge argues that A and B are indeed equivalent with respect to rates of fall. In an insightful Gedanken experiment, the lightest thread can apparently make all the difference.

<u>A bridge used by Newton</u>: One of the most extraordinary scientific analogies of all time was propounded by Robert Hooke and Isaac Newton in the seventeenth century. They claimed that the moon falls toward the earth just as an everyday object (such as an apple) does. To a modern physicist, this may seem more like an obvious fact than a creative analogy, but to advocate such an idea in Newton's time was not an obvious step at all. One has only to imagine the consternation that would be produced by telling someone ignorant of science that the moon is falling.

The proposed analogy relation is represented by the dotted line in Figure 10. Essentially this conjecture says that the same causal mechanism is involved in making the moon revolve around the earth and making an apple fall. A multiple bridge used by Newton to support this analogy in his <u>Principia</u> is shown in Figure 10c. This is the idea of a cannonball fired faster and faster until it enters into orbit around the earth — a premonition of modern rocketry.

students who had not taken physics, we found that 75% of the students did not believe that a table pushes up on a book. The physicist believes that since all materials are deformable, the table will deform, acting like a spring, and provide an upward force just large enough to balance the weight of the books. The difficulty here is not just one of students lacking a particular fact. Pilot tutoring interviews and class discussions indicate that many students express disbelief in the physicist's view and have a deeply held belief that stationary objects are rigid barriers which cannot exert a force on their own. On the other hand, 96% of the students did believe that a spring pushes up when it is compressed with one's hand. The contrast between these results is interesting since the physicist views these two situations as essentially identical. diSessa (1983) refers to the concept of springiness as a "phenomenological primitive" and discusses the evolution of the individual's intuitions that is needed to become skilled in physics.

An interesting conjecture is that the right bridge may help a student believe in the validity of an analogy proposed in instruction. The following instructional strategy, suggested by our analysis of experts, attempts to build basic conceptual models that are grounded in intuitions the student already has.

Teaching Strategy:

(1) Draw out the conceptual difficulty in a concrete target situation A where the student makes a statement that is in conflict with accepted theory.
(2) Search for a simple analogous situation B where the student has a reliable intuition that is in agreement with physical theory and that is a relevant starting point for the area of difficulty.
(3) Stimulate discussion and encourage students to look for and match key features by asking them to describe how A and B are alike and different.
(4) Students often will still not believe that A is analogous to B. Finding an apt bridge that appeals to the students' intuitions is an important technique for remedying this.

The hand on the spring situation is a potential starting point for instruction since it draws out a correct intuition from students. For this reason we call it an "anchor". Figure 11 shows multiple bridging cases used to help convince students that the analogy between the "hand on the spring" anchor and the targeted "book on the table" case is valid. We have attempted to use this strategy in tutoring studies and class sessions with high school students and find that: 1) Students readily understand the anchoring case; 2) many students indeed do not initially believe that the anchor and the target cases are analogous; 3) the bridging cases sparked an unusual amount of argument and constructive thinking in class discussions; and 4) the bridging cases helped many students to believe in the analogy. We are in the process of collecting data to confirm these preliminary results. Minstrell (1982) has had some success with a related approach.

In this method then, we are trying to ground the student's understanding on a physical intuition about a familiar case. The strategy is to build on and extend the intuition by using analogical reasoning. The problem is that students will not be able to understand how other cases can possibly be analogous to the familiar case. Presenting the right analogy is not enough — the student must also come to believe in the validity of the analogy. The technique of bridging by using chains of analogies combined with discussion to encourage active thinking, appears to be helpful for this purpose.

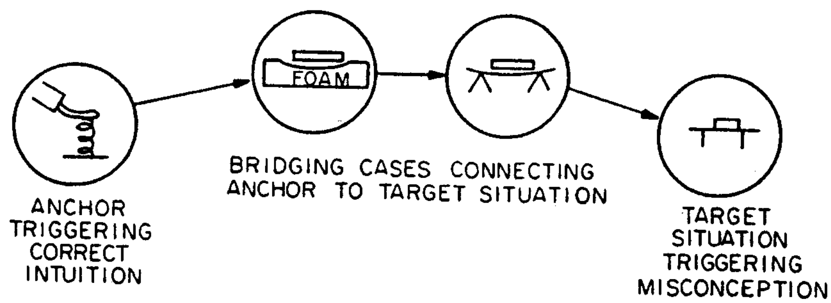

Figure 11

CONCLUSION

The ability to evaluate the validity of analogies appears to play as important a role in insightful problem solutions as the ability to generate analogies. Three methods for evaluating validity were discussed: matching key features, bridging, and finding a conserving transformation. These patterns of reasoning were observed to occur in different problem contexts in both science and mathematics. These empirical studies have allowed us to construct a more detailed theory of how analogies can be used effectively in instruction. A method for evaluating analogies used by expert scientists as well as by Newton and by Galileo and his predecessors appears to be a promising strategy for helping students overcome misconceptions in science.

Note 1. The subject may also use bridging recursively by bridging again between C and B or C and A as in the case of the square and hexagonal doughnuts.

*Preparation of this paper was supported by a grant from the National Science Foundation #MDR8470579.

References

Benedetti, G. (1969) Diversarum speculationum mathematicarum (1585), in S. Drake and I. Drabkin, Mechanics in sixteenth-century Italy (p. 206), Madison, Wisconsin: U. of Wisconsin Press.

Clement, J. (1981) "Analogy generation in scientific problem solving". Proceedings of the Third Annual Meeting of the Cognitive Science Society, Berkeley, California.

Clement, J. (1982). "Analogical reasoning patterns in expert problem solving". Proceedings of the Fourth Annual Meeting of the Cognitive Science Society (pp. 79-81). Ann Arbor, Michigan.

Clement, J. (1983). A conceptual model discussed by Galileo and used intuitively by physics students. In A. Stevens and D. Gentner (Eds.), Mental models (pp. 325-340), New Jersey: Lawrence Erlbaum Associates.

diSessa, A. (1983). Phenomenology and the evolution of intuition. In A. Stevens and D. Gentner (Eds.), Mental models (pp. 15-33), New Jersey: Lawrence Erlbaum Associates.

Galileo, G. (1954) Two New Sciences (1638). Translated by H. Crew and A. de Salvio (p. 62), New York: Dover.

Gentner, D., & Gentner, D. (1983). Flowing waters or teeming crowds: mental models of electricity. In A.L. Stevens and D. Gentner (Eds.), Mental models (pp. 99-101), New Jersey: Lawrence Erlbaum Associates.

Gick M. and Holyoak, K. (1980). Analogical problem solving. Cognitive Psychology, 12, 306 - 355.

Minstrell, J. (1982). Explaining the 'at rest' condition of an object. The Physics Teacher, 20, p. 10.

Rummelhart, D. and Norman, D. (1980). Analogical processes in learning. In J.R. Anderson (Ed.), Cognitive skills and their acquisition. Hillsdale, N.J.: Lawrence Erlbaum Associates.