When one encounters a problem that one has no adequate way of representing, a new mental model for the problem may have to be found. When a possible model is found, an important next step is to evaluate the validity of the model. In this document, examples of subjects finding and evaluating mental models used as problem representations are discussed in several case studies which depict expert scientists solving problems. This paper focuses on a non-deductive strategy called "bridging" that is used to evaluate the validity of a mental model and that has been observed in solutions to both science and mathematics problems. In constructing a bridge, the subject creates an intermediate case that is seen as "in between" the proposed model and the problem situation because it shares important features of both. A bridge can help the subject confirm or deny the validity of the analogy relation between the model and the problem. It is suggested that the bridging strategy observed in experts can be used to help students construct and refine new mental models. Numerous figures are provided. (Author/TW)
A METHOD EXPERTS USE TO EVALUATE THE VALIDITY OF MODELS
USED AS PROBLEM REPRESENTATIONS IN SCIENCE AND MATHEMATICS

John Clement

Department of Physics and Astronomy
University of Massachusetts
Amherst, MA 01003

March 26, 1985


Preparation of this paper was supported in part by a grant from the National Science Foundation program for Research in Science Education, SED80-16567 and by a grant from the EXXON Education Foundation.
Abstract

When one encounters a problem that one has no adequate way of representing, one may have to find a new mental model for the problem. Once a possible model is found, an important next step is to evaluate the validity of the model. Examples of subjects finding and evaluating mental models used as problem representations are discussed in several case studies of expert scientists solving problems. In particular, this paper focuses on a creative, non-deductive strategy called bridging that is used to evaluate the validity of a mental model and that has been observed in solutions to both science and mathematics problems. In constructing a bridge, the subject creates an intermediate case that is seen as "in between" the proposed model and the problem situation because it shares important features of both. A bridge can help the subject confirm or deny the validity of the analogy relation between the model and the problem. A bridge works by reducing the conceptual distance across which the analogy is being made. Most of the bridges observed appeared to be novel inventions.

One of the motives for this study was to see whether observations of experts struggling to resolve conceptual difficulties of their own could inform attempts to help students resolve conceptual difficulties. It is conjectured that the bridging strategy observed in experts can be used to help students construct and refine new mental models.
INTRODUCTION

In comparison to our knowledge of strategies for searching within a problem representation, much less is known about how people construct or find a problem representation in creative problem solving. When one encounters a problem that one has no adequate established way of representing, one may have to find a new mental model for the problem situation. Once a possible model is found, an important next step is to evaluate the validity of the model. Several case studies of mental models being proposed and evaluated by different subjects working on different problems will be presented here.

A number of authors have discussed the role of mental models in problem solving and learning, including de Sessa (1983), Gentner and Gentner (1983), Larkin (1983), Norman (1983), Williams, Hollan and Stevens (1983), and Collins, (1983). This paper discusses some specific findings relevant to the following general questions:

(1) How do expert subjects attack problems when they do not have an adequate understanding of the situation described in the problem?
(2) How do subjects use qualitative mental models (as opposed to mathematical formalisms) to solve problems?
(3) When the applicability of a model to a given problem is questionable, how does one establish the validity of the model for the problem?
(4) What are the implications of findings in this area for teaching?

METHOD

Four case study solutions from three expert subjects solving physics and mathematics problems will be presented. The three problems discussed in this
paper are shown in Figs. 1, 2, and 8. The subjects were professors with graduate training in mathematics. Problem protocols were transcribed from video or audio tapes of the interviews.

The subjects were told that the purpose of the interview was to study problem solving methods. They were given instructions to solve the problem "in any way that you can," and were asked to give a rough estimate of confidence in their answers. 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.

This paper will not attempt to present a nomothetic description of behavior frequencies. Rather, the purpose is to identify an important reasoning strategy that occurred across different subjects and problems and to propose an initial description of its form and function. Our knowledge of creative problem solving strategies used by experts is still in a very primitive state. We lack basic process models in this area. In fact, we lack even a basic set of well defined observational and theoretical concepts in this area. Developing these concepts is an important task, and is a major goal of the present set of case studies.

FINDINGS

The Use of Analogue Mental Models

The problems used in this study were chosen so that many of the subjects: (1) lacked an established formal procedure for solving the problem; and (2) lacked a sufficient understanding of relationships in the problem situation. Thus, the first step in their solutions was to develop a viable mental model for the problem situation. In the physics problems these were qualitative
physical models, and in the math problems they were geometric models.

A lever as a mental model for a wheel. An example of a qualitative physical model occurs in the "Wheel Problem" illustrated in Fig. 1, a question about whether one can exert a more effective uphill force (parallel to the ground) on a wheel at the top or at the level of the axle (as in pushing on the wheel of a covered wagon, for example). Subject S1 compared the wheel to the analogous case of pushing on a heavy lever hinged to the hill (Fig. 3b). 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 as a mental model for thinking about what was happening in the wheel.

By saying that someone is using a "mental model" for a situation like the wheel problem, we mean that they have a cognitive structure or schema that allows them to make predictions and formulate explanations about the problem. As used here, a mental model is not necessarily identical with a naive, practical representation of the target system. The model may be analogous to the target or more abstract than the target. It may also represent hidden features such as forces or molecules that are not directly observable in the target. A model can also be dynamic if it involves moving elements, and this can be the case even when the target system is ordinarily thought of as stationary (e.g., molecules in solids).

We will use the term "established scientific model" for a mental model which has become accepted as a useful part of a scientific theory. The mental models discussed in this paper appear not to be established scientific models for the subjects, but rather are conjectured models being used with the problem situation for the first time. The validity of the models will be discussed in a later section.

A rod as a model for a spring. A second example concerns the "Spring
Problem shown in Fig. 2. Essentially, the problem is to decide whether a wide spring will stretch more than a narrow spring, other factors being equal. Several subjects have conjectured that this problem might be analogous to the simpler case in Fig. 4b of comparing long and short horizontal rods bent by equal weights hung at their ends (Clement 1983b). Usually, 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 initial mental model for thinking about the spring.

A pulley as a model for the wheel. Subject S2 dealt with the wheel problem in a different way. He first thought about the extreme case of rolling the wheel up an extremely steep hill that was almost vertical (shown in Fig. 5a) and was trying to decide intuitively whether it would be harder to push on the edge of the wheel or on the middle. He also used the well-known physicist's technique of adding gear teeth to the wheel and the cliff so as to prevent slipping, as required by the problem:

101 S: Suppose it were tilted steeply and you did that; so steep as to be almost vertical. (Draws Fig. 5a).

103 S: It seems like it would skid out from under you the other way (down). This would get away from you here. Let's assume it's gear toothed and that it won't slip or that the friction is strong enough here that it's not going to slip under you.

This extreme case then inspires him to make an analogy to a pulley:

105 S: ...So if you do that...it now er, what it feels like is the weight of it--; pretty close to parallel with what you've got if you go well, [to] a complete vertical. It now begins to feel like a pulley...

106 S: ...And you're over here pulling like this. That feels like you're on the outside of a pulley pulling up. (Draws Fig 5b).

107 S: Oh, it says you have to push and not pull, but I'll ignore that for now. Let's see how that feels--almost vertical.

108 S: [In] this new point of view, it feels like working at X [on the edge of the wheel] is better [than at Y].
In this case the pulley model has led to a prediction that it would be easier to push the wheel on the outside.

**Summary of findings on mental models.** Based on these examples, we can make the following interpretations:

1. The subjects use qualitative reasoning strategies that appear to be non-deductive inferences from physical intuitions rather than deductive arguments from equations or formal principles.

2. The extensive use of drawings provides evidence that spatial representations are being used.

3. In each of the above examples the subject was apparently faced with an insufficient understanding of the problem situation. In information processing terms they did not have an adequate problem representation. Their knowledge of causal or other essential factors in the system was not sufficient to yield a prediction about the effects of operators in the problem. They therefore had to try to increase their qualitative understanding of the problem situation. The first major subproblem for these subjects was: "What is a familiar model I know something about that applies to this problem?"

4. The mental models they use can also be viewed as analogous cases, just as many established scientific models can be viewed as depending on an analogy (Hesse, 1966; Campbell, 1957). In each example, the model B is an analogous case in the sense that one or more features ordinarily assumed to be fixed in A are different in B, and yet the subject treats structural or functional relationships as equivalent in A and B. Finding a new problem representation in the form of a mental model involves making an analogy between the problem and the model in these cases. Thus, we can refer to them as analogue mental models. Although established scientific models are often thought of as being at a higher level of abstraction than an analogy, this may
not always be the case for an initial mental model of a problem situation. Each subject finds the model to be suggestive but is uncertain as to whether the model is a valid one for the problem in question. Their uncertainties about the model's validity and their strategies for resolving this uncertainty are examined in the next section.

**Evaluation of Mental Models via Bridging**

Model evaluation is an important process, because as soon as a candidate model is proposed, the question arises: "Does the problem really fall within the domain of applicability of the model?" This is because such domains are often fuzzy and not well-defined (Norman, 1983). Another way to ask this is: "Does the relationship between this model and the problem situation constitute a valid analogy?" Thus, three essential processes in using such a model are: (a) accessing (or constructing) a well-understood model; (b) evaluating the validity of the model for this problem; (c) applying the model to the problem.

Methods for generating analogies that can be used as mental models (step (a) above), are discussed in Clement (1983a). The remainder of this paper discusses "bridging", a creative method subjects can use to evaluate a model (step (b) above). Five examples of "bridges" constructed by experts will be examined here.

**Evaluating models for the wheel.** In the "Wheel Problem" the subject S1 discussed earlier was confident that it would be easiest to move the heavy lever in Fig. 3 by pushing at point X, but he questioned whether there was a valid analogy relationship between the case of the wheel and the model of 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? A bridge generated by this subject is the spoked wheel without a rim shown in fig. 6C. The spoked wheel allows one to view the wheel as a collection of many levers. It
is a bridge in the sense of being an "in between" case which shares features with both the wheel and the lever. This bridge 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. The spoked wheel, then, is an example of a bridging case constructed by the subject in order to confirm the validity of the lever as a mental model for the wheel.1

Another bridge in the wheel problem. Recall that S2 used the mental model of a pulley to predict the correct answer for the case of rolling the wheel up an almost vertical wall. However, in the following passage he goes on to evaluate and criticize the validity of the pulley model -- that is, he questions the validity of the analogy relation between the pulley and the wheel:

155 S: The pulley analogy may be totally wrong, and misleading me because of the way we think of a pulley going around under and holding it. But we have no problem in gripping here at the edge you say. Then we have a good grip for whatever reason.

157 S: ...my problem with the pulley is, I don't feel like it's a comparable experiment. That is, somehow this rope wrapping around here and pulling doesn't feel to me necessarily like the problem which is stated here, which is pushing on the outside of a wheel.

159 S: Push on the outside; and put a rope around it and use a pulley there. Is that the same problem? The answer is I don't know.

Apparently the subject is quite unsure of the validity of the analogy between the pulley and the wheel problem. In the following section, he generates a bridge in order to help him to evaluate the analogy relation further. The bridge in this case takes the form of a rope tied to the wheel at a point opposite to the wall, as shown in Fig. 7. This case seems to share some characteristics with the pulley and some with the vertical wheel problem, so
it is an intermediate case.

162 S: [Looking at the wheel on a vertical wall] My instinct tells me [push at] X again but that er, but again it’s in terms of a pull and not a push. I’d have to get a grip. Assuming...we attach a rope to one of the teeth. Now it becomes more like the pulley problem which I was thinking before. (draws Fig 7c).

163 S: ...a lot easier than getting down here behind it [at "Y"] and pushing. Why? because of that coupling pulley effect. It seems like it would be a lot easier to hold it here [at "X"] for a few minutes than it would be to get behind it or even to attach a rope here [at "Y"] and--; yeah, my confidence here is much higher now, that it’s right.

164 I: Can you recall what made you think of the rope on the tooth?

167 S: Let me think. It was something about holding it steady instead of trying to pull it up.

169 S: You put a rope right there and the tooth down here holding on will play the same role as the rope here. And so the pull; I don’t know. It just felt right.

175 S: I was trying to...I did not want that feeling of [the rope] going all the way around [the pulley]. I wanted to use the internal strength of the wheel.

178 I: OK. And do you have a sense of where your increased confidence is coming from? Is it this example?

179 S: It’s the pulley analogy starting to feel right.

183 S: I guess I decided to take the pulley argument more seriously and I just had to throw away the rope. I had to abandon the part of the pulley argument that had been bothering me. All along, which is the rope going around...

185 S: ...and decide no, there's no problem with that. You just attach it there. Um, put a little eyelet in a little tooth; I mean a little thing there, just tie it up.

In lines 163 and 179 above, we have evidence that the bridge has increased the subject's confidence in the pulley analogy. The pulley appears to be gaining credibility for the subject as a mental model for the problem situation. There is also evidence in line 162 that the bridge of the rope tied to the wheel is an intermediate case for the subject in that he says that it is more like the pulley problem than the original problem was. The subject seems to
be worried about the fact that in the original problem one was applying a
force only at one point of the wheel whereas in the pulley model, the rope may
be applying force to the wheel at every point where it touches the wheel on
its circumference. The bridge helps him resolve this criticism. Such
protocols serve to illustrate an alternating cycle of productive (imaginative
and divergent) processes and evaluative (critical and convergent) processes in
scientific thinking.

Bridging from doughnuts. Another example of a bridge occurred in a
solution to the mathematics problem (shown in Fig. 8) 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 "straightened out" doughnut). He thought the appropriate length for the
cylinder would be equal to the central or "average" circumference of the torus
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. 9c. This is a doughnut made of four straight cylinders, where
the small cross-section of the doughnut is a circle and the outside and inside
perimeters of the doughnut are squares. He then showed that the four sides of
the square doughnut could be reassembled into a long cylinder with slanted
ends. He reasoned that the volume of this horizontal cylinder would be its
perpendicular cross section times its central length and that the appropriate
length to use in the square doughnut was the average of its inner and outer
perimeters. This raised his confidence in his 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.
Further examples of bridging: the zig-zag and square springs. In the spring problem subject S1 had generated the model of a horizontal bending rod. However, he was concerned about the apparent lack of a match between the non-constant slope a bug would experience walking down a bending rod and the constant slope the bug would experience walking down a stretched spring. This led him to question whether the rod was a valid model for the spring. Apparently, in order to help evaluate the analogy relation between the spring and the bending rod, he constructed a bridge in the form of the "zig-zag spring" shown in fig 10c. Unfortunately the zig-zag spring proved to be indecisive for the purpose of evaluating the analogy between the spring and the model of the bending rod, and at this point the subject dropped the idea. Presumably this happened because he still could not reconcile the bending going on in elements of the zig-zag spring with the lack of change in slope in the original helical spring (i.e., he was unable to confirm link 3 in fig. 10). Thus the zig-zag spring is an example of an unsuccessful bridge which failed to yield either a confirmation or a rejection of the original bending rod model for the spring.

A second extremely successful attempt at a bridge between the case of a single coil of the spring and the bending rod model occurred when this subject generated the idea of a square-shaped coil. 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—a major breakthrough in his solution which corresponds to the way in which engineering specialists view springs. In this case the square spring eventually acquired the role of a preferred mental model which changed his conception of how springs work. This significantly increased the subject's confidence concerning whether he had a good understanding of the spring.
We can now summarize our view of the bridging process as follows:

Subjects using a bridge start from the following context: a mental model is proposed as possibly applicable to the problem situation A. Typically the situation described in the model is considerably simpler than the problem situation, as in the case of using the lever as a model for the wheel. The model may yield a prediction for the problem but the problem solver is unsure that the model is valid for the problem.

They then employ the following strategy:

(1) The subject constructs a representation for an intermediate bridging situation C which in his view shares important features with both the problem situation, A, and the model, B.

(2) The subject asks whether C and A are analogous by examining whether he believes significant structural features are equivalent in C and A.

(3) The subject also asks this question about C and B.

(4) If the answer to both questions is yes, this constitutes evidence for the validity of the model.

The case studies described in this paper are evidence that the above pattern of reasoning can occur in different contexts.

The following additional features of bridges are noted: (1) A conception, Cc, can be thought of as a "qualitative interpolation" between the original situation A and a proposed mental model B for that situation. A bridge is used for the task of evaluating the validity of the analogy relation between A and B. The bridge makes it easier to do this by changing this task to evaluating analogy relations between two pairs of cases that are in some sense "closer" to each other. Thus the bridge reduces the "conceptual distance" across which analogies are being made. (2) Bridges are not deductive arguments, but experts have been observed to use them as a powerful intuitive argument. The reasoning involved here appears to be not at a formal level but
at a phenomenological level that appeals to the subject's physical or spatial intuitions. (3) Many bridges 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 not simply been retrieved from memory. (4) As stated earlier, the models used initially by the subjects can be viewed as analogous cases. However, the bridging cases discussed in this paper can also be viewed as analogous cases. Thus a bridge can be thought of as an analogy used to evaluate a previous analogy.

In summary, bridging is a creative non-deductive method for evaluating mental models. Evidence for bridging has been observed in experts solving both physics and mathematics problems.

EDUCATIONAL IMPLICATIONS

Bridging may be an important learning process because it provides a way to build a firm link between ideas that are only tentatively connected. Although I will not present data here on the use of bridging in instruction, I will propose a strategy for its use that may have value in science teaching.

One of the motives for this study was a hope that observations of experts resolving conceptual difficulties of their own could inform attempts to help students resolve their conceptual difficulties. Recently, a fairly large literature has appeared on the problem of persistent misconceptions in science and mathematics (e.g., Helm and Novak, 1983; Clement, 1983; McDermott, 1983). diSessa (1983) discusses the evolution of the individual's intuitions that is needed to become skilled in physics. Typically in these areas the standard curriculum jumps too quickly into formulas and formal arguments without paying
enough attention to the development of basic qualitative models. The result is that the principles do not make sense intuitively to the student.

An interesting conjecture is that the right bridge may help a student see why a physical model B is a useful way of viewing phenomenon A. The following strategy, suggested by our observations 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 problem 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 subject has a reliable intuition and that is a relevant starting point for the area of difficulty.
3. Students often will not believe that A is analogous to B. Find a way to help the student see the analogy. Finding an apt bridge that appeals to the student's intuitions is one important technique for doing this.
4. Refine B and relate it to other situations in a similar way, so that it becomes a general model.

Data from tutoring experiments employing this strategy as one of the methods used to help students overcome misconceptions in physics will be presented in a future paper.
CONCLUSION

Several case studies have been presented drawn from protocols of expert scientists solving problems. We have focused on a strategy used by experts when they cannot adequately represent a problem situation and are attempting to find a viable mental model for the problem.

In the case of analogue models this depends on setting up an analogy relation between the problem and the model. Inventing a bridging case is a creative, non-deductive method the subject can use to evaluate the validity of such an analogy relation. In most cases this leads to the acceptance, rejection, or modification of the model. This same pattern of reasoning can occur in different contexts in both science and mathematics. The reasoning pattern suggests a teaching strategy for grounding ideas in physical intuitions that should be useful in helping students to construct and refine new mental models.

Bridges appear to be an important tool for stretching the domain of applicability of an intuitively grounded model to a new situation, i.e. for making the model more general and powerful. Furthermore bridging seems to work at the level of a person's intuitions, not just at a formal symbolic level. It may therefore be an important tool for developing and refining one's physical and mathematical intuitions.
(1) A recorder was not available for this interview and data consists of careful notes made by the interviewer or the subject.

(2) Or the subject may use bridging recursively by bridging again between C and B or C and A (as in the case of the square and hexagonal doughnuts).
References


**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.

![Diagram](image)

*Fig. 1*
SPRING PROBLEM

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 2

Figure 3

Figure 4
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.